Studies on free energy calculations. II. A theoretical approach to molecular solvation

Haluk Resat and Mihaly Mezei
Department of Biophysics and Physiology, Mount Sinai School of Medicine, CUNY, New York, New York 10029-6574

(Received 22 February 1994; accepted 23 June 1994)

Using the concepts of scaled particle theory, an analytical theory is developed to investigate the limiting behavior of solvation free energies at the particle creation limit. The new theory directly incorporates the weakly attractive, dispersion interaction terms into the analytical calculations. For neutral molecular systems, the effects of longer ranged electrostatic interactions are also incorporated, albeit in an *ad hoc* way, and the validity of the utilized assumptions are then demonstrated with numerical examples. It is shown that it is possible to blend the numerical and analytical methods to increase the reliability of quantitative results, and, at the same time, to achieve savings on computational expenditure for certain types of calculations. Different methods of performing the thermodynamic integration in solvation free energy calculations are also compared. Studied examples clearly show the importance of proper treatment of the divergence at the particle creation limit in obtaining quantitatively reliable results for the solvation free energies.

I. INTRODUCTION

Understanding the behavior of many chemical systems requires the determination of the chemical equilibria in solution which can be calculated if the free energy differences between the involved chemical states are known. For this reason, chemical equilibrium studies have mainly concentrated on the calculation of free energy differences.[1–8] Examples of systems investigated include the free energy difference between the two conformations of the same solute, or the free energy change associated with a transmutation, or the free energy of oxidation or activation. Free energy of solvation in which an ideal gas particle is transformed into a fully interacting particle in solution also belongs to this category. Calculation of the solvation free energies is particularly important, because, for relatively weak attractive systems, it is related to the thermodynamics of hydrophobic solvation, an essential but not so well understood subject in biological sciences.[9]

There have been numerous attempts to theoretically calculate solvation free energies using both numerical and analytical methods. For example, analytical scaled particle theory (SPT)[10–12] has often been used in studying the hydrophobic solvation.[10–17] In the original implementation of the SPT, the solute and the solvent molecules were represented as hard-sphere particles. In later applications, a perturbation treatment was used to allow for realistic repulsive interactions between the solute and the solvent particles,[15] but the weakly attractive dispersive interactions were not incorporated. Another often used analytical approach which is more empirical in character compared to the SPT is based on solvent accessible surface areas.[18–20] Also, some approximation schemes have been introduced for including the effects of the weakly attractive solute–solvent interactions to the solvation free energies.[11,21]

Solvation free energies can, in principle, be calculated for any system using numerical molecular simulations. But, it is now well established that the molecular simulation approaches are subject to numerical uncertainties which arise from the possible undersampling of the phase space. This is especially true for the class of solvation free energy calculations done in a way that a null point solute particle is allowed to gradually grow until it reaches the desired size and the shape of the studied solute. In this gradual growth approach to numerical calculation of the solvation free energies, the most serious of the encountered problems arises at the limit of creating a cavity from an uninteracting point particle. This is the limit when the size of the gradually created solute is considerably smaller than a cavity which can accommodate the full sized solute. Availability of an analytical method at this small cavity limit would make it possible to combine the analytical and numerical methods. Use of a combined approach would increase the reliability and the quantitative accuracy of the calculated solvation free energies.

The issue of eliminating the encountered problems at the very small solute limit from the numerical calculations is the subject of this report. Using the concepts of SPT, we first develop an analytical formula for the limiting value of the solvation free energy of a weakly attractive solute molecule. In the second part, we combine the new analytical approach with numerical simulation methods to study the solvation properties of some model solutes. The robustness and the reliability of this combined approach is demonstrated by several examples.

The layout of this report is as follows. In Sec. II, we first present the theory of creating a Lennard-Jones particle in a molecular solvent, and the theory is then generalized to molecular solutes. Section III presents the details of the numerical calculations employed in this study. Section IV develops and compares the different schemes for combining the analytical and the numerical simulation methods in calculating the solvation free energies. It is followed by the results presented in Secs. V and VI. Section V deals with the limiting behavior of solvation of various solutes. Numerical and analytical approaches are contrasted, and the convergence characteristics of molecular simulations and the applicability of

0021-9606/94/101(7)/6126/15/$6.00

the new analytical theory are discussed. Section VI investigates the appropriateness of the use of Gaussian quadratures in the thermodynamic integration (TI) method of calculating the solvation free energies. Section VII summarizes the important aspects and results of this work, and establishes the connection to our previous work, Ref. 7.

II. THEORY

A. Single site solute

1. Free energy of cavity formation: Perturbation treatment

The main aim of this report is to calculate an analytical formula for the solvation free energies at the limit when the created solute is very small. At this limit, the interaction between the solute and the solvent particles in the Hamiltonian can be presented as a perturbing potential energy term.[15] Separating the Hamiltonian in this way, the Helmholtz free energy of creating a cavity in solution, ΔA, can be expressed in terms of the following rigorous expression:[22]

$$\beta\Delta A = -\ln\langle\exp(-\beta V)\rangle_0. \tag{1}$$

Here $\beta = 1/k_BT$ is the inverse of the temperature expressed in energy units, and $\langle\cdots\rangle_0$ corresponds to a canonical ensemble average consisting of pure solvent at constant volume, number density, and temperature. V is the solute–solvent interaction potential energy (throughout this report, V and v, respectively, correspond to solute–solvent interaction potential and to volume, with subscripts specifying certain contributions). In a wide range of studies, an interaction site model (ISM) representation[23] is used to accommodate the appropriate solute–solvent and solvent–solvent intermolecular interactions. Within ISM representation, and for a Lennard-Jones-type interaction, the solute–solvent interaction potential is expressed as

$$V(\lambda) = \sum_{i,j} V_{ij}(\lambda) = \lambda^k \sum_{i,j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right], \tag{2}$$

where the sum j goes over all distinct sites of all the solvent molecules. Similarly, the sum i is over the distinct sites of the solute molecule. λ is the so-called *coupling parameter* and as it is changed from 0 to 1, the null solute particle gets transformed into a fully interacting solute. Therefore the size (or the radius) of the solute particle is governed by the value of the coupling parameter. The exponent k is an arbitrary positive exponent included for convenience in studying the asymptotic behavior of the free energy change curve as the solute particle is created.

For a single site spherical solute, the required ensemble average appearing in Eq. (1) may be expressed as

$$\langle e^{-\beta V}\rangle_0 = \sum_i \int \mathbf{d}^3\mathbf{r}_i \int d\bar{\Omega}_i\, e^{-\beta V} p_{i,0}(\mathbf{r}_i, \Omega_i), \tag{3}$$

where $d\bar{\Omega}_i = d\Omega_i/\bar{\Omega}$ ($\bar{\Omega} \equiv \int d\Omega_i = 8\pi^2$ is the normalization constant over the Euler angle integrals). If the solute is located at the origin, $p_{i,0}$ $(\mathbf{r}_i, \Omega_i)$ in Eq. (3) is the normalized probability distribution function for having the ith solvent molecule with its center of mass at $\mathbf{r}_i$ with an orientation Ω_i in terms of its Euler angles. Since $\Sigma_i p_{i,0}(\mathbf{r}_i, \Omega_i)$ is normalized (number of solvent molecules is conserved), the above equation may also be written as

$$\langle e^{-\beta V}\rangle_0 = 1 + \sum_i \int \mathbf{d}^3\mathbf{r}_i \int d\bar{\Omega}_i [e^{-\beta V} - 1] p_{i,0}(r_i, \Omega_i). \tag{4}$$

Let us suppose that the cavity created by the solute particle is small enough that the solute can only interact with a single site of one of the solvent molecules at a time. Then, since it is the probability distribution function with respect to *pure* solvent state, we have[15] $p_{i,0}(\mathbf{r}_i, \Omega_i) = p_0(\mathbf{r}, \Omega) = \rho/N_s$. Here ρ is the number density of solvent molecules, and N_s is the number of solvent molecules. If the longest ranged pair interaction decays faster than r^{-3}, for ISMs, integrating Eq. (4) by parts results in a more convenient form:[24]

$$\langle e^{-\beta V}\rangle_0 = 1 + \frac{\rho}{3} \int \mathbf{d}^3\mathbf{r}\, r\, \frac{\partial \beta V}{\partial r} e^{-\beta V}. \tag{5}$$

The above integral has the form of a second virial coefficient and can be done analytically for some forms of the interaction potential.

Using this approach, Postma *et al.* have derived the necessary expressions for a repulsive r^{-12}-type potential between the solute and the solvent molecules (soft sphere solute).[15] To see the effect of the inclusion of the longer ranged weak attractive interaction, we generalize their derivation to a 6-12-type potential which is a more general form of Lennard-Jones-type potential (the following derivation is quite general, and the C_6 parameters can have negative values as well). Let us rewrite the solute–solvent interaction potential energy, Eq. (2), between a single site solute and a multisite solvent molecule as

$$\beta V = \sum_j \lambda^{k_{12}} \frac{C'_{j,12}}{r_j^{12}} - \lambda^{k_6} \frac{C'_{j,6}}{r_j^{6}}, \tag{6}$$

where summation over j goes over the distinct solvent sites, and $C'_{j,m} \equiv 4\beta\epsilon_j\sigma_j^m$ (for the rest of this report, script m is equal either to 6 or to 12, and stands, respectively, for r^{-6} or r^{-12} terms). In Eq. (6), to make it more general, separate exponents for the coupling parameter are included. This generalization would be needed when a polynomial path thermodynamic integration is employed.[7,25] Notice that, because of the assumption that the solute particle can only interact with one site of a solvent molecule at a time, the ensemble averages with respect to the pure solvent state require only the knowledge of the singlet solvent distribution function.[10] The singlet solvent distribution, as stated above, is uniform and corresponds to a constant $p_{i,0}(\mathbf{r}_i, \Omega_i)$, i.e., without any spatial or angular dependence, which makes further simplification possible. Therefore, Eq. (6) may be written in a mathematically more convenient form as

$$\beta V = \lambda^{k_{12}} \frac{C_{12}}{r^{12}} - \lambda^{k_6} \frac{C_6}{r^6}, \tag{7}$$

with $C_m \equiv 4\beta\Sigma_j\epsilon_j\sigma_j^m = \Sigma_j C'_{j,m}$. For positive C_6, one can similarly define the *molecular* solute–solvent Lennard-Jones interaction parameters ϵ_{mol} and σ_{mol} as

$$4\beta\epsilon_{\mathrm{mol}}\sigma^{m}_{\mathrm{mol}}\equiv C_m . \tag{8}$$

Before proceeding with performing the integrals, let us define an *effective cavity radius*, r_c, which would also set a length scale for the problem. One obvious choice for r_c would be to use the thermal fluctuations of the system,[15] thus r_c may be defined using the relation $\beta V\cong 1$. Rather than using this definition, we use only the shortest ranged term in the interaction potential V_{12} and simply set $\beta V_{12}=1$. This results in a much simpler expression for the cavity radius:[26]

$$r_c^{12}\equiv C_{12}\lambda^{k_{12}}. \tag{9}$$

We would like to point out that the cavity radius always factors out of the final expressions, and is defined only to introduce a length scale to the problem with the aim of conceptual simplification.

One proceeds by substituting the partial derivative of Eq. (7) into Eq. (5). The resulting integral can be expressed in terms of the degenerate hypergeometric and gamma functions.[27,28] Using the polynomial expansion of the degenerate hypergeometric functions for small argument, the expression for the free energy of cavity formation can be written as a series expansion in terms of the coupling parameter (details of the derivation are given in Appendix A). Thus, the required ensemble average appearing in the free energy of cavity formation, Eq. (1), can be expressed as [Eq. (A12)]

$$\langle e^{-\beta V}\rangle_0 = 1-\rho v_c\Gamma(3/4)\sum_{n=0}^{\infty}\xi_n\left(\frac{C_6}{2\sqrt{C_{12}}}\lambda^{k_\Delta}\right)^n . \tag{10}$$

Here ξ_n's are the series expansion coefficients and their values for small n are tabulated in Appendix A. The exponent k_Δ and the cavity volume v_c are defined as

$$k_\Delta\equiv k_6-\frac{k_{12}}{2}, \tag{11}$$

$$v_c\equiv\frac{4\pi r_c^3}{3}=\frac{4\pi}{3}(\lambda^{k_{12}}C_{12})^{1/4}. \tag{12}$$

Equation (10) is valid if and only if $k_\Delta>0$ which is the case for most commonly utilized exponent sets. For example, as discussed in a previous communication, Ref. 7, for $k_{12}=4$, the choice $k_6\simeq 3$ gives an almost linear TI integrand in the polynomial path TI approach (PP-TI).[25] Therefore, it was concluded that in most cases $\{k_{12},k_6\}=\{4,\approx 3\}$ would be a pretty good choice for the PP-TI approach, and with this choice of the exponent set, $k_\Delta>0$. k_Δ is also positive when the exponents are equal, $k_{12}=k_6$. For this reason, our interest concentrates on the case when $k_\Delta>0$. But, it should be noted that for $k_\Delta\leq 0$, the integral in Eq. (5) is still finite, and an asymptotic limiting behavior similar to Eq. (10) may be obtained (Appendix A).

By turning off the longer range attractive term $C_6=0$ and by setting $k_{12}=1$, we obtain the result of Postma *et al.*[15]:

$$\langle e^{-\beta V}\rangle_0=1-\frac{4\pi\rho}{3}\Gamma(3/4)C_{12}^{1/4}\lambda^{1/4}=1-\rho v_c\Gamma(3/4). \tag{13}$$

A general derivation of ΔA for a purely repulsive solute–solvent intermolecular interaction of type $V\propto 1/r^n$ ($n>3$) is reported in Appendix B. Notice that the series expansion expression in Eq. (10) has a close resemblance with the corresponding SPT expression for a hard-sphere solute dissolved in a hard-sphere solvent.[10] As noted by Postma *et al.*,[15] to obtain an *equivalent* hard-sphere solute having volume v_{HS}, the volume of the *soft* solute particle v_c needs to be scaled [e.g., $\Gamma(3/4)$ factor in Eq. (13)]. As shown in Appendix B, the scaling factor to determine the equivalent hard-sphere solute depends on the exponent of the potential employed,[29] and for $\beta V=\lambda^{k_n}C_n/r^n$, it is given as [Eq. (B2)]

$$v_{\mathrm{HS}}\rightarrow g_s v_c ,$$

$$\text{where}\quad g_s=\frac{2\sqrt{\pi}}{2^{\gamma}}\frac{\Gamma(\gamma)}{\Gamma[(\gamma+1)/2]}\quad\text{and}\quad \gamma=2-\frac{6}{n}. \tag{14}$$

As it should, $g_s\rightarrow 1$ at the $n\rightarrow\infty$ limit (which corresponds to hard-sphere interactions).

One important feature of the series expansion in Eq. (10) is that the leading term of the series is not affected by the inclusion of the attractive $1/r^6$ term. This observation in a way assures the applicability of the perturbation treatments to calculate the free energy of cavity formation for a soft-core solute with an attractive tail. To have an understanding of how many terms to keep in the series expansion, we consider, as an example, a carbon-like particle in aqueous solution. Here the solute interacts only with the oxygen site of the water molecules, and typical solute–solvent interaction values are $\sigma=3.25$ Å and $\beta\epsilon=0.25$. Using the experimental structure data for water, the closest distance that the oxygens of the different water molecules can come together is approximately 2.4 Å, thus the assumption that the solute can interact with one water molecule only would be valid up to the cavity radius of 1.2 Å.[12] Using the definition of the cavity radius, Eq. (9), an upper limit for λ expansion, $\lambda_{\max}$, may be determined:[30]

$$r_c\approx 1.2\ \text{Å}\rightarrow\lambda_{\max}^{k_{12}}=r_c^{12}/C_{12}\approx 6\times 10^{-6}. \tag{15}$$

Using $k_{12}=k_6=1$, i.e., $k_\Delta=1/2$, gives the following series expansion for our example case:

$$\langle e^{-\beta V}\rangle_0=1-\frac{4\pi\rho\sigma^3}{3}\Gamma(3/4)\lambda^{1/4}\sum_{n=0}^{\infty}\xi_n\left(\frac{\lambda}{4}\right)^{n/2}. \tag{16}$$

Substituting the respective values, we see that at $\lambda=\lambda_{\max}=6\times 10^{-6}$, the contribution of the $n=1$ term is less than 0.2%, and the contribution of $n>1$ terms is still smaller and, therefore, negligible.

2. Thermodynamic integration method

In the thermodynamic integration (TI) approach to calculating the free energy of solvation one starts with the constitutive equation[31]

$$\beta\Delta A=\beta(A_1-A_0)=\int_0^1\left\langle\frac{\partial\beta E(\lambda)}{\partial\lambda}\right\rangle_{\lambda}d\lambda, \tag{17}$$

where $\langle\cdots\rangle_\lambda$ corresponds to an ensemble average of the enclosed quantity with energy function $E(\lambda,\mathbf{X}^N)$ in the Boltzmann factor. If $E(\lambda,\mathbf{X}^N)=V(\lambda,\mathbf{X}^N)+E(\lambda=0,\mathbf{X}^N)$ then it can be rigorously shown, for any observable $\mathscr{O}$, that

$$\langle\mathscr{O}\rangle_\lambda=\frac{\langle\mathscr{O}e^{-\beta V(\lambda)}\rangle_{\lambda=0}}{\langle e^{-\beta V(\lambda)}\rangle_{\lambda=0}}. \tag{18}$$

With the defining path given by Eq. (7), and using Eq. (18), the integrand in Eq. (17), $\mathscr{T}(\lambda)$, may be expressed as[7] (m $=6,12$)

$$\beta\mathscr{T}(\lambda)\equiv\left\langle\frac{\partial\beta E(\lambda)}{\partial\lambda}\right\rangle_\lambda$$

$$=\sum_m\beta\mathscr{T}_m(\lambda)\equiv\sum_m(-)^{m/6}\frac{\mathscr{N}_m(\lambda)}{\mathscr{D}(\lambda)}, \tag{19a}$$

with

$$\mathscr{N}_m(\lambda)\equiv\frac{k_m}{\lambda}\lambda^{k_m}C_m\langle r^{-m}e^{-\beta V}\rangle_{\lambda=0},$$

$$\mathscr{D}(\lambda)\equiv\langle e^{-\beta V}\rangle_{\lambda=0}. \tag{19b}$$

In the small λ limit, the analytical expression for the ensemble average appearing as the denominator in Eq. (19a), $\mathscr{D}(\lambda)$, is already reported in the previous section, Eq. (10). The calculation of the numerators, $\mathscr{N}(\lambda)$, is quite similar and the details of their calculation may be found in Appendix A. Quoting from Eq. (A14), we have

$$\mathscr{N}_m(\lambda)=\frac{k_m}{4\lambda}\rho v_c\Gamma(3/4)\sum_{n=0}^{\infty}\zeta_{m,n}\left(\frac{C_6}{2\sqrt{C_{12}}}\lambda^{k_\Delta}\right)^n. \tag{20}$$

The coefficients $\zeta_{m,n}$ are given in Appendix A. We note that $\zeta_{6,0}=0$, so the series expansion for $\mathscr{N}_6$ starts with the $n=1$ term. Therefore, the limiting behavior of the TI integrand is solely determined by the leading term of the repulsive $1/r^{12}$ contribution as in the case of the purely repulsive interaction potential, and has a λ dependence of the form $\lambda^{(k_{12}/4)-1}$. This finding shows, as stated in Ref. 1, that when the k exponents are equal, the divergent behavior of the TI integrand can be deduced from the shortest ranged intermolecular solute–solvent interaction. Thus, as can be seen from Eq. (20), choosing the exponent $k_{12}\geqslant 4$ effectively gets rid of the divergence in the TI integrand.

Substitution of Eqs. (10) and (20) into Eq. (19a) gives the final expression for the TI integrand:

$$\beta\mathscr{T}(\lambda)=\frac{k_{12}}{4\lambda}\frac{\rho v_c\Gamma(3/4)\Sigma_{n=0}^{\infty}\zeta_n[(C_6/2\sqrt{C_{12}})\lambda^{k_\Delta}]^n}{1-\rho v_c\Gamma(3/4)\Sigma_{n=0}^{\infty}\xi_n[(C_6/2\sqrt{C_{12}})\lambda^{k_\Delta}]^n}, \tag{21}$$

where $\zeta_n=\zeta_{12,n}-\zeta_{6,n}$.

B. Solute with multisites

In generalizing the analytical expressions in Eqs. (10), (20), and (21) for the solvation free energy of a small single-site solute to multisite solutes, we notice that difference would arise from the exclusion of the solvent molecules not only from the volume occupied by the *labeled* solute site located at the origin but also from the volume occupied by the *other* solute sites as well. It is clear from the derivation given in Appendix A that to calculate the solvation free energy or the corresponding TI integrand, the ensemble averages of the solute–solvent site–site interaction potential terms with respect to pure solvent state is needed. Within ISM presentation, the solute–solvent molecular interaction potential is site–site additive, or in other words

$$V=\sum_{i,j}V_{ij}=\sum_i V_i=\sum_j V_j, \tag{22}$$

where the sum over j and i goes over the distinct sites of the solvent and the solute molecules, respectively. V_i is the interaction potential energy of the solute site i with the solvent, as given by Eq. (7) for each solute site. For an observable of the form $\mathscr{O}=\Sigma_i\mathscr{O}_i$, the integrals

$$I_{O,i}=\int \mathbf{d}^3\mathbf{r}\,\mathscr{O}_i e^{-\beta V} \tag{23}$$

may be expressed as

$$I_{O,i}=\int \mathbf{d}^3\mathbf{r}\,\mathscr{O}_i\prod_k e_k(\mathbf{r})$$

$$=\int \mathbf{d}^3\mathbf{r}\,\mathscr{O}_i e_i(\mathbf{r})\prod_{k\neq i} e_k(\mathbf{r})$$

$$=\int \mathbf{d}^3\mathbf{r}\,\mathscr{O}_i e_i(\mathbf{r})\prod_{k\neq i}[1+f_k(\mathbf{r})], \tag{24}$$

where the product over k goes over the distinct solute sites. $f_k(\mathbf{r})$ is the Mayer f function for the solute site k and defined as $f_k(\mathbf{r})=e_k(\mathbf{r})-1=e^{-\beta V_k(\mathbf{r})}-1$. $f_k(\mathbf{r})$ is approximately equal to $-\beta V_k$ when V_k is small, and is approximately equal to -1 when V_k is very large, i.e., when $\mathbf{r}$ is inside the repulsion core of solute site k. Expanding the product results in

$$I_{O,i}=\int \mathbf{d}^3\mathbf{r}\,\mathscr{O}_i e_i\mathscr{E}_i(\mathbf{r}), \tag{25a}$$

$$\mathscr{E}_i(\mathbf{r})\equiv 1+\sum_{k\neq i}f_k+\sum_{\substack{k,k'\neq i\\ k\neq k'}}f_kf_{k'}+\cdots. \tag{25b}$$

Since the f functions approximately vanish outside the cavities, the terms of the function $\mathscr{E}_i(\mathbf{r})$ would converge very rapidly. A close look at the $\mathscr{E}_i(\mathbf{r})$ function reveals that it is approximately equal to unity except at points $\mathbf{r}$ belonging to a cavity formed by solute sites other than the site i, and at these points $\mathscr{E}_i(\mathbf{r})$ approximately vanishes. Second and higher order terms in Eq. (25b) correspond to the overlap of two or more solute sites. For very small λ, the cavities created by solute sites are small and the overlap of the site cavities would not be considerable. Therefore, the contribution of the second or higher order terms are expected to have little importance.

In the implementation in this report, rather than truncating the series for $\mathscr{E}_i(\mathbf{r})$ at a certain level, we use Eq. (25a) and numerically calculate the value of the required integrals (a three-dimensional integral for each solute site) at each λ point. Even though we use numerical methods to calculate the necessary integrals, analytical treatment of some simple cases helps to understand the methodology.

To give an example, let us consider a two-site solute assuming that the f_k functions can be approximately represented as equal to -1 inside the cavities of solute sites and 0 elsewhere (as in the case of hard-sphere interactions). Two sites of the solute are separated by a bond distance r_l. When the sum of the cavity radius of the two sites is smaller than r_l, the site cavities do not overlap. If this is the case, then the contribution of the sites to the ensemble average of any operator, $\mathcal{O}=\mathcal{O}_1+\mathcal{O}_2$, may be calculated as [total solute–solvent interaction energy may similarly be split as $V(\mathbf{r})=V(i=1;\mathbf{r})+V(i=2;\mathbf{r})$]

$$\begin{aligned} I_{O,1}&=\int \mathbf{d}^3\mathbf{r}\ \mathcal{O}_1(\mathbf{r})e^{-\beta V(1;\mathbf{r})}e^{-\beta V(2;\mathbf{r}+\mathbf{r}_l)} \\ &\simeq -v_2\mathcal{O}_1(-\mathbf{r}_l)\int \mathbf{d}^3\mathbf{r}\ \mathcal{O}_1(\mathbf{r})e^{-\beta V(1;\mathbf{r})}, \\ I_{O,2}&=\int \mathbf{d}^3\mathbf{r}\ \mathcal{O}_2(\mathbf{r})e^{-\beta V(1;\mathbf{r}-\mathbf{r}_l)}e^{-\beta V(2;\mathbf{r})} \\ &\simeq -v_1\mathcal{O}_2(\mathbf{r}_l)\int \mathbf{d}^3\mathbf{r}\ \mathcal{O}_2(\mathbf{r})e^{-\beta V(2;\mathbf{r})}. \end{aligned} \tag{26}$$

Since the site cavity volumes v_1 and v_2 in Eq. (26) are small, and assuming a smooth functional form for the operand, the integrand values would not vary appreciably over the cavities. This argument justifies the approximation in passing from the first to the second equality in the above equations. Comparison of Eq. (26) with Eq. (23) shows that the solvation properties of a multisite solute may be studied by considering the contribution of the individual solute sites separately as if the other sites do not exist, and the required corrections can, to a good degree, be incorporated at the end.

The above example clearly shows that the necessary corrections are proportional to site cavity volumes, hence, they vanish as $\lambda\rightarrow 0$. The $\lambda=0$ limit corresponds to the limit at which the molecular solute has chemically bonded *point* interacting sites. Since site cavities do not occupy any volume, corrections to the ensemble average of any operator $\mathcal{O}_i$ would vanish, and the system would be equivalent to an *independent* set of spherical point cavities placed into the solvent.

When λ is increased, the spheres will start to overlap at a certain value. Although the calculation of the overlap volume is possible in most cases, the calculation becomes rather cumbersome as the number of sites increases.[32–35] Note that at around $\lambda_{\max}$, which corresponds to the upper limit for the perturbation treatment to hold, a typical site cavity radius is in the same range as the bond lengths of chemically occurring systems. Thus around $\lambda\approx\lambda_{\max}$, the overlap between the solute sites, and hence the necessary corrections, may be considerable. Therefore, approximations introduced in Eq. (26) can only be used at much smaller λ values.

An additional ambiguity for the solvation of a multisite solute is the determination of the upper bound $\lambda_{\max}$, for which the perturbation treatment would be valid. Fortunately, Eq. (13) gives a hint that the upper bound criterion may be based on the cavity volume. This point is further discussed in Sec. IV.

C. Molecular solutes with site charges

In most of the studies in the literature, and in some cases of this report, the molecular interaction sites also carry charges. Unfortunately, a similar analytical derivation when the solute molecule also interacts with the solvent molecules via Coulomb forces is not currently feasible. But it is well established that the Coulomb interaction can be *renormalized* giving rise to a damped $1/r$-type behavior.[36] As our derivation in this report shows, the addition of the longer ranged weakly attractive $1/r^6$ term to the pair interaction does not change the divergence of the total TI integrand. Therefore, we make the *ad hoc* assumption that the inclusion of the site charges would not change the conclusion that the divergence of the TI integrand is determined by the shortest ranged repulsive interaction term. Even though we cannot justify it rigorously at this point, it is tempting to conclude, in analogy with Eq. (11), that the above stated assumption would be valid as long as

$$12k_1>k_{12}. \tag{27}$$

Concerning the contribution of the Coulomb term to the solvation free energy or to the TI integrand, the ensemble average of the Coulomb interaction between any one of the charged solute sites and any one of the solvent sites diverges when treated separately. But, if the solute and solvent molecules are charge neutral and if λ is very small, the pure solvent state ensemble average of the total solute–solvent Coulomb interaction

$$\langle\beta V_1\rangle_{\lambda=0}=\lambda^{k_1}\sum_{i,j}\left\langle\frac{q_iq_j}{r}\right\rangle_0 \tag{28}$$

vanishes. This vanishing is exact at $\lambda=0$ and holds only approximately for nonzero λ's. Because of the uniform solvent distribution assumption (see Sec. II A), the pure solvent state ensemble average of the local solvent charge density is conserved at each space point around the solute. Due to this local solvent charge neutrality, site–site solute–solvent Coulomb interaction contributions are cancelled, and therefore, the ensemble average in Eq. (28) vanishes. This predicted cancellation depends on the degree of goodness of the utilized assumptions and on the validity of the perturbation treatment. Thus, the deviation of the ensemble average $\langle V_1e^{-\beta V}\rangle_{\lambda=0}$ (which is proportional to the corresponding TI integrand, $\mathcal{T}_1$) from nullity would actually reflect on the goodness of the underlying assumptions and on the validity of perturbation treatment. The vanishing of the Coulomb term contribution to TI at the $\lambda=0$ limit was numerically shown for the case of aqueous solvation of alanine-dipeptide in a recent communication (see Figs. 4 and 5 of Ref. 7). It is further investigated in Sec. V of this study.

TABLE I. Interaction potential and molecular geometry parameters.[a]

Site	C_{12} ($\times 10^{-3}$)	C_6	q
Solute[b]			
SS	2516.6	0	0.0
LJ	2516.6	1228.8	0.0
UD[c]	2516.6	1228.8	0.0
DD[c]	2516.6	1228.8	±0.5
Solvent[d]			
O	582	595	−0.834
H	0	0	0.417

[a] Site i of a molecule interacts with site j of a different molecule as $C_{12,ij}/r^{12}-C_{6,ij}/r^6+q_iq_j/r$, where $C_{12,ij} = \sqrt{C_{12,i}C_{12,j}}$ and $C_{6,ij} = \sqrt{C_{6,i}C_{6,j}}$. C_{12} and C_6 are in units of kcal/mol Å^{12} and kcal/mol Å^6, respectively.

[b] For the molecular parameters of the alanine-dipeptide solute, see Ref. 7.

[c] Dumbbells have two chemically bonded sites which are separated by $l=1.5$ Å. The sites have the same short range potential parameters.

[d] The water bond lengths are $l_{OH}=0.9572$ Å and $l_{HH}=1.5139$ Å.

III. COMPUTATIONS

In this report, the aqueous solvation properties of a soft sphere particle, of a Lennard-Jones particle, and of a dumbbell with two fused equivalent sites separated by a bond length of 1.5 Å are studied. The solvent liquid water is represented by the TIP3P model.[37] To investigate the contribution and effects of the Coulomb term, the aqueous solvation of a dipolar dumbbell which has the same molecular geometry and short range potential parameters of the dumbbell solute but also carries partial site charges $\pm 0.5e$ is also studied. The molecular model parameters of all simulated systems are reported in Table I. In addition, to investigate the solvation properties of a *complex* solute, solvation properties of alanine-dipeptide at C_7 and α_R conformations were simulated at a few small coupling parameter values. Alanine-dipeptide was the subject of a recent study in our group, Ref. 7, which tabulates the solute molecular model parameters.

All simulations were performed using constant volume canonical ensemble Monte Carlo (MC) simulations. Face-centered-cubic periodic boundary conditions were employed. Water–water interactions were truncated at a spherical cutoff of 7.75 Å, and the minimum image convention was used for the solute–solvent interactions. All simulations involved a single solute and 215 solvent molecules at $T=298$ K. Additional relevant details of the simulation runs are tabulated in Tables III–IX.

The sampling rate was accelerated by using force biasing,[38] as well as by using a preferential sampling of type $1/R_{ns}$, where R_{ns} is the distance between the nearest solute interaction site and the solvent molecules. The convergence problem in the molecular simulations is more pronounced when the coupling parameter is small. The encountered convergence problem partly arises from the fact that around $\lambda \simeq 0$ the solute–solvent interaction contribution becomes very small, and the motion of the molecules are mainly determined by the interaction of the solvent molecules among themselves. Thus, to enhance the contribution of the solute to the underlying dynamics of the system, certain *methodological tricks* may prove to be useful. In addition to preferential sampling and force bias, we have employed a selection scheme in which the solute has an eight times higher selection probability with respect to the selection probability of the solvent molecules. When the particular studied property is strongly related to the solute–solvent interactions, the employment of a selection rule in which the solute is moved as much as possible is important in improving the statistics and the convergence.[39] Availability of such enhanced solute selection schemes in calculating the solvation free energies is probably the most important advantage of the Monte Carlo simulation method over the corresponding molecular dynamics method.

IV. QUADRATURE INTEGRATION

As we have shown in Ref. 7, the TI approach using Gaussian quadrature integration with four to nine points were quite successful in studying the solvation free energies at different conformations of a small biomolecule. Based on this success and on the results reported in the following sections, we propose to incorporate the analytical theory developed in this report into the TI approach to calculate the solvation free energies using quadratures. This section gives the details of how this incorporation can be achieved, and states a particular integration scheme which will be employed later in this report.

As is shown in Sec. II A 1, and also in Sec. V, the analytical theory is valid up to λ_{max}, and typical values of λ_{max} for realistic systems are smaller than 10^{-6}. This approximately corresponds to the first Gaussian quadrature point with exponent $k=4$. Let us say that an n-point Gaussian quadrature is chosen to perform molecular simulations, and various TI integrand terms corresponding to different solute–solvent interaction contributions, namely $\mathcal{T}_i$ with $i=12$, 6, and 1 in this report, are calculated. In the most straightforward implementation, the simulation results at n different λ points can be supplemented by analytical results at p small λ points. Overall $n+p$ results can then be fitted to a certain functional form, or $(n+p)$-point quadrature integration can be employed. The drawback of this approach is that the first root point of the Gaussian quadratures is generally very small and additional use of analytical results at p additional very small λ points would bias the integration toward the small λ range. Another drawback is that when two of the root points are very close to each other, the errors arising from the statistical fluctuations produces unphysical oscillations in the resulting polynomial fit to the integrand. Simulations at very small λ generally carry the largest statistical error, and therefore, the mismatch between the results at two very close adjacent points may be considerable. Due to the statistical fluctuations, this type of mismatches between two simulation results, or a simulation result and an analytical result, is actually unavoidable. When this mismatch happens, forcing the fitting polynomial to go through all the data points (as in quadrature integration) gives rise to pronounced oscillations in the polynomial fit.

To avoid these complications, we propose and implement the following approach. The simulation at the first root point of the n-point Gaussian quadrature is simply omitted, and the $n-1$ simulation results are supplemented by the analytical result at $\lambda=0$, which is exact. This approach again

TABLE II. Root points and weight functions of five-point quadratures.[a]

Gaussian quadrature		Modified Gaussian quadrature	
λ_i	$\mathscr{C}_i$	λ_i	$\mathscr{C}_i$
0.046 91	0.118 46	0.000 00	0.076 35
0.230 76	0.239 31	0.230 76	0.307 51
0.500 00	0.284 46	0.500 00	0.243 74
0.769 24	0.239 31	0.769 24	0.259 77
0.953 09	0.118 46	0.953 09	0.112 63

[a] For an n-point quadrature integration, the integral over λ of a function $f(\lambda)$ is given as $\Sigma_i^n \mathscr{C}_i f(\lambda_i)$. The modified Gaussian quadrature is explained in Sec. IV.

TABLE III. Soft sphere and Lennard-Jones solutes.[a,b]

Solute	Theory	$\lambda^{3/4}\mathscr{T}_{12}$	$\lambda^{1/4}\mathscr{T}_6$	$\lambda^{3/4}\mathscr{T}$
$\lambda=10^{-7}$				
SS	AT	1.08	⋯	1.08
	MC	1.08±0.03	⋯	1.08±0.03
LJ	AT	1.08	−3.23	1.08
	MC	1.11±0.03	−3.22±0.03	1.10±0.03
$\lambda=10^{-6}$				
SS	AT	1.20	⋯	1.20
	MC	1.21±0.03	⋯	1.21±0.03
LJ	AT	1.20	−3.60	1.20
	MC	1.21±0.02	−3.47±0.03	1.21±0.02
$\lambda=4.8424\times10^{-6}(c)$				
SS	AT	1.37	⋯	1.37
	MC	1.33±0.02	⋯	1.33±0.02
LJ	AT	1.38	−4.11	1.37
	MC	1.32±0.03	−3.68±0.04	1.32±0.03

[a] TI integrand values are in kcal/mol. $\mathscr{T}_{12}$ and $\mathscr{T}_6$ are defined in Eq. (13). AT is the prediction of the scaled particle type theory, Eq. (15). SS and LJ stand, respectively, for the soft sphere and Lennard-Jones solutes.

[b] The estimated statistical errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 6×10^6 steps $\approx2.78\times10^4$ cycles.

[c] This λ value is the first root point of five-point Gaussian quadrature with $k=4$, and these results are repeated in Tables VI and VII.

results in an n-point quadrature approach. But this time, since the first root point is replaced with zero, the choice of the sampled coupling parameter points results in a *modified Gaussian* quadrature integration. Analytical expression for the TI integrand simplify considerable at $\lambda=0$, and may be calculated using Eqs. (10) and (20). If $k_{12}=4$ and $k_\Delta>0$, Eq. (10) gives

$$\langle e^{-\beta V}\rangle_0(\lambda=0)=\mathscr{D}(0)=1, \tag{29}$$

and, using Eq. (20), we obtain

$$\begin{aligned}\beta\mathscr{T}_{12}(\lambda=0)&=\mathscr{N}_{12}(0)=\frac{\pi\rho k_{12}}{3}\Gamma(3/4)C_{12}^{1/4},\\ \beta\mathscr{T}_6(\lambda=0)&=-\mathscr{N}_6(0)=-\frac{\pi\rho k_6}{3}\Gamma(1/4)C_6C_{12}^{-1/4}\lambda^{k_\Delta}\\ &\rightarrow0,\end{aligned} \tag{30}$$

and as assumed $\beta\mathscr{T}_1(\lambda=0)=0$. Since one of the simulation points is omitted, this approach also allows for some savings in computational expenditure. For example, for the five-point quadrature TI, the savings is roughly equal to 20%. In this report, we employ five-point quadratures, and compare the performance of the Gaussian quadratures with its above stated modified form. The root points and the integration weight coefficients for both cases are tabulated in Table II.

V. LIMITING BEHAVIOR

A. Soft sphere and Lennard-Jones solutes

In Table III, the results for the solvation of a single site Lennard-Jones (LJ) solute are reported and compared to the solvation of a soft sphere (SS) solute. Note that these two solutes have the same C_{12} parameter, but $C_6=0$ for the SS solute (Table I). Comparison of the SS and LJ results reveal that, as expected, $\mathscr{T}_{12}$ for the LJ solute lies slightly above the one for the SS solute. The existence of the negative $1/r^6$ solute–solvent energy term for the LJ case allows the solute to come closer to the solvent molecules, thus sampling more of the higher $1/r^{12}$ solute–solvent energy range. Since the $\mathscr{T}_{12}$ contribution is proportional to the $1/r^{12}$ solute–solvent energy term, the $\mathscr{T}_{12}$ for the LJ solute is expected to be larger than the $\mathscr{T}_{12}$ for the SS solute. But, as predicted by the analytical theory, the inclusion of the attractive $1/r^6$ interaction term has little effect on the contribution of the $1/r^{12}$ interaction term, $\mathscr{T}_{12}$, which is expected to be nearly equal for both SS and LJ solutes at small λ's.

The upper limit for the applicability of perturbation treatment, $\lambda_{\max}$, for SS and LJ solutes may approximately be obtained by using Eq. (15). Substituting the value for C_{12} (Table I) with $r_c=1.2$ Å gives $\lambda_{\max}=4.36\times10^{-6}$ (unless stated otherwise, from this point on in this report, for simplicity, k_{12} and k_6 are treated as if they are equal to 1). As can be seen from Table III, the agreement between the analytical and the simulation results for $\lambda\leq\lambda_{\max}$ is well within the statistical error bars which are estimated to be about 3%. The only noticeable exception is $\mathscr{T}_6$ at $\lambda=10^{-6}$, for which the difference is 3.5%. Although it may be the result of insufficient statistical sampling, this disagreement is most probably due to the inappropriateness of the perturbation treatment in the analytical theory at this coupling parameter. This conclusion is based on the observations that the performed simulation runs are comparatively long and seem to be well converged. Also notice that the assumed upper bound of the perturbation treatment, Eq. (9), can only be taken as a sensible estimate and its validity needs to be further justified using independent tests. As expected, the agreement between the analytical theory and the simulation results gets worse as λ is increased. For example, at $\lambda=4.84\times10^{-6}$, which is slightly larger than the estimated $\lambda_{\max}$, the disagreement in $\mathscr{T}_6$ values is approximately 10%. This sudden worsening of the agreement between the results when there is a small increase in the coupling parameter implies that the validity criteria employed for the perturbation treatment is actually pretty good, but it may need slight alterations. Therefore, it can be safely stated that the criteria given by Eq. (9) with "some" allowance of indeterminance is a quite successful

TABLE IV. Uncharged dumbbell and dipolar dumbbell solutes.[a,b]

Solute	Theory	$\lambda^{3/4}\mathscr{T}_{12}$	$\lambda^{1/4}\mathscr{T}_6$	$\mathscr{T}_1$	$\lambda^{3/4}\mathscr{T}$
$\lambda=10^{-8}$					
UD, DD	AT	2.14	−6.29	c	2.14
UD	MC	2.14±0.06	−6.17±0.08	···	2.14±0.06
DD	MC	2.15±0.05	−6.19±0.07	0.12±0.27	2.15±0.05
$\lambda=10^{-7}$					
UD, DD	AT	2.19	−6.57	c	2.18
UD	MC	2.17±0.04	−6.27±0.06	···	2.17±0.04
DD	MC	2.13±0.05	−6.27±0.07	0.05±0.30	2.13±0.05
$\lambda=10^{-6}$					
UD, DD	AT	2.44	−7.57	c	2.43
UD	MC	2.30±0.04	−6.61±0.06	···	2.30±0.04
DD	MC	2.28±0.04	−6.63±0.07	0.07±0.22	2.28±0.04

[a] TI integrand values are in kcal/mol. $\mathscr{T}_{12}$ and $\mathscr{T}_6$ are defined in Eq. (13), and $\mathscr{T}_1$, the contribution of the Coulomb term to the TI integrand for the DD solute, is defined analogously (see Ref. 7). AT is the prediction of the scaled particle type theory, where the required ensemble average integrals are numerically calculated using Eq. (18). UD and DD stand, respectively, for the (uncharged) dumbbell and dipolar dumbbell solutes.

[b] The estimated statistical errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 6×10^6 steps.

[c] Analytical theory assumes that the contribution of solute–solvent Coulomb interaction vanishes.

criteria for establishing the upper bound of the perturbation treatment.

B. Dumbbell and dipolar dumbbell solutes

Possibly the simplest models to study the solvation properties of multisite solutes are dumbbells. As discussed in Sec. II B, the determination of $\lambda_{\max}$ is not unambiguous for multisite solutes. One possibility for estimating $\lambda_{\max}$ is the use of a criterion based on the cavity volumes. This criterion can be justified on the physical grounds that it is the pressure of the solvent exerted on the surface of the cavity that requires the work, and thus gives rise to solvation free energies. Of course other criteria, such as equal energy surface areas, may also be employed. Since we are only interested in a rough $\lambda_{\max}$ criteria, and since it is conceptually simpler and easy to implement, we prefer to utilize the $\lambda_{\max}$ values obtained by using the criterion based on the cavity volumes.

To obtain the desired approximate relation, we start with Eq. (10). For small βV,

$$\langle e^{-\beta V}\rangle_0 \cong \langle 1-\beta V\rangle_0 \cong 1-\rho v_c\Gamma(3/4), \tag{31a}$$

or equivalently

$$\langle\beta V\rangle_0 \cong \rho v_c\Gamma(3/4) = \cup_i \frac{4\pi\rho}{3}\Gamma(3/4)(\lambda C_{12,i})^{1/4}. \tag{31b}$$

The above equation shows that two solutes will have approximately equal cavity volumes if the ensemble averages of the respective solute–solvent interaction potentials, $\langle\beta V\rangle_0$'s, are equal. $\cup_i$ in Eq. (31b) stands for the summation over individual site cavity volumes while taking into consideration the possible overlap of the site cavities to avoid the overcounting.

$\lambda_{\max}$ for uncharged dumbbell (UD) and dipolar dumbbell (DD) solutes may be calculated by utilizing Eq. (31b) with LJ solute as the reference. As reported in Table I, C_{12} parameters for the UD and DD solute sites and LJ solute are the same. Therefore, Eq. (31b) establishes the estimate for $\lambda_{\max,D}$ for the UD and DD solutes as

$$\varphi_D^4\lambda_{\max,D} \cong \lambda_{\max,\mathrm{LJ}}. \tag{32}$$

Depending on the amount of overlap between site cavities of the dumbbell solute, the factor φ_D appearing in Eq. (32) can have values between one (completely overlapping dumbbell site cavities) and two (nonoverlapping cavities). For some small molecules, analytical formulas to calculate the volume of overlapping spheres exist,[17,23,35] but, its implementation gets complicated as the number of sites increase. Note that the solute sites are kept apart by the chemical bonds and the involved site cavities are relatively small, therefore, a likely value of φ_D should be closer to two. Since we are only interested in an approximate upper bound, the use of the most conservative estimate ($\varphi_D=2$) would be safer in establishing the coupling parameter range for which the perturbation treatment should be valid. Using the $\lambda_{\max}$ for LJ solute, and using $\varphi_D=2$, we obtain $\lambda_{\max,D}\approx 2.7\times10^{-7}$ as the upper bound estimate for dumbbell solutes.

In the analytical theory calculations for DD solute, it is implicitly assumed that (Sec. II C), for very small λ, the contribution of the solute–solvent Coulomb interaction to the solvation free energy vanishes. Because of this assumption, and since they have the same short range potential parameters, the analytical theory results for UD and DD solutes are the same.

As in the case of LJ and SS solutes, the agreement between the analytical theory and the simulation results (Table IV) is quite satisfactory within the valid λ range implying a good enough convergence of the numerical calculations. At $\lambda=10^{-8}$, MC and analytical results are in agreement within the statistical fluctuations. Similarly, for $\lambda=10^{-7}\approx\lambda_{\max}$, $\mathscr{T}_{12}$ results are well within the error bars. As in the case of LJ solute, disagreement between the $\mathscr{T}_6$ results, 4.6%, is worse when compared to $\mathscr{T}_{12}$ results, but an uncertainty of this

TABLE V. Alanine-dipeptide solute.[a,b]

Conf.	Theory	$\lambda^{3/4}\mathscr{T}_{12}$	$\lambda^{1/4}\mathscr{T}_6$	$\mathscr{T}_1$	$\lambda^{3/4}\mathscr{T}$
$\lambda=10^{-13}$					
C_7, α_R	AT	9.12	−26.44	c	9.12
C_7	MC	9.19±0.37	−26.14±0.54	−0.33±0.97	9.19±0.37
α_R	MC	8.92±0.34	−25.94±0.51	0.88±1.05	8.92±0.34
$\lambda=10^{-12}$					
C_7, α_R	AT	9.37	−27.17	c	9.37
C_7	MC	8.90±0.34	−25.60±0.41	0.06±1.05	8.90±0.34
α_R	MC	8.87±0.28	−25.72±0.49	−1.05±1.07	8.87±0.28
$\lambda=10^{-11}$					
C_7, α_R	AT	9.86	−28.54	c	9.86
C_7	MC	8.97±0.28	−26.03±0.45	−0.95±0.87	8.97±0.28
α_R	MC	9.25±0.27	−26.34±0.45	−0.64±1.06	9.25±0.27
$\lambda=10^{-10}$					
C_7, α_R	AT	10.86	−31.30	c	10.86
C_7	MC	9.46±0.26	−26.90±0.35	−1.84±0.73	9.46±0.26
α_R	MC	9.47±0.23	−26.83±0.30	−1.06±0.58	9.47±0.23

[a] TI integrand values are in kcal/mol. $\mathscr{T}_{12}$ and $\mathscr{T}_6$ are defined in Eq. (13), and $\mathscr{T}_1$, the contribution of the Coulomb term to the TI integrand, is defined analogously (see Ref. 7). AT is the prediction of the scaled particle type theory, where the required ensemble average integrals are numerically calculated using Eq. (18). For the λ range of this table, the analytical theory results are the same for both conformations.

[b] The estimated statistical errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 6×10^6 steps.

[c] Analytical theory assumes that the contribution of solute–solvent Coulomb interaction vanishes.

magnitude is not unusual in solvation free energy simulations.[40] When λ is further increased to $10^{-6}(>\lambda_{\max})$, the disagreement between the analytical and simulation results becomes quite large, as expected.

Comparison of UD and DD results (Table IV) shows that the existence of the solute–solvent Coulomb interaction for the DD solute has only a slight effect on the contributions of the shorter ranged terms to the TI integrand. Especially the contribution of the r^{-6} interaction term, $\mathscr{T}_6$, is almost unchanged. Although there is a small but noticeable change in $\mathscr{T}_{12}$, 2% at $\lambda=10^{-7}$, this difference between UD and DD results is still less than the statistical fluctuations and cannot be conclusive. This observation strongly supports the validity of the underlying assumption of the analytical theory in regard to the effects of the solute–solvent Coulomb interaction as discussed in Sec. II C.

C. Alanine-dipeptide solute

To further investigate the limiting behavior of solvation free energies, simulations were performed at a few coupling parameter values for the alanine-dipeptide solute at C_7 and α_R conformations with the results tabulated in Table V. Comparison of the analytical theory with the numerical method can be used to test if the thermodynamic integration approach is capable of retaining the correct limiting behavior, and if the analytical theory is appropriate to study the solvation of this relatively complex system. As in the DD case, analytical theory calculations assume that the contribution of the solute–solvent Coulomb interaction to the solvation free energy and to the TI integrand vanishes. The solvation properties of alanine-dipeptide were studied in detail in an earlier work,[7] and it was numerically shown that the contribution of the Coulomb interaction term to the TI integrand vanishes at the small λ limit (see Figs. 4 and 5 of Ref. 7). Table V further shows that, at small λ's, contribution of the solute–solvent Coulomb interaction to the TI integrand is almost zero for both C_7 and α_R conformations. This vanishing of the solute–solvent Coulomb interaction contribution further justifies the underlying assumption of the analytical theory.

An analysis[41] similar to the one given in Eqs. (31) and (32) for UD and DD solutes estimates $\varphi_{\text{pep}}=9.2$, and establishes the upper bound for the perturbation treatment as

$$\varphi_{\text{pep}}^4\lambda_{\max,\text{pep}}\approx\lambda_{\max,\text{LJ}} \quad \text{or as} \quad \lambda_{\max,\text{pep}}\approx6\times10^{-10}. \tag{33}$$

Table V reflects that, at $\lambda=10^{-13}$, the disagreement between the simulation and the analytical results is about 2%. This is to be compared to the statistical fluctuations in the simulations which are roughly 4%. But as λ is increased, the agreement between the numerical and the analytical results rapidly worsens, becoming approximately 6%, 9%, and 14% for $\lambda=10^{-12}$, 10^{-11}, and 10^{-10}, respectively. There are two possible sources of this rather large disagreement. First is the possibility that the phase space is undersampled in the numerical methods. The second and the more likely case is due to the omission of the Coulomb interactions, especially that of the hydrogen sites, in the analytical theory.

The results of Tables IV and V justify that, for small enough λ, the overall contribution of the solute–solvent electrostatic interaction to the solvation free energies and to the TI integrand can be neglected as is done in the analytical theory. Because of this neglect, it is intrinsically assumed in the analytical theory that the water hydrogens do not play an explicit role in the expulsion of the water molecules from the cavities formed by the solute molecule, and it is also assumed that the water molecules around solute sites can have any orientation. These deductions follow from the model used to represent the water molecules, in which the partially charged hydrogen sites do not have any short range repulsion cores around them. Due to the complex structure of the dipeptide molecule, the charge interaction between the water hydrogens and the dipeptide sites would prohibit the water molecules from having certain unfavorable orientations (especially when they are in the solute pockets). However, these orientational effects are overlooked in the analytical theory. These additional effects coming from the solute–solvent Coulomb interactions are expected to vanish as $\lambda\rightarrow 0$ as discussed in Sec. II. The effects of the electrostatic interactions of the hydrogen sites can be mimicked in the analytical calculations by assigning very small cores to water hydrogens. This makes the solute sites "see" the water hydrogens, and, if the assigned hydrogen cores are tiny, the overall modeling of the solute–solvent interaction would be left almost unchanged. Assignment of very small repulsion core parameters to water hydrogens is in effect equivalent to considering the analytical theory to be valid up to a smaller cavity radius,[30] $r_c\approx 0.75$ Å. Thus a different $\lambda_{\max}$ could be derived, and the new $\lambda_{\max}$ would be considerable smaller than the one obtained in Eq. (33). In fact, the use of $r_c=0.75$ Å establishes the upper bound of validity as $\lambda_{\max,\text{pep}}=2\times 10^{-12}$, below which the observed agreement between the analytical and the molecular simulation results is very good. This observation strongly points to the effects of the solute–solvent charge interactions.

Based on the results for the aqueous solvation of alanine-dipeptide, it can be safely concluded that the analytical theory developed in this report using the concepts of scaled particle theory can be used for studying the solvation properties of complex and of site-charged solutes. But, whenever complex solutes with charged interaction sites are involved, a much more conservative range of validity criteria may need to be employed.

Another outcome of the analytical theory for the multisite solutes, as discussed in Sec. II B, is that as $\lambda\rightarrow 0$, the corrections to the TI integrands coming from the additional solute sites [Eqs. (23)–(26)] asymptotically vanish. This effectively gives an *independent* set of solute sites. Therefore, as long as the same short range site potential parameters, C_{12} and C_6, are used in studying the solvation properties of a certain solute at different conformations, then the asymptotic limiting behavior should be independent of conformation. This provides an additional test of the convergence of the molecular simulations. In our alanine-dipeptide test case, the interaction potential parameters are the same at both C_7 and α_R conformations. As Table V reflects, for small enough λ, the simulations at these two different conformations are in very good agreement with each other, thus confirming the success of the analytical theory in predicting this particular aspect of solvation.

VI. TEST OF GAUSSIAN QUADRATURES

In some recent contributions from our laboratory,[7,25,42] we have shown that Gaussian quadratures capture all the essentials of the TI integrand, and quadrature integration can be successfully utilized in taking the required integrals to obtain the solvation free energies. The most important advantage of quadrature thermodynamic integration over comparative schemes such as slow growth is that simulations are necessary at only a few coupling parameter points which makes it an economical method.

One intrinsic problem with quadrature integration is that a certain functional form is assumed from the beginning, and the results are fitted to that form. For example, in Gaussian quadratures the integrand is fitted to a polynomial. The actual TI integrand is of course much more complex, and its functional form for the whole λ range is mostly unknown. For the small λ range, our analytical derivation actually gives some clues regarding the functional form of the TI integrand. Using Eqs. (10) and (20) in Eq. (19a), and performing the polynomial division, we obtain

$$\begin{aligned}\beta\mathscr{T}_{12}(\lambda)&\propto\lambda^{k_{12}/4-1}\sum_{n=0}^{\infty}\gamma_{12,n}(1+\tau_{12,n}\lambda^{k_{12}/4})\lambda^{nk_\Delta},\\ \beta\mathscr{T}_{6}(\lambda)&\propto\lambda^{k_6-k_{12}/4-1}\sum_{n=0}^{\infty}\gamma_{6,n}(1+\tau_{6,n}\lambda^{k_{12}/4})\lambda^{nk_\Delta}.\end{aligned}\tag{34}$$

It was shown that the use of $k_{12}\geqslant 4$ was important in eliminating the divergencies from the integrand function. Equation (34) shows its other aspect: when $k=k_{12}=k_6=4$ ($k_\Delta=2$), the TI integrand is an exact polynomial. Hence, the use of Gaussian quadrature would be appropriate. This analysis also predicts that if the set $k=k_{12}=k_6=1$ ($k_\Delta=1/2$) is employed, the resulting TI integrand is again a polynomial, but this time for the variable $\lambda^{1/4}$, and the TI integrand has an overall multiplying $\lambda^{-1/4}$ factor making it divergent at $\lambda=0$. Therefore, even though the divergent factor can be accounted for analytically, as done in Ref. 43, for $k=1$, a blind application of Gaussian quadratures should be avoided. Alternatively, a variable change from λ to $\lambda^{1/4}$ can be made, and Gaussian quadratures can subsequently be used. This section discusses these points.

To test how well the Gaussian quadratures perform in evaluating the required integrals, we did some additional MC simulations for SS, LJ, UD, and DD solutes. Two different sets of five-point Gaussian quadratures with equal k exponents were used.[44] Also simulations were performed for each *full* solute, i.e., at $\lambda=1$. Tables VI–IX present the results of these additional simulations. These simulations combined with the ones at smaller λ's (Tables III and IV) add up to a total of 13 simulation points for LJ and SS solutes, and to a total of 14 simulation points for UD and DD solutes.

For each solute, the combined set of MC simulation results were first least-square fitted to the functions[45]

$$\begin{aligned}\lambda^{3/4}\mathscr{T}_{12}&=a_{12,0}+\sum_{j=1}^{N_j}a_{12,j}\lambda^{j/4},\\ \lambda^{1/4}\mathscr{T}_{6}&=a_{6,0}+\sum_{j=1}^{N_j}a_{6,j}\lambda^{j/4},\end{aligned}\tag{35}$$

and then analytically integrated from zero to one to obtain the solvation Helmholtz free energies. Results of the analytical theory at $\lambda=0$ were used for the constants $a_{12,0}$ and $a_{6,0}$ to assure that the fitting functions have the correct limiting

TABLE VI. Soft sphere solute.[a]

$\lambda^{1/k}$	$\mathcal{T}=\mathcal{T}_{12}$
$k=1$	
0.046 91	20.18±1.77
0.230 76	6.16±0.56
0.500 00	3.56±0.30
0.769 24	2.49±0.20
0.953 09	2.14±0.16
$k=4$	
0.046 91	12 837±218
0.230 76	161.70±8.25
0.500 00	16.28±1.49
0.769 24	4.54±0.43
0.953 09	2.40±0.25
1.	1.97±0.14
ΔA (fit to 13 MC points)	7.81
ΔA (five-point Gaussian quadrature with $k=4$, method I)[b]	7.81
ΔA (five-point Gaussian quadrature with $k=4$, method II)[b]	7.80
ΔA (five-point Gaussian quadrature with $k=1$, method I)[b]	5.73
ΔA (five-point Gaussian quadrature with $k=1$, method II)[b]	8.23

[a] TI integrand and ΔA values are in kcal/mol. The estimated errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 3×10^6 steps, except for the run at $\lambda=0.046\,91$ with $k=4$, for which the run length is 6×10^6 steps.

[b] For the definition of methods I and II, see the text (Sec. IV B).

TABLE VII. Lennard-Jones solute.[a]

$\lambda^{1/k}$	$\mathcal{T}_{12}$	$\mathcal{T}_6$	$\mathcal{T}$
$k=1$			
0.046 91	25.33±2.02	−10.45±0.47	14.88±1.56
0.230 76	8.45±0.53	−7.25±0.24	1.20±0.30
0.500 00	4.91±0.40	−5.87±0.25	−0.96±0.17
0.769 24	4.67±0.35	−5.99±0.23	−1.32±0.14
0.953 09	3.99±0.30	−5.69±0.21	−1.71±0.10
$k=4$			
0.046 91	12 826 ± 257	−78.36±0.77	12 748 ± 256
0.230 76	164.89±6.82	−19.42±0.50	145.46±6.34
0.500 00	18.74±1.43	−9.33±0.39	9.41±1.06
0.769 24	6.85±0.54	−6.78±0.31	0.07±0.24
0.953 09	3.96±0.31	−5.60±0.21	−1.65±0.11
1.	3.79±0.37	−5.52±0.27	−1.73±0.10
ΔA (fit to 13 MC points)			3.09
ΔA (five-point Gaussian quadrature with $k=4$, method I)[b]			3.03
ΔA (five-point Gaussian quadrature with $k=4$, method II)[b]			3.03
ΔA (five-point Gaussian quadrature with $k=1$, method I)[b]			1.26
ΔA (five-point Gaussian quadrature with $k=1$, method II)[b]			3.91

[a] TI integrand and ΔA values are in kcal/mol. The estimated errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 3×10^6 steps, except for the run at $\lambda=0.046\,91$ with $k=4$, for which the run length is 6×10^6 steps.

[b] For the definition of methods I and II, see the text (Sec. IV B).

TABLE VIII. Uncharged dumbbell solute.[a]

$\lambda^{1/k}$	$\mathcal{T}_{12}$	$\mathcal{T}_6$	$\mathcal{T}$
$k=1$			
0.046 91	32.67±3.03	−16.57±0.81	16.11±2.24
0.230 76	10.67±1.02	−11.32±0.54	−0.64±0.51
0.500 00	7.11±0.58	−10.00±0.44	−2.89±0.19
0.769 24	6.12±0.38	−9.75±0.27	−3.63±0.14
0.953 09	5.37±0.31	−9.37±0.29	−4.00±0.10
$k=4$			
0.046 91	24 388 ± 414	−149.83±1.62	24 238 ± 413
0.230 76	248.8 ±17.1	−33.01±1.38	215.8 ±15.7
0.500 00	23.50±1.89	−14.54±0.60	8.97±1.33
0.769 24	9.08±0.85	−10.97±0.58	−1.90±0.32
0.953 09	5.86±0.27	−9.64±0.23	−3.78±0.10
1.	5.82±0.31	−9.85±0.25	−4.02±0.11
ΔA (fit to 13 MC points)			2.70
ΔA (five-point Gaussian quadrature with $k=4$, method I)[b]			2.62
ΔA (five-point Gaussian quadrature with $k=4$, method II)[b]			2.57
ΔA (five-point Gaussian quadrature with $k=1$, method I)[b]			−0.41
ΔA (five-point Gaussian quadrature with $k=1$, method II)[b]			2.40

[a] TI integrand and ΔA values are in kcal/mol. The estimated errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 3×10^6 steps, except for the run at $\lambda=0.046\,91$ with $k=4$, for which the run length is 6×10^6 steps.

[b] For the definition of methods I and II, see the text (Sec. IV B).

behavior, and the series expansion was truncated at $N_j=8$. A simple calculation, Eq. (30), gives $a_{12,0}=0.958$ kcal/mol for SS and LJ solutes, and 1.915 kcal/mol for UD and DD solutes. Similarly $a_{6,0}=-2.862$ kcal/mol for the LJ solute, and −5.723 kcal/mol for UD and DD solutes. Varying N_j between 6 and 11 did not have any considerable effect, and the change in solvation Helmholtz energies calculated with different N_j's was at most 0.05 kcal/mol for all the solute cases. These "correct" values were then compared to Gaussian quadrature integration results.

As discussed above, the use of $k=1$ results in divergent TI integrands. For this reason, five-point Gaussian quadrature with $k=1$ results were integrated in two different ways. The first way involves a *bold* application of the quadrature integration ignoring the divergence, which is obviously the wrong approach to integrate an improper function (in Tables VI–IX, the label for this approach is $k=1$, method I). In the second approach (with label $k=1$, method II) $\mathcal{T}$, the total TI integrand, was first multiplied with $\lambda^{3/4}$ to eliminate the divergence. The corresponding integrand was then generated using the five-point Gaussian quadrature fit polynomial. The resulting polynomial function was later multiplied by $\lambda^{-3/4}$ to account for the initial correction and then analytically integrated.

For $k=4$ results, we again used two different ways to perform the quadrature integral. The first way (with label $k=4$, method I) is the usual direct application of five-point Gaussian quadrature integration. The second approach uses the scheme reported in Sec. IV. To briefly summarize, the

TABLE IX. Dipolar dumbbell solute.[a]

$\lambda^{1/k}$	$\mathcal{T}_{12}$	$\mathcal{T}_6$	$\mathcal{T}_1$	$\mathcal{T}$
$k=1$				
0.046 91	33.52±2.36	−17.00±0.65	−0.41±0.43	16.11±1.87
0.230 76	12.21±0.94	−12.08±0.48	−2.90±0.45	−2.77±0.42
0.500 00	8.75±0.56	−10.87±0.36	−6.19±0.47	−8.31±0.39
0.769 24	8.12±0.41	−10.66±0.31	−9.37±0.45	−11.91±0.37
0.953 09	8.09±0.37	−10.56±0.29	−11.93±0.47	−14.40±0.40
$k=4$				
0.046 91	24 021 ± 409	−148.68±1.48	0.14±0.23	23 873 ± 408
0.230 76	242.1±16.6	−32.88±1.34	−0.19±0.29	209.0 ±15.4
0.500 00	27.85±2.37	−15.83±0.69	−1.16±0.45	10.86±1.69
0.769 24	10.32±0.86	−11.60±0.60	−4.74±0.44	−6.02±0.48
0.953 09	7.65±0.43	−10.29±0.34	−10.22±1.34	−12.85±0.36
1.	7.39±0.48	−10.23±0.39	−12.00±0.67	−14.85±0.58
ΔA (fit to 13 MC points)				−2.61
ΔA (five-point Gaussian quadrature with $k=4$, method I)[b]				−2.72
ΔA (five-point Gaussian quadrature with $k=4$, method II)[b]				−2.79
ΔA (five-point Gaussian quadrature with $k=1$, method I)[b]				−5.67
ΔA (five-point Gaussian quadrature with $k=1$, method II)[b]				−2.90

[a] TI integrand and ΔA values are in kcal/mol. The estimated errors are calculated using a method of batch means with batch sizes of 10^5 MC steps, and correspond to two standard deviations. The MC results are for run lengths of 3×10^6 steps, except for the run at $\lambda=0.046\ 91$ with $k=4$, for which the run length is 6×10^6 steps.

[b] For the definition of methods I and II, see the text (Sec. IV B).

first quadrature λ root point was replaced with $\lambda=0$, and the analytical results at $\lambda=0$ were utilized to obtain the equivalent modified Gaussian quadratures. The results of this approach is labeled as $k=4$, method II.

When compared to the solvation free energy results obtained using the set of all the MC simulation points and the analytical result at $\lambda=0$, the results (Tables VI–IX) for each studied solute clearly show that both five-point quadrature integration approaches with $k=4$ perform superbly. It is important to note that the exact solvation free energies are not known, and the results obtained by fitting to all MC simulation points can only be taken as the "best estimated" values. For this reason, small deviations from these results cannot be completely conclusive. Both approaches (methods I and II) with $k=4$ seem to be performing equally well, and the deviations from the correct results are much smaller than the statistical errors which are estimated to be about 0.5–0.8 kcal/mol depending on the solute and the exponent employed. This success is mainly due to the fact that the use of $k=4$ results in a better sampling of low λ range, and therefore, to the correct estimation of the divergent behavior. Another not so surprising outcome is that, if $k=1$, a straightforward application of (five-point) Gaussian quadratures (method I) performs very poorly. But the results of the second method with $k=1$ described above are considerably better which is ought to be expected, because the divergence of the integrand is accounted for from the beginning.

The overall outcome of this section reinforces earlier conclusions that in order to obtain accurate and reliable quantitative results for the solvation free energies using the TI method, it is very important to choose a thermodynamic path which avoids the singularity at $\lambda=0$. It is equally important to choose a fitting and integration scheme which can properly represent the integrand. As is shown with examples here, the analytical theory developed in this report can be combined with numerical simulation methods via a combination scheme introduced in Sec. IV to achieve improvements in correct characterization of the limiting behavior at small coupling parameters, and at the same time to achieve some savings in computational expenditure.

VII. SUMMARY AND DISCUSSION

Using the concepts of scaled particle theory, an analytical method is developed to study the limiting behavior of solvation free energies. The derivation is based on a perturbation-type treatment, and successfully incorporates the effects of the weakly attractive dispersion interaction between the solute and the solvent molecules. An approximate criterion for the upper bound of validity for the perturbation treatment is established in terms of the interaction potential parameters, and the possible shortcomings of the validity criterion are investigated.

In the first paper of the series,[7] we have shown that the use of a polynomial path TI (PP-TI) makes it possible to "linearize" the integrand hence lowering the integration errors in free energy simulation studies. The focus of this earlier work was the whole coupling parameter range, and it was shown that the curved nature of the TI integrand (and the solvation free energy as a function of λ) is mainly due to the Coulomb interactions. In this work, the limiting behavior of the solvation free energies is analytically determined. As is shown, the Coulomb term does not have any important role in the limiting behavior which is essentially determined by the shortest ranged interaction term. In this respect, this study complements our previous work. Therefore, combining the achievements of both works, i.e., using PP-TI to linearize the integrand and incorporating the analytical theory to obtain the correct limiting behavior, would further minimize the integration errors. Although only the equal k exponents are used in this report, the analytical theory is derived for an arbitrary combination of exponents, thus, it can be used as easily in the PP-TI approach which employs a mixed-exponent set.

The analytical theory also introduces an independent test for the convergence and the appropriateness of the simulation method employed, where the latter has been a major concern in free energy simulations.[1–6,8] As the comparison of the analytical theory and the numerical results indicate, the use of the thermodynamic integration method to calculate the solvation free energies is quite successful in accounting for the predicted divergence. This finding supports the previous observations that the thermodynamic integration method using (five-point) Gaussian quadratures with $k_{12} \geqslant 4$ is a robust, powerful, and economical method for studying the free energy of solvation. The theory developed in this report will make it even sounder by assuring the correct limiting behavior of the integrand. As is shown in Sec. IV, the analytical theory results at $\lambda=0$ can be used as the first root point of a quadrature integration, thereby further lowering the computational expenses. An equally useful application of the analytical theory would be the comparison of the numerical result at $\lambda=0$ determined by extrapolating the simulation results to that of the analytical result at $\lambda=0$. This comparison would form an independent test criterion for the convergence of such molecular simulations.

Comparison of different integration schemes shows that Gaussian quadrature with exponents $k_{12}=4$ and $k_\Delta>0$ is quite successful and should be the preferred choice. The poor performance of a bold application of Gaussian quadrature integration with exponents $k_{12}=k_6=1$ clearly shows the importance of properly accounting for the divergence of the TI integrand. Another important outcome of this study is that when creation of particles are involved, it is important to sample the low coupling parameter values with more weight. For example, the approaches $k=4$, method I and $k=1$, method II discussed in Sec. VI both take into account the divergent behavior at $\lambda=0$. The major difference between these two approaches is in the selection of the quadrature root points at which molecular simulations are performed. Comparison of these two approaches show that $k=4$, method I which samples the low λ range with more weight is more successful than the $k=1$, method II approach. This is an important point to consider in designing future simulations.

Although, the new analytical theory is incapable of incorporating the solute–solvent Coulomb interaction, using physical arguments and deducing from earlier results, some assumptions about the role of Coulomb interaction in the asymptotic limiting behavior are made and tested using molecular simulations. For this the aqueous solvation of a dipolar dumbbell solute and of alanine-dipeptide at two different conformations were studied. For charge neutral solute and solvent molecules, analytical theory predicts that the Coulomb term should not effect the asymptotic limiting behavior. The results of the studied cases confirm the predictions of the analytical theory, and thus justify the validity of its underlying assumptions in relation to the treatment of solute–solvent Coulomb interactions. Based on the results of the alanine-dipeptide study, when "coreless" proton sites are involved, or when strong orientational correlations may exist, we suggest to use the analytical theory only at extremely small values of the coupling parameter.

Note added in proof. After the work has been completed we came across Ref. 46 also dealing with the analysis of the singularity discussed in this report but using an alternative approach.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R55-GM43500. Computing resources were provided in part by the City University of New York, University Computing Center. It is our great pleasure to acknowledge helpful discussions with Chung F. Wong at several stages of this project.

APPENDIX A

1. Perturbation treatment formulation

In this appendix we give a detailed derivation of Eqs. (10) and (20). Starting with Eq. (5), and by substituting the partial derivative of the interaction potential, Eq. (7), the ensemble average of the exponential term can be written in terms of ensemble averages of solute–solvent interaction energy terms as

$$\langle e^{-\beta V}\rangle_0 = 1 - 4\pi\rho \int_0^\infty r^2\, dr \left[4\, \frac{\lambda^{k_{12}} C_{12}}{r^{12}} - 2\, \frac{\lambda^{k_6} C_6}{r^6} \right] e^{-\beta V}$$
$$= 1 - 4\lambda^{k_{12}} C_{12} \langle r^{-12} e^{-\beta V}\rangle_0 + 2\lambda^{k_6} C_6 \langle r^{-6} e^{-\beta V}\rangle_0. \tag{A1}$$

Subsequent variable changes to unitless $s \equiv r/r_c$ and $x \equiv 1/s^6$ gives $(m=6,12)$

$$\langle r^{-m} e^{-\beta V}\rangle_0 = \frac{\rho v_c}{2 r_c^m} \int_0^\infty dx\, x^{(m/6-3/2)} e^{-B_{12}x^2 + B_6 x}, \tag{A2}$$

where $B_{12} \equiv C_{12}\lambda^{k_{12}}/r_c^{12}$, $B_6 \equiv C_6 \lambda^{k_6}/r_c^6$, and $v_c \equiv 4\pi r_c^3/3$ is the volume of the spherical cavity. Note that $B_{12}=1$ if r_c is defined as in Eq. (9). But, to reserve generality of the derivation, we assume an arbitrary definition for r_c in the rest of this Appendix. Also note that B_{12} always appears in combination with the cavity volume as $v_c B_{12}^{1/4}$, and therefore, there is no explicit r_c dependence in any of the final results.

Let us define the generalized integral

$$I_\nu(B_{12}, B_6) \equiv \int_0^\infty dx\, x^{\nu-1} e^{-B_{12}x^2 + B_6 x}. \tag{A3}$$

Then Eq. (A1) may be written as

$$\langle e^{-\beta V}\rangle_0 = 1 - 2\rho v_c B_{12} I_{3/2}(B_{12}, B_6)$$
$$+ \rho v_c B_6 I_{1/2}(B_{12}, B_6). \tag{A4}$$

The results of the $I_\nu(B_{12}, B_6)$ integrals for $\mathrm{Re}[\nu]>0$ and for $B_{12}>0$ are given in terms of the parabolic cylinder functions, $D_{-\nu}(z)$, as [Ref. 27, p. 337, Eq. (3.462.1)]

$$I_\nu(B_{12},B_6)=(2B_{12})^{-\nu/2}\Gamma(\nu)e^{z^2/4}D_{-\nu}(z), \tag{A5}$$

where $z \equiv -B_6/\sqrt{2B_{12}}$, and $\Gamma(\nu)$ is the gamma function. The parabolic cylinder functions may be expressed in terms of the degenerate hypergeometric functions, $\Phi(\alpha,\gamma;z)$ [Ref. 27, p. 1064, Eq. (9.240)]

$$D_p(z)=2^{p/2}e^{-z^2/4}\left\{\frac{\sqrt{\pi}}{\Gamma[(1-p)/2]}\,\Phi\left(-\frac{p}{2},\,\frac{1}{2};\,\frac{z^2}{2}\right)\right.$$
$$\left.-\frac{\sqrt{2\pi}z}{\Gamma(-p/2)}\,\Phi\left(\frac{1-p}{2},\,\frac{3}{2};\,\frac{z^2}{2}\right)\right\}. \tag{A6}$$

Defining $y\equiv z^2/2=B_6^2/4B_{12}$, thus $z = -\sqrt{2y}$, and the subsequent substitution gives

$$I_{3/2}(B_{12},B_6)=\frac{\Gamma(3/4)}{2B_{12}^{3/4}}\left\{\Phi\left(\frac{3}{4},\,\frac{1}{2};y\right)\right.$$
$$\left.+\sqrt{y}\,\frac{\Gamma(1/4)}{2\Gamma(3/4)}\,\Phi\left(\frac{5}{4},\,\frac{3}{2};y\right)\right\},$$
$$I_{1/2}(B_{12},B_6)=\frac{\Gamma(3/4)}{2B_{12}^{1/4}}\left\{\frac{\Gamma(1/4)}{\Gamma(3/4)}\,\Phi\left(\frac{1}{4},\,\frac{1}{2};y\right)\right. \tag{A7}$$
$$\left.+2\sqrt{y}\Phi\left(\frac{3}{4},\,\frac{3}{2};y\right)\right\}.$$

Eliminating B_6 in favor of y, $B_6 = 2\sqrt{yB_{12}}$, then using Eqs. (A7) in Eq. (A4), and utilizing the relation [Ref. 27, p. 1058, Eq. (9.212.3)]

$$(\alpha-\gamma)\Phi(\alpha,\gamma+1;y)=\alpha\Phi(\alpha+1,\gamma+1;y)-\gamma\Phi(\alpha,\gamma;y), \tag{A8}$$

we obtain

$$\langle e^{-\beta V}\rangle_0=1-\rho v_c\Gamma(3/4)B_{12}^{1/4}\left\{\Phi\left(\frac{3}{4},\,\frac{1}{2};y\right)\right.$$
$$\left.-2y\Phi\left(\frac{3}{4},\,\frac{3}{2};y\right)-\sqrt{y}\,\frac{\Gamma(1/4)}{2\Gamma(3/4)}\,\Phi\left(\frac{1}{4},\,\frac{3}{2};y\right)\right\}. \tag{A9}$$

Using the expression of y in terms of the coupling parameter λ:

$$y=\frac{B_6^2}{4B_{12}}=\frac{C_6^2}{4C_{12}}\,\lambda^{2k_6-k_{12}}, \tag{A10}$$

and defining $k_\Delta=k_6-k_{12}/2$, it is obvious that, as long as $k_\Delta>0$, small λ corresponds to small y. For a small argument, the series expansion of the degenerate hypergeometric functions [Ref. 27, p. 1058, Eq. (9.210.1)]

$$\Phi(\alpha,\gamma;y)=1+\frac{\alpha}{\gamma}\frac{y}{1!}+\frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{y^2}{2!}$$
$$+\frac{\alpha(\alpha+1)(\alpha+2)}{\gamma(\gamma+1)(\gamma+2)}\frac{y^3}{3!}+\cdots \tag{A11}$$

will prove to be useful. Making use of the expansion for the degenerate hypergeometric functions, as given above, we derive a series expansion for the required ensemble average

$$\langle e^{-\beta V}\rangle_0=1-\rho v_c\Gamma(3/4)B_{12}^{1/4}\sum_{n=0}^{\infty}\xi_n\left(\frac{C_6}{2\sqrt{C_{12}}}\,\lambda^{k_\Delta}\right)^n=1$$
$$-a_0C_{12}^{1/4}\lambda^{k_{12}/4}\sum_{n=0}^{\infty}\xi_n\left(\frac{C_6}{2\sqrt{C_{12}}}\,\lambda^{k_\Delta}\right)^n, \tag{A12}$$

with $a_0\equiv4\pi\rho\Gamma(3/4)/3$, $\Gamma(3/4)=1.225\,42$, and $\Gamma(1/4)/\Gamma(3/4)=2.958\,67$. A few leading ξ coefficients are

$$\xi_0=1,\quad \xi_1=-\frac{1}{2}\frac{\Gamma(1/4)}{\Gamma(3/4)},$$
$$\xi_2=-\frac{1}{2},\quad \xi_3=-\frac{1}{12}\frac{\Gamma(1/4)}{\Gamma(3/4)},$$
$$\xi_4=-\frac{1}{8},\quad \xi_5=-\frac{1}{48}\frac{\Gamma(1/4)}{\Gamma(3/4)}.$$

For the $k_\Delta\leqslant0$ case, the generalized integrals $I_\nu(B_{12},B_6)$ are again finite, and their asymptotic expressions can be obtained (see, e.g., Sec. 19 of Ref. 28, or Sec. 9.246 of Ref. 27). But these asymptotic expressions, and the resulting expressions for $\langle e^{-\beta V}\rangle_0$ and for $\mathcal{T}$, are relatively complicated. Since our main interest is in using exponent sets for which $k_\Delta>0$, the $k_\Delta\leqslant0$ case is not pursued in this report.

2. Thermodynamic integration integrand

To calculate the numerator factors in Eqs. (19), the ensemble averages $\langle r^{-m}e^{-\beta V}\rangle_0$ with $m=6$ and 12 are needed. These averages are already calculated in the first part of this Appendix. After some straightforward algebra, the numerators in Eq. (19b), $\mathcal{N}_{12}$ and $\mathcal{N}_6$, can be expressed as

$$\mathcal{N}_{12}=\frac{k_{12}}{2\lambda}\,\rho v_cB_{12}I_{3/2}(B_{12},B_6),$$
$$\mathcal{N}_6=\frac{k_6}{2\lambda}\,\rho v_cB_6I_{1/2}(B_{12},B_6). \tag{A13}$$

For small y, the series expansion of the generalized integrals, Eqs. (A7), with $r\equiv k_6/k_{12}$ give ($m=6,12$)

$$\mathcal{N}_m=\frac{k_{12}}{4\lambda}\,\rho v_cB_{12}^{1/4}\Gamma(3/4)\sum_{n=0}^{\infty}\zeta_{m,n}\left(\frac{C_6}{2\sqrt{C_{12}}}\,\lambda^{k_\Delta}\right)^n, \tag{A14}$$

where $\zeta_{12,n}=(1-2n)\xi_n$ and $\zeta_{6,n}=-4nr\xi_n$.

APPENDIX B

This appendix deals with the purely repulsive potentials of the form $\beta V = \lambda^{k_n}C_n/r^n$ for $n>3$. As in the case of the 6-12 potential (Appendix A), the following variable changes simplify the derivation: $B_n \equiv \lambda^{k_n}C_n/r_c^n$, $s\equiv r/r_c$, and $x\equiv s^{-n/2}$. As in Appendix A, $B_n=1$ if the cavity radius is determined by using the criterion $\beta V=1$. To reserve generality of the derivation, we again assume an arbitrary definition for r_c in the rest of this Appendix. Following the steps of Appendix A, it is straightforward to show that

$$\langle e^{-\beta V}\rangle_0 = 1 - 2\rho v_c B_n \int_0^\infty dx\, x^{(1-6/n)} e^{-B_n x^2}$$
$$= 1 - 2\rho v_c B_n I_{2-6/n}(B_n, 0), \qquad \text{(B1)}$$

where we have used the definition of integrals $I_\nu(B_{12}, B_6)$ given in Eq. (A3) with $B_{12} = B_n$ and $B_6 = 0$. Therefore $z = 0$, and hence, the final equations simplify considerably to

$$\langle e^{-\beta V}\rangle_0 = 1 - \rho v_c B_n^{3/n} \frac{2\sqrt{\pi}}{2^\gamma} \frac{\Gamma(\gamma)}{\Gamma[(\gamma+1)/2]}, \qquad \text{(B2)}$$

or equivalently to

$$\langle e^{-\beta V}\rangle_0 = 1 - \frac{4\pi\rho}{3} (\lambda^{k_n} C_n)^{3/n} \frac{2\sqrt{\pi}}{2^\gamma} \frac{\Gamma(\gamma)}{\Gamma[(\gamma+1)/2]}, \qquad \text{(B3a)}$$

with

$$\gamma \equiv 2 - \frac{6}{n}. \qquad \text{(B3b)}$$

Similarly, one can also easily derive that [see Eqs. (19)]

$$\beta \mathcal{F} = \frac{1}{\langle e^{-\beta V}\rangle_0} \times \frac{3k_n}{n\lambda} \rho v_c B_n^{3/n} \frac{2\sqrt{\pi}}{2^\gamma} \frac{\Gamma(\gamma)}{\Gamma[(\gamma+1)/2]}. \qquad \text{(B4)}$$

---

[1] M. Mezei and D. L. Beveridge, Ann. Acad. Sci. (N.Y.) **482**, 1 (1986).

[2] D. L. Beveridge and F. M. DiCapua, Annu. Rev. Biophys. Chem. **18**, 431 (1989).

[3] W. L. Jorgensen, Acc. Chem. Res. **22**, 184 (1989).

[4] M. Karplus and G. A. Petsko, Nature **347**, 631 (1990).

[5] P. M. King, C. A. Reynolds, J. W. Essex, G. A. Worth, and W. G. Richards, Mol. Simul. **5**, 265 (1990).

[6] T. P. Straatsma and J. A. McCammon, Annu. Rev. Phys. Chem. **43**, 407 (1992).

[7] H. Resat and M. Mezei, J. Chem. Phys. **99**, 6052 (1993).

[8] D. A. Pearlman, J. Comput. Chem. **15**, 105 (1994).

[9] (a) A. Ben-Naim, *Hydrophobic Interactions* (Plenum, New York, 1980); (b) C. Tanford, *The Hydrophobic Effect* (Wiley, New York, 1973).

[10] H. Reiss, H. L. Frisch, and J. L. Lebowitz, J. Chem. Phys. **31**, 369 (1959); H. Reiss, H. L. Frisch, E. Helfand, and J. L. Lebowitz, *ibid.* **32**, 119 (1960); E. Helfand, H. Reiss, H. L. Frisch, and J. L. Lebowitz, *ibid.* **33**, 1379 (1960).

[11] R. A. Pierotti, Chem. Rev. **76**, 717 (1976).

[12] F. H. Stillinger, J. Solut. Chem. **2**, 141 (1973).

[13] R. M. Gibbons, Mol. Phys. **17**, 81 (1969).

[14] T. Boublik, Mol. Phys. **27**, 1415 (1974); Adv. Chem. Ser. **204**, 173 (1983).

[15] J. P. M. Postma, H. J. C. Berendsen, and J. R. Haak, Faraday Symp. Chem. Soc. **17**, 55 (1982).

[16] A. Ben-Naim, K. L. Ting, and R. L. Jernigan, Biopolymers **28**, 1309 (1989).

[17] M. Irisa, K. Nagayama, and F. Hirata, Chem. Phys. Lett. **207**, 430 (1993).

[18] C. Chothia, Nature (London) **248**, 338 (1974).

[19] D. Eisenberg and A. D. McLachlan, Nature (London) **319**, 199 (1986).

[20] T. Ooi, M. Oobatake, G. Nemethy, and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 3086 (1987).

[21] C.-L. Lin and R. H. Wood, J. Comput. Chem. **15**, 149 (1994).

[22] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).

[23] C. G. Gray and K. E. Gubbins, *Theory of Molecular Fluids* (Oxford University, New York, 1984), Vol. 1.

[24] D. A. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976), p. 247.

[25] M. Mezei, J. Comput. Chem. **13**, 651 (1992); Mol. Simul. **10**, 225 (1993).

[26] It is easy to show, for some typical values of C_{12} and C_6, and for the range of λ considered in this report, that the difference between the cavity radius values calculated using Eq. (9) and using $\beta V \equiv 1$ is less than 0.1%. Therefore the choice made for the definition of r_c is not important.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. (Academic, San Diego, 1980).

[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, 9th printing (Dover, New York, 1972).

[29] The scaling factor also depends on how the effective cavity radius is defined. If the definition $\beta V = 1$ is not used, then $B_n \neq 1$, and Eq. (14) needs to be rewritten in terms of Eq. (B3) which is independent of the cavity radius definition.

[30] To be consistent with the underlying assumptions, one must consider the closest distance that any two sites of distinct solvent molecules can occupy in calculating the upper bound $\lambda_{\max}$. For example, if the hydrogen sites of water molecules also interact with the solute molecule via short range interaction forces, then the closest distance that the hydrogen site of a water molecule can come closer to an oxygen or to a hydrogen site of another water molecule has to be taken into account, too. Using the available experimental and theoretical water structures, this closest distance can be estimated to be around 1.5 Å, which is shorter than the closest positions that the oxygens of two water molecules can occupy. Therefore, if the hydrogen sites of water molecules also involve in the interaction with the solute, the upper bound of validity for the analytical theory would be $r_c \approx 0.75$ Å.

[31] J. G. Kirkwood, *Theory of Liquids*, edited by B. J. Alder (Gordon and Breach, New York, 1968).

[32] B. Lee and F. M. Richards, J. Mol. Biol. **55**, 379 (1971); F. M. Richards, Annu. Rev. Biophys. Bioeng. **6**, 151 (1977).

[33] T. J. Richmond, J. Mol. Biol. **178**, 63 (1984).

[34] R. Lustig, Mol. Phys. **55**, 305 (1985); **59**, 195 (1986).

[35] L. R. Dodd and D. N. Theodorou, Mol. Phys. **72**, 1313 (1991).

[36] (a) H. L. Friedman, *Ionic Solution Theory* (Wiley, New York, 1962); (b) G. Stell and J. L. Lebowitz, J. Chem. Phys. **49**, 3706 (1968); (c) H. C. Andersen and D. Chandler, *ibid.* **57**, 1918 (1972); D. Chandler and H. C. Andersen, *ibid.* **57**, 1930 (1972).

[37] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Chem. Phys. **79**, 926 (1983).

[38] C. Pangali, M. Rao, and B. Berne, J. Chem. Phys. **71**, 2975, 2982 (1979).

[39] R. H. Kincaid and H. A. Scheraga, J. Comput. Chem. **3**, 525 (1982).

[40] See the extensive list of references given in Refs. 1–8.

[41] The value of φ_{pep} is obtained by summing $C_{12}^{1/4}$ values for the dipeptide sites–water oxygen interaction over the dipeptide interaction sites, and then taking the ratio with respect to the $C_{12}^{1/4}$ for the LJ solute–water oxygen interaction. As in UD and DD solutes, this assumes that the solute site cavities do not overlap, and therefore, it is a conservative estimate for the upper bound.

[42] M. Mezei, Mol. Simul. **2**, 201 (1988); Mol. Phys. **47**, 1307 (1982) [Erratum: **67**, 1205 (1989)].

[43] M. R. Mruzik, F. F. Farid, F. Abraham, D. E. Screiber, and G. M. Pound, J. Chem. Phys. **64**, 481 (1976). For a similar earlier application see D. R. Squire and W. G. Hoover, *ibid.* **50**, 701 (1969).

[44] For the transition ideal gas→full solute, the use of exponent k in Eq. (7) is arbitrary. This is due to the fact that a variable change to λ^* can be made such that the defining path would have a linear λ^* dependence. The only difference the use of $k \neq 1$ makes is in the choice of the n-root points defining the n-point Gaussian quadratures. Therefore, the second set of simulations reported in Tables VI–IX are equivalent to simulations using $k_{12} = k_6 = k_1 = 1$ at λ' points with $\lambda' = \lambda^4$.

[45] Simulation results with $k = 4$ were first scaled to obtain the equivalent $k = 1$ results, as discussed in Ref. 44, then Eq. (35) was subsequently used. For the DD solute, the Coulomb term contribution $\mathcal{F}_1$ was fitted to the same functional form of $\mathcal{F}_6$ with the constant term, $a_{6,0}$, set equal to zero.

[46] T. Simonson, Mol. Phys. **80**, 441 (1993).