The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water

M. Mezei
Chemistry Department and Center for Study of Gene Structure and Function, Hunter College–CUNY, New York, New York 10021

(Received 22 December 1986; accepted 10 March 1987)

A new technique is proposed to compute by Monte Carlo (or molecular dynamics) computer simulation the hydration free energy differences. The method, called finite difference thermodynamic integration, is a combination of the thermodynamic integration and the perturbation method. It was compared with thermodynamic integration over two different paths and the perturbation method on computing the solvation free-energy difference between the dilute aqueous solution of acetone and dimethyl amine. Finite difference thermodynamic integration was found to have the best convergence characteristics among the methods tested.

I. INTRODUCTION

The calculation of free energy from computer simulations presents special difficulties since the free energy is directly related to the partition function and the simulations avoid computing the partition function. Several methods have been proposed to circumvent this difficulty that have been periodically reviewed in recent years.[1–6] Two methods gained relatively wide usage: the thermodynamic integration and the so-called perturbation method. The purpose of this paper is to develop a variant of the thermodynamic integration where individual quadrature points are evaluated using the perturbation method. This combined technique will be compared with two different forms of thermodynamic integration and with the perturbation method to allow an assessment of its performance. As a byproduct of these calculations a comparison between thermodynamic integration and the perturbation method will be obtained as well.

II. BACKGROUND

The common feature of the methods calculating free energy differences between two systems is that a path has to be defined that connects the two systems in the configuration space. In most cases a coupling parameter λ is introduced that varies from 0 to 1 and it continuously brings the initial system 0 into the final system 1. Clearly, in general several such paths exist. The thermodynamic integration (TI), originated from Kirkwood[7] considers the free energy as the function of λ and employs the fundamental theorem of calculus

$$A_1 - A_0 = \Delta A \int_0^1 \frac{\partial A(\lambda)}{\partial \lambda}\, d\lambda, \tag{1}$$

where

$$A(\lambda) = -kT \ln\left\{\int \cdots \int \exp[-E_N(\lambda)/kT]\, dV^N\right\}, \tag{2}$$

where k is the Boltzmann factor and T is the absolute temperature. Substituting Eq. (2) into Eq. (1) gives

$$\Delta A = \int_0^1 \left\langle \frac{\partial E_N(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda, \tag{3}$$

where the subscript λ implies a Boltzmann average based on the function $E_N(\lambda)$. In the following the subscript N will be dropped for simplicity. Most applications to date used an important simple particular case where $E(\lambda)$ is linear in λ:

$$E(\lambda) = (1-\lambda)E_0 + \lambda E_1 = E_0 + \lambda \Delta E \tag{4}$$

giving the free energy difference as

$$\Delta A = \int_0^1 \langle \Delta E \rangle_\lambda d\lambda. \tag{5}$$

Equation (3) or (5) is evaluated by carrying out a series of simulations corresponding to successive discrete values of λ. The final integration over λ is carried out by numerical integration.

Equation (4) has been generalized by Mezei and Beveridge to decrease the rate of coupling of newly created particles[6]:

$$E(\lambda) = (1-\lambda)^k E_0 + \lambda^k E_1, \tag{6}$$

defining a path that is nonlinear in λ. Use of TI on the path of Eq. (6) will be called k th degree TI. The integrand of Eq. (3) corresponding to the k th degree TI has been derived[6] as

$$-k(1-\lambda)^{k-1}E_0 + k\lambda^{k-1}E_1. \tag{7}$$

k th degree TI has been used recently to determine the free energy difference between the A and B forms of a DNA tetramer duplex.[8]

It is also possible to transform linearly the coordinates of the molecule(s) changed along with the potential coefficients, resulting in another nonlinear path [the dependence of $E(\lambda)$ is nonlinear in λ]. For this path, at each λ the coordinates of the "hybrid molecule" are obtained as

$$\mathbf{R}(\lambda) = (1-\lambda)\mathbf{R}^{\mathrm{A}} + \lambda \mathbf{R}^{\mathrm{B}}, \tag{8}$$

where $\mathbf{R}^{\mathrm{A}}$ and $\mathbf{R}^{\mathrm{B}}$ are the coordinates of the original molecules A and B, respectively. Similar coupling is performed on the solute–solvent potential coefficients. Couplings of the type of Eq. (8) have been proposed by Mezei and Beveridge,[6] Cross,[9] and tested by Cross on a model system.[9]

0021-9606/87/127084-05$02.10

An alternative expression for ΔA can be obtained by evaluating Eq. (2) at $\lambda = 0$ and at $\lambda = 1$ and inserting one as $\exp(E_0/kT)\exp(-E_0/kT)$. This leads to the equation[10,11]

$$\Delta A = -kT \ln \langle \exp[-(E_1 - E_0)/kT] \rangle_0. \qquad (9)$$

Reversing the role of systems 0 and 1 the mirror expression of Eq. (9) is obtained:

$$\Delta A = kT \ln \langle \exp[-(E_0 - E_1)/kT] \rangle_1. \qquad (10)$$

Bennett called Eqs. (9) and (10) an infinite order perturbation expansion.[10] Use of these equations have been also referred to as the perturbation method (PM) by Lybrand, Wipf, and McCammon[12] and by Jorgensen and Ravimohan[13] in their recent work using the PM since E_1 and E_0 differ by a small "perturbation" only in successful application. They do not, however, correspond to a "perturbation theory" in the usual sense of the word since they are exact expressions.

The use of the PM involves essentially a simulation carried out based on the energy function E_0 (or E_1), in which E_1 (or E_0) is also computed at each step and the average of $\exp[\mp(E_1 - E_0)/kT]$ is formed. Successful numerical calculation of the ensemble average of $\exp[\mp(E_1 - E_0)/kT]$ via simulations requires that states 0 and 1 be not too dissimilar. In the event of difficulties with the direct application of Eqs. (9) and (10) it is possible to use the coupling parameter approach to define a numerically viable path between states 0 and 1 [using, for example, Eq. (8)] involving intermediate states that are sufficiently close to each other and compute the free energy difference as

$$\Delta A = \sum_i -kT \ln \langle \exp[-(E_{i+1} - E_i)/kT] \rangle_i. \qquad (11)$$

The advantage of TI lies in the fact that it is based on the ensemble average of the energy (and possibly its derivative) that are known to converge relatively fast and it is able to take interpolate between the state points chosen. Furthermore, for the case of linear λ dependence, the integrand in Eq. (5) is a monotonic function of λ[6,14] making the interpolation inherent in the numerical integration rather reliable. Its main weakness lies in the possible divergence of the integral for certain (mostly linear) paths,[6,15] necessitating the use of nonlinear paths.[6,9] This, at least in principle, would make the interpolation less reliable since the guaranteed monotonicity of the integrand is lost.

The PM is very appealing due to its conceptual simplicity. For systems that are very similar, the free-energy difference can be obtained from a single simulation.[12] However, Eqs. (9) and (10) are very sensitive to fluctuations in $E_1 - E_0$, requiring in general several intermediate states due to the presence of the exponential in the quantity to be averaged requires that the "neighboring" states be rather close.

A third approach to calculate free-energy differences is based on the fundamental relation

$$A(\lambda) = -kT \ln P(\lambda). \qquad (12)$$

It has been used to compute free energy differences first on a conformational problem.[16] Currently calculations are under way to test its efficiency on chemical changes.

The proposed method, called finite difference thermodynamic integration (FDTI), is a combination of TI and PM. Write the free-energy difference between two states as the function of the coupling parameters:

$$A(\lambda_i) - A(\lambda_j) = A(\lambda_i, \lambda_j - \lambda_i). \qquad (13)$$

Clearly,

$$\Delta A = \sum_{i=1}^{n-1} A(\lambda_i, \lambda_{i+1} - \lambda_i) \qquad (14)$$

for $\lambda_1 = 0 < \lambda_2 < \cdots < \lambda_n = 1$. By multiplying and dividing each term by $\lambda_{i+1} - \lambda_i$ and decreasing the difference between the successive λ_i's, Eq. (1) is recovered through the definition of the derivative and the definition of the integral. Furthermore, substituting Eq. (9) into Eq. (13) and letting the differences between the successive λ_i's become arbitrarily small leads to

$$A(\lambda, \Delta) \rightarrow \langle E(\lambda, \Delta\lambda) \rangle_\lambda, \qquad (15)$$

where $E(\lambda, \Delta\lambda)$ is defined analogously to $A(\lambda, \Delta\lambda)$. Thus Eq. (3) can also be recovered at the small $\Delta\lambda$ limit. The suggested FDTI would evaluate $A(\lambda, \Delta\lambda)$ using Eqs. (9) and (10) for a small $\Delta\lambda$ at selected λ_i values and extrapolate for the rest of the interval. This extrapolation is most conveniently done if the λ_i values are selected by some quadrature prescription and the quadrature is used to evaluate the integral in Eq. (1). In this case the computed $A(\lambda, \Delta\lambda)$ values are divided by $\Delta\lambda$ to provide finite difference approximation to the integrand of Eq. (1). It is also possible (although not done in this work) to compute $A(\lambda, \Delta\lambda)$ with several different $\Delta\lambda$ values to give an estimate of the higher order derivatives of $A(\lambda)$ that would improve the precision of the numerical integration.

For larger $\Delta\lambda$, FDTI is related to the graphic interpolation idea of Bennett.[10] In the limit of infinitesimal $\Delta\lambda$, the nonlinear TI is recovered. The conceptual advantage of FDTI over a general nonlinear TI[6,9] lies in the fact that integration in general smoothes a function; therefore, local fluctuations and the nonmonotonicity of the integrand of Eq. (3) are likely to cause less damage. Also, the analytical evaluation of the λ derivatives is eliminated, thereby saving programming time. Finally, the author suspects that the convergence characteristics of $A(\lambda, \Delta\lambda)$ are likely to be better than those of $\langle \partial E(\lambda)\rangle/\partial\lambda \rangle$, although this remains to be demonstrated.

The PM calculation described here between the states $E(\lambda_i)$ and $E(\lambda_{i+1})$ used a reference state with $[E(\lambda_i) + E(\lambda_{i+1})]/2$. Simultaneous use of Eqs. (9) and (10) gave the free-energy difference $A(\lambda_{i+1}) - A(\lambda_i)$. This technique is formally equivalent[6] to the half-umbrella sampling by Scott and Lee[17] who arrived at the same expression by using Eq. (9) only (that is, looking for averages in the $\langle\,\rangle_\lambda$ ensemble) but employing non-Boltzmann (umbrella) sampling with a weighting function $\exp\{-[E(\lambda_{i+1}) - E(\lambda_i)]/2kT\}$.

An alternative to the half-umbrella sampling computes $A(\lambda_{i+1}) - A(\lambda_i)$ by running a simulation using $E(\lambda')$ where λ' falls between λ_i and λ_{i+1} and using Eqs. (9) and (10).[14] Assuming, however, that both $[E(\lambda_i) + E(\lambda_{i+1})]/2$ and $E(\lambda')$ sample the configuration space adequately, the half-umbrella sampling would in general provide smaller energy differences than the use of $E(\lambda')$ since

$$E(\lambda_{i+1}) - E(\lambda') + E(\lambda') - E(\lambda_i) = E(\lambda_{i+1}) - E(\lambda_i), \tag{16}$$

$$|E(\lambda_{i+1}) - E(\lambda')| + |E(\lambda') - E(\lambda_i)| \geqslant 2|[E(\lambda_{i+1}) - E(\lambda_i)]/2|. \tag{17}$$

Also,

$$\max[|E(\lambda_{i+1}) - E(\lambda')|, |E(\lambda') - E(\lambda_i)|] \geqslant |[E(\lambda_{i+1}) - E(\lambda_i)]/2|. \tag{18}$$

Therefore, the half-umbrella sampling is expected to have better convergence properties (although the difference is probably rather small). Also, the use of the intermediate state λ' requires an additional solute–solvent energy term.

The FDTI method was tested and compared with linear and nonlinear TI and with the PM on computing the difference between the solvation free energy of the aqueous solution of acetone and dimethylamine. In problems of this kind, the total energy of the system, $E_N(\lambda)$, can be separated into two terms: one for the solute–solvent interactions and another, much larger term for the solvent–solvent interactions. Since the solvent–solvent term is independent of λ, it will drop from consideration whenever $E_1 - E_0$ or $\partial E(\lambda)/\partial\lambda$ is formed. Therefore, all the methods discussed above are expected to have good convergence characteristics.

The test system chosen involves significant reorganization of the solvent around the solute during the transition between the two states, necessitating the use of several intermediate states. It is an interesting problem since the free-energy difference between the two systems models the relative contributions of the $>$CO and $>$NH groups to the hydration free energy of biomolecules. Successful reproduction of the experimental result would increase the confidence in the accuracy of computer simulation results on aqueous solution of peptides (using this potential).

III. CALCULATIONS

The Metropolis algorithm,[18] modified by incorporating the force bias procedure[19] and preferential sampling[20] for convergence acceleration, shown to have a synergistic effect,[21] was employed. The system for study in each case was comprised of 216 rigid molecules; one solute and 215 water molecules. The size of the simulation cell was chosen based on the experimentally determined partial molar volumes of acetone, dimethylamine and water, 67,[22] 61,[23] and 18.02 ml/mol, respectively, the average of the two solute molar volumes, 64 ml/mol was used. The condensed phase environment of the simulation cell was provided by face-centered cubic periodic boundary conditions. The solute–water interactions were treated under the minimum image convention while a 7.75 Å spherical cutoff was applied on the water–water interactions. Individual simulations involved at least 1000 K configurations (after equilibration). Calculations of this length with the force biased technique are equivalent to 2000–4000 K configurations generated by the regular Metropolis method.[19,24]

All simulations were based on the OPLS functions developed by Jorgensen and co-workers for the solute–water interactions[25,26] and the TIP4P water–water potential.[27,28] The OPLS potentials contain 6–12 interactions between the solute atoms and the water oxygen and electrostatic terms between the solute atoms and water hydrogens as well as between the solute atoms and a negative center near the water oxygen. Methyl groups are represented by a single center (united atom representation). In this calculation the methyl group charge on the dimethylamine was chosen to be 0.1 to maintain electroneutrality—a necessary, but to some extent arbitrary, deviation from the original OPLS parameter set. The molecular geometry used for acetone is from Herzberg[29] and for dimethylamine from Beagley and Hewitt.[30]

The free-energy calculations were performed along three different paths. A five-point Gaussian quadrature was used on the linear path [Eq. (4)] to evaluate Eq. (5). The function $A(\lambda,\Delta\lambda)$ was also estimated on a 0.02 wide interval around the quadrature points selected. Next, third degree TI calculations were carried out using the same quadrature as for the linear path. On the third path, defined by Eq. (8), both PM and FDTI calculations were performed. The FDTI calculations again used the same five λ values as the linear TI. The PM calculations used as wide intervals as appeared possible. Two consistency checks were applied on the PM calculations: (1) The expectation values $\langle E(\lambda_i)\rangle_\lambda$ were computed for the endpoints λ_i of the PM intervals from both runs[14] whose endpoint is λ_i and (2) for some PM calculations, the interval was split up two and independent calculations were performed on both segments whose sum should give the same result as on the full interval.

IV. RESULTS AND DISCUSSION

The results using the TI and FDTI with the five-point Gaussian quadrature are collected in Tables I and II. Plots of the computed $E_1 - E_0$ and $A(\lambda,\Delta\lambda)$ (not shown) indicate that both have an inflexion point, thus the minimum number of quadrature point to reproduce the correct shape of the integrand is 4 in this case. Table I gives the results at the end of 500, 1000, and 2000 K (1 K represents 1000 attempted moves) on the linear path to give an assessment of the convergence of the results. The results of FDTI calculations on the nonlinear paths described by Eq. (8) and of the third

TABLE I. Linear TI results. c_i: quadrature coefficients. The numbers in parentheses give the change in the contribution to the integral as the number of steps were doubled. All energies are in kcal/mol. $\lambda = 0$: acetone, $\lambda = 1$: dimethylamine.

		$E_1 - E_0$		
λ_i	c_i	500 K	1000 K	2000 K
0.046 910	0.118 463	13.63	14.86(0.14)	14.43(0.05)
0.230 765	0.239 314	8.45	10.21(0.42)	9.49(0.17)
0.5	0.284 444	2.69	3.04(0.10)	3.49(0.07)
0.769 235	0.239 314	− 2.01	− 1.87(0.03)	− 2.45(0.14)
0.953 089	0.118 463	− 17.29	− 14.27(0.36)	− 13.40(0.07)
ΔA:		1.87	2.93	2.80

TABLE II. Nonlinear TI results. All energies are in kcal/mol. The numbers in parentheses give the change in the contribution to the integral as the number of steps were doubled. $\lambda = 0$: acetone, $\lambda = 1$: dimethylamine.

	FDTI $A(\lambda_i + 0.01) - A(\lambda_i - 0.01)$		Third degree TI TI integrand	
λ_i	500 K	1000 K	500 K	1000 K
0.046 910	0.2712	0.2727(0.0002)	42.36	44.08(0.20)
0.230 765	0.2046	0.2060(0.0004)	16.33	16.90(0.14)
0.5	0.0605	0.0624(0.0005)	− 0.81	− 1.01(0.06)
0.769 235	− 0.0722	− 0.0767(0.0011)	− 13.48	− 13.73(0.06)
0.953 089	− 0.3232	− 0.3181(0.0006)	− 30.25	− 32.55(0.27)
ΔA:	2.14	2.17	1.89	1.84

degree TI calculations at 500 and 1000 K are shown in Table II.

The free-energy difference is obtained as 2.80, 1.84, 2.17, and 2.11 kcal/mol (favoring acetone) using linear TI, third degree TI, FDTI, and the PM, respectively. The value obtained by the FDTI displayed the most stable behavior when the intermediate results are compared with the final results on each of the contributing simulations, as demonstrated by the change in the contributions to the integrand from the individual simulations as the number of steps is doubled (also shown in Tables I and II). The magnitude of these changes indicate also that the uncertainty of the estimated integral is larger than one would estimate it based on the changes in the integral alone. The convergence characteristics of the linear TI are the worst, in accordance with the fact that the value obtained with linear TI differs most from the others.

The various PM calculations along the nonlinear path are described in Table III. Again, the convergence can be assessed by comparing the results at 500 and 1000 K. The polynomial fitting obtained from the FDTI calculation was integrated for the various λ intervals and also displayed in Table III for comparison. The values obtained from integrating the Gaussian fit generally agree with the values obtained from the PM calculation within the observed uncertainties.

The comparison of the computed $\langle E(\lambda)\rangle_\lambda$ at the end of the various λ intervals used in PM calculations show that they are a rather stringent indicator of the adequacy of sampling: Really good agreement was found only at $\lambda = 0.95$ between the (0.95,1.0) and (0.9,0.95) runs and at $\lambda = 0.6$ between the (0.3,0.6) and (0.1,0.3) runs. The comparison at $\lambda = 0.9$ is particularly instructive: the agreement with the value from the (0.9, 0.95) run progressively gets worse as the interval is lengthened.

The $\Delta E_{\max}$ values listed give the difference between the smallest and largest sampled values of $E(\lambda_{i+1}) - E(\lambda_i)$. It follows then that the smallest and largest terms in the Monte Carlo average of Eq. (6) differ by a factor of $\exp(\Delta E_{\max}/2kT)$. The PM run between 0.8 and 1.0 shows that $\Delta E_{\max}$ values over 12 kcal/mol introduce large fluctuations into the Monte Carlo average and thus are to be avoided.

The differences between the computed $A(\lambda,\Delta\lambda)$ and $E(\lambda,\Delta\lambda)$ values (not shown) were consistently below 1% for both paths, showing that in this system Eq. (11) holds to

TABLE III. Perturbation method calculations. λ_-, λ_+ are the endpoint coupling parameter values. ΔA_{+-} is the free-energy difference computed between λ_+ and λ_-. $U(\lambda_-)$ is $\langle E(\lambda_-)\rangle_\lambda$ and $U(\lambda_+)$ is $\langle E(\lambda_+)\rangle_\lambda$; $\Delta E_{\max}$ is the maximum of the sampled $|E_1 - E_0|$ values. ΔA_{int} is the free-energy difference for the interval obtained by integration, the polynomial of the Gaussian quadrature computed from FDTI. All energies are in kcal/mol. $\lambda = 0$: acetone, $\lambda = 1$: dimethylamine.

		500 K results				1000 K results				
λ_-	λ_+	$U(\lambda_-)$	$U(\lambda_+)$	$\Delta E_{\max}$	ΔA_{+-}	$U(\lambda_-)$	$U(\lambda_+)$	$\Delta E_{\max}$	ΔA_{+-}	ΔA_{int}
0.0	0.1	− 17.12	− 14.34	3.7	1.45	− 15.90	− 12.93	4.0	1.34	1.35
0.1	0.3	− 14.21	− 9.61	7.1	2.09	− 14.03	− 9.81	7.5	2.09	2.20
0.3	0.6	− 9.74	− 7.21	9.2	1.28	− 9.48	− 7.03	9.2	1.26	1.33
0.6	0.9	− 7.09	− 8.73	11.2	− 0.89	− 7.12	− 8.76	11.9	− 0.88	− 1.13
0.7	0.9	− 7.05	− 8.72	9.4	− 1.12	− 7.06	− 8.77	9.4	− 1.00	− 1.11
0.8	0.9	− 7.80	− 8.67	4.5	− 0.51	− 7.57	− 8.52	5.7	− 0.53	− 0.79
0.8	1.0	− 7.26	− 9.71	12.0	− 1.59	− 7.41	− 12.62	16.2	− 2.28	− 2.37
0.9	0.95	− 8.25	− 9.72	4.2	− 0.59	− 8.28	− 9.32	4.5	− 0.50	− 0.66
0.95	1.0	− 9.19	− 12.56	12.2	− 1.45	− 9.34	− 12.16	12.2	− 1.30	− 0.92
0.9	1.0	− 8.39	− 12.08	9.6	− 1.67	− 8.49	− 11.74	9.6	− 1.70	− 1.58
ΔA:					2.25				2.11	

a good approximation for $\Delta\lambda = 0.02$. Therefore, on the linear path, both TI and FDTI gives the same result (within 1%).

Experimentally, the free-energy difference between the dilute aqueous solutions of dimethylamine and acetone is computed from the data in Ref. 31 as -0.21 kcal/mol (favoring dimethylamine), showing a discrepancy of at least 2 kcal/mol between theory and experiment. In light of the recent sucess of Jorgensen and Ravimohan in closely reproducing the free-energy difference between the aqueous solutions of ethane and ethanol with potentials from the same library, this discrepancy is surprisingly large. It should be kept in mind, however, that the parameter set was extended without checking the exact consequences of this extension (even if it was done in a "reasonable" way). Also, dimethylamine is likely to undergo larger conformation changes due to solvent effects than acetone and as a consequence the correspondence between the simulated system and the experimental system is worse for dimethylamine than for acetone.

V. SUMMARY AND CONCLUSIONS

Free-energy Monte Carlo simulations have been performed to compute the difference between the free energy of aqueous hydration of acetone and dimethylamine using several techniques: Linear TI, integration, k th degree TI, the PM, and FDTI, the method suggested in the present paper.

The comparison of the different techniques lead to the following conclusions: (a) the interpolation inherent in any TI technique appears to work well; (b) the convergence of the linear TI is surprisingly slow; (c) the proposed FDTI was found to converge significantly faster than the other techniques tried.

For systems involving a change larger than the one treated here, the number of PM calculations required increases rapidly due to the limitation of $\Delta E_{\max}$, making the TI techniques in general preferable. In view of the comparison of the performance of the various TI techniques and the relative simplicity of its implementation, the recommended technique for problems of this kind is FDTI.

ACKNOWLEDGMENTS

This work was supported by an RCMI Grant No. SRC5G12RR03037 from NIH to Hunter College, CUNY and NSF Grant No. CHE-8203501.

[1] J. A. Barker and D. Henderson, Rev. Mod. Phys. **48**, 587 (1976).

[2] N. Quirke, *Proceedings of the NATO Summer School on Superionic Conductors* (Plenum, New York, 1980).

[3] K. S. Shing and K. E. Gubbins, *Molecular Based Study of Fluids*, edited by J. M. Haile and G. A. Mansoori (American Chemical Society, Washington, D.C., 1983), Vol. 204.

[4] Levesque, J. J. Weiss, and J. P. Hansen, in *Monte Carlo Methods in Statistical Physics II*, edited by K. Binder (Springer, Berlin, 1984).

[5] A. Pohorille and L. R. Pratt, Methods Enzymol. **127**, 64 (1986).

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

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

[8] M. Mezei and D. L. Beveridge (in preparation).

[9] A. Cross, Chem. Phys. Lett. **128**, 1981 (1986).

[10] C. H. Bennet, J. Comput. Phys. **22**, 245 (1976).

[11] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).

[12] T. P. Lybrand, I. Gosh, and J. A. McCammon, J. Am. Chem. Soc. **107**, (1986).

[13] W. L. Jorgensen and C. Ravimohan, J. Chem. Phys. **83**, 3050 (1985).

[14] M. Mezei, Mol. Phys. **47**, 1307 (1982).

[15] M. R. Mruzik, F. F. Abraham, D. E. Schreiber, and G. M. Pound, J. Chem. Phys. **64**, 481 (1976).

[16] M. Mezei, P. K. Mehrotra, and D. L. Beveridge, J. Am. Chem. Soc. **107**, 2239 (1985).

[17] H. L. Scott and C. Y. Lee, J. Chem. Phys. **73**, 4591 (1980).

[18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[19] M. Rao, C. S. Pangali, and B. Berne, Mol. Phys. **37**, 1773 (1979).

[20] J. C. Owicki and H. A. Scheraga, Chem. Phys. Lett. **47**, 600 (1979); J. C. Owicki, in *Computer Modeling of Matter*, edited by P. G. Lykos (American Chemical Society, Washington, D.C., 1978).

[21] P. K. Mehrotra, M. Mezei, and D. L. Beveridge, J. Chem. Phys. **78**, 3156 (1983).

[22] J. T. Edwards, P. G. Fand, and F. Shahidi, J. Chem. Soc. Faraday Trans. **73**, 705 (1977).

[23] M. V. Kaulgud and K. J. Patil, J. Phys. Chem. **78**, 714 (1974).

[24] M. Mezei, S. Swaminathan, and D. L. Beveridge, J. Chem. Phys. **71**, 3366 (1979).

[25] W. L. Jorgensen and C. J. Swenson, J. Am. Chem. Soc. **107**, 569 (1985).

[26] W. L. Jorgensen and C. J. Swenson, J. Am. Chem. Soc. **107**, 1489 (1985).

[27] W. L. Jorgensen, J. Chandrashekar, J. D. Madura, R. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[28] W. L. Jorgensen and J. Madura, Mol. Phys. **56**, 1381 (1985).

[29] G. Herzberg, *Electronic Spectra of Polyatomic Molecules* (Van Nostrand, Princeton, 1967).

[30] B. Beagley and T. G. Hewitt, Trans. Faraday Soc. **64**, 2565 (1968).

[31] J. Hine and P. K. Mookerjee, J. Org. Chem. **40**, 292 (1975).