Convergence acceleration in Monte Carlo computer simulation on water and aqueous solutions

Prem K. Mehrotra, Mihaly Mezei, and David L. Beveridge

Chemistry Department, Hunter College of the City University of New York, New York, New York 10021

(Received 4 January 1982; accepted 14 June 1982)

The convergence characteristics of standard Metropolis Monte Carlo calculations on liquid water and aqueous solutions are described, and documentation of the need for convergence acceleration procedures is presented. The acceleration procedures are additional importance sampling schemes added to the Metropolis method. The recently proposed "forced bias" and "preferential sampling" procedures are specifically considered. Comparative studies on liquid water at $T = 25\,^{\circ}\mathrm{C}$ using the force bias method show that convergence is improved by a factor of 2–3 over standard Metropolis results. For aqueous solutions, force bias and preferential sampling procedures used separately and together were studied on a prototype aqueous solution problem based on the liquid water simulation. Neither method alone was found to be adequate for describing structural characteristics of solutions in realizations of the length presently used for simulations on pure liquids. A combination of the force bias and preferential sampling methods was found to be quite successful, and makes aqueous solutions accessible to simulation studies at levels of rigor commensurate with that obtained for pure liquids. Preliminary convergence acceleration results on $[CH_4]_{aq}$ using the combined force bias-preferential sampling acceleration methods are also presented.

I. INTRODUCTION

Computer simulation is presently emerging as a powerful and increasingly popular computational means for the study of liquids and solutions. Applications range from liquid state chemical physics[1–4] to water and aqueous solution structure,[5–8] and most recently, to organic liquids and solutions,[9–10] and to water in biological systems.[11–14] Simulation procedures are currently being extended to further study the solvation of macromolecules, the organization of water in crystals of hydrated macromolecules, and the structure of liquids and solutions at interfaces. Considerable progress in further understanding of the structure and properties of molecular liquids and solutions via computer simulation is anticipated in the next few years.

One present problem with liquid state computer simulation is that proper calculations on systems of chemical interest require very long realizations and are thus very expensive. For example,[15] a Metropolis Monte Carlo calculation on liquid water represented as 216 water molecules under periodic boundary conditions requires at least 1500 K configurations and roughly 15 h on a computer such as the IBM 370/168 to compute stable values for internal energy and radial distribution functions. Fluctuation properties such as heat capacity may require twice as much sampling effort. Properties such as dipole correlation functions required for the calculations of the dielectric properties are even more slowly convergent, due in part to another statistical factor lost when less than full orientational averaging is involved. Molecular dynamics calculations, for analogous reasons in the deterministic realm, require roughly equivalent amounts of computer time. The situation is analogous but much worse in computer simulations on aqueous solutions, due to the loss of a statistical factor in computing properties referenced to the solute.

The basis of Monte Carlo computer simulation for fluids is a method first proposed by Metropolis *et al.*[17] (referred to as Metropolis or standard Metropolis method), now used in virtually all liquid state work. Following the seminal paper of Ceperly, Chester, and Kalos[18] on quantum liquids, several methods have recently been suggested to accelerate the convergence of Monte Carlo calculations, the main idea being to append additional importance sampling criteria to the standard Metropolis method. The principal procedures currently under consideration are the force bias method by Pangali, Rao, and Berne[19–21] and the preferential sampling approach first used by Owicki and Scheraga[22–23] for simulation studies of hydrophobic hydration, which has subsequently been generalized by Owicki.[24] An alternative force bias procedure based on Brownian dynamics has been proposed by Rossky, Doll, and Friedman.[25]

A recent paper from this laboratory[15] described in detail the convergence characteristics of the standard Metropolis method applied to liquid water as described by the quantum mechanical MCY-CI potential and by the empirical ST2 potential. We report herein comparable calculations on liquid water using force bias and preferential sampling methods. The force bias method is applicable to simulations on both pure liquids and aqueous solutions, and the preferential sampling method is most appropriate for aqueous solutions. The results from our previous study by the standard Metropolis method will serve as a point of comparison to quantify the extent of improvement achieved by additional importance sampling criteria and the computational overhead incurred. Preliminary studies in this investigation have been reported in proceedings of a recent workshop conference on computer simulation.[26] Parallel studies on liquid water and aqueous solution of methane in conjunction with Metropolis and preferential sampling schemes are being reported by Kincaid and Scheraga.[27]

II. BACKGROUND

The liquid state Monte Carlo computer simulation methods presented herein involve the determination of

0021-9606/83/063156-11$2.10

average properties of an N-molecule system under canonical ensemble conditions, with temperature T, volume V, and number density N/V fixed. Monte Carlo simulations in other ensembles (T, P, N)[9,23] and (T, V, μ)[2,28] are also being performed. We shall limit our discussion here to the (T, V, N) case, but the results have clear implications for the work in other ensembles. For general references, see the recent books by Hansen and McDonald,[29] and by Ben-Naim.[30]

An N-molecule configuration of the system is specified by the configurational coordinate vector $\mathbf{X}^N$,

$$\mathbf{X}^N = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \ldots, \mathbf{X}_N\} , \tag{1}$$

where $\mathbf{X}_i = (\mathbf{R}_i, \mathbf{\Omega}_i)$ represents the configurational coordinates of the molecule i. Here $\mathbf{R}_i$ denotes the positional coordinates and $\mathbf{\Omega}_i$ the angular coordinates, respectively. An average property of the system such as internal energy U is given by the configurational integral

$$U = \int \cdots \int E(\mathbf{X}^N) P(\mathbf{X}^N)\, d\mathbf{X}^N = \langle E(\mathbf{X}^N) \rangle , \tag{2}$$

where $E(\mathbf{X}^N)$ is the configuration energy and $P(\mathbf{X}^N)$ is the probability of finding the system in configuration $\mathbf{X}^N$,

$$P(\mathbf{X}^N) = \exp[-\beta E(\mathbf{X}^N)] \Big/ \int \cdots \int \exp[-\beta E(\mathbf{X}^N)]\, d\mathbf{X}^N . \tag{3}$$

The constant volume heat capacity in this notation is given by the expression

$$C_V = (1/kT^2)\,[\langle E^2(\mathbf{X}^N)\rangle - \langle E(\mathbf{X}^N)\rangle^2] \tag{4}$$

and analogous expressions may be formed for other properties of the system.[30]

The configurational integral of Eq. (2) is well known to be ill conditioned for direct numerical integration. In the Metropolis method, the integration is carried out by means of a stochastic walk through configuration space, by generating a realization of an irreducible Markov chain whose unique limiting stationary distribution is the Boltzman distribution,

$$\mathbf{P}(\mathbf{X}^N) = \{P(\mathbf{X}_0^N), P(\mathbf{X}_1^N), P(\mathbf{X}_2^N), \ldots\} , \tag{5}$$

where $P(\mathbf{X}_i^N)$ denotes the Boltzmann probability [Eq. (3)] of the ith configuration of the system. Thus, in a realization of this process, the configurations $\mathbf{X}^N$ are sampled with a frequency proportional to $P(\mathbf{X}^N)$, and the determination of average properties reduces to a simple summation over the energy of the individual configurations $\mathbf{X}^N$,

$$\bar{E} = (1/M) \sum_{t=1}^{M} E[X(t)] . \tag{6}$$

Here the $X(t)$ denotes the state of the N-particle system at the step t of the Markov chain,

$$X(t) \in \{\mathbf{X}_i^N ,\ i = 0, 1, 2, \ldots\} . \tag{7}$$

As $M \to \infty$, $\bar{E} \to \langle E(\mathbf{X}^N)\rangle$, the cumulative average energy becomes an increasingly good estimator of the energy expectation value. The computation of heat capacity and other properties of the system take analogous form.

The convergence and statistical error bounds of the calculations are generally monitored according to the method of block averages (also known as the method of batch means).[3] Here the Monte Carlo realization is partitioned into several nonoverlapping blocks of equal lengths, and the averages of the property under consideration (e.g., mean energy) are computed over each block. Let $\bar{f}_i$ denote the property f computed over the block i. Under the assumptions that the $\bar{f}_i$'s are independent and normally distributed and that the Markov chain is ergodic, the error bounds for the property f as a 95% confidence level are $\pm 2\hat{\sigma}$, where

$$\hat{\sigma}^2 = [1/K(K-1)] \sum_{i=1}^{K} [\bar{f}_i^2 - (\bar{f}_i)^2] \tag{8}$$

and the summation runs over the K blocks. In computer simulations of small lengths, the above assumptions are honored more in the breach than in observance, and thus computed error bounds by the method of batch means are to be taken with caution.

The details of the Metropolis method and the convergence acceleration procedures under consideration herein may all be specified in a convenient general notation based on the work of Hastings.[31] Here the elements of the one-step transition probability matrix of the Markov chain, p_{kl},

$$p_{kl} = Pr\{X(t+1) = l \mid X(t) = k\} \tag{9}$$

are written as a product of the two terms,

$$p_{kl} = q_{kl}\, \alpha_{kl} . \tag{10}$$

The first term q_{kl} is dependent on the method of generating the state l from the state k in a single step transition. The way in which state l is accepted such that the microscopic reversibility conditions are satisfied when sampling from the Boltzman distribution $\mathbf{P}(\mathbf{X}^N)$ [Eq. (5)] leads to expressions for the second term.

Hastings discusses several choices for the α_{kl}. The Metropolis choice,

$$\alpha_{kl} = \min(1, P(\mathbf{X}_l^N) q_{lk} / P(\mathbf{X}_k^N)\, q_{kl}) \tag{11}$$

has been shown by Peskun[32] to be asymptotically optimum. The elements of the one-step transition probability matrix for the Markov chain can be rewritten into a more popular notation

$$p_{kl} = q_{kl} \min(1, p_l q_{lk} / p_k q_{kl}) , \quad k \neq l , \tag{12}$$

and

$$p_{kk} = 1 - \sum_{k \neq l} p_{kl} , \quad k = l . \tag{13}$$

The various sampling methods discussed herein differ essentially in the definition of q_{kl}, i.e., the way in which state l is generated from state k in a single step transition. In principle, all the sampling schemes allow for more than 1-particle moves. However, in practice, for convergence efficiencies, the moves are restricted to a single particle. Thus, the configurational coordinates of the state l are related to the configurational coordinates of state k,

$$\mathbf{X}_l^N = \mathbf{X}_k^N + \delta^N , \tag{14}$$

where

TABLE I. Results of the Monte Carlo simulations on MCY-CI water using Barker–Watts and Euler treatments for sampling the angular configurational space. The statistical averages are computed from 10 K long runs.

Method	Step-sizes Δr (ang)	$\Delta\alpha$ (deg)	Acceptance rate	Mean angular displacement $\langle\delta\alpha\rangle$	Mean square displacement[a] $\langle\delta r^2\rangle$
Metropolis Barker–Watts	0.25	20.0	0.53	4.65	0.231
Metropolis Euler	0.25	20.0	0.52	4.65	0.210
Force bias Barker–Watts	0.50	30.0	0.64	7.92	0.615
Force bias Euler	0.50	30.0	0.65	8.02	0.474

[a] $\langle\delta r^2\rangle = (1/N)\sum_{i=1}^{N}(x_i^f - x_i^0)^2 + (y_i^f - y_i^0)^2 + (z_i^f - z_i^0)^2$, where N is the number of molecules, f denotes the last configuration, 0 the initial configuration, and x, y, z are the coordinates of the c.m. of the molecule.

$$\delta^N = \{0, 0, \ldots, \delta(\mathbf{X}_m), 0, \ldots\}, \qquad (15)$$

and $\delta(\mathbf{X}_m)$ is a displacement vector for the molecule m selected for the move. For rigid polyatomic molecules,

$$\delta(\mathbf{X}_m) = \{\delta x_{\text{c.m.}}, \delta y_{\text{c.m.}}, \delta z_{\text{c.m.}}, \delta\alpha, \xi\}, \qquad (16)$$

where $\delta x_{\text{c.m.}}$, $\delta y_{\text{c.m.}}$, $\delta z_{\text{c.m.}}$ are the displacements for the center-of-mass and $\delta\alpha$ is the rotation around a chosen axis ξ, passing through the center-of-mass of the molecule m. The magnitudes for the center-of-mass displacement and for the rotation angle are further restricted by certain step-size parameters Δr and $\Delta\alpha$, which are optimized in the initial stages of the simulation.

In Metropolis sampling, the components of the displacement vector δ are obtained by uniformly sampling from the domain D, centered at the coordinates of the molecule m in the state k, and defined by the step-size parameters Δr and $\Delta\alpha$. The elements q_{kl} of the transition probability matrix $\mathbf{Q}$ are then

$$q_{kl} = \text{a constant}, \quad \mathbf{X}_l^N \in \mathbf{X}_k^N + D, \qquad (17)$$

and

$$q_{kl} = 0, \quad \mathbf{X}_l^N \notin \mathbf{X}_k^N + D. \qquad (18)$$

Moreover, Q is a symmetric matrix.

The methods for the selection of the rotation axis ξ need further elaboration. The two most common methods employed in the literature are a method due to Barker and Watts[33–34] and another generally referred as the Euler angle treatment.[20,35] In the procedure of Barker and Watts, the axis ξ is selected by uniformly sampling from the x, y, and z axes defined in a fixed frame of reference. In the Euler angle treatment, the axis ξ and the rotation angle $\delta\alpha$ are obtained by uniformly sampling from the three Euler angles. The choice of the Euler angles often varies. In the method discussed herein, the polar angles θ and ϕ of the ξ with respect to a fixed frame of reference are two of the Euler angles, and the rotation angle $\delta\alpha$ around the ξ axis forms the third. The Euler angle θ is uniformly sampled from the $\cos(\theta)$ (-1 to 1) distribution, ϕ from 0 to 2π, and $\delta\alpha$ from $-\Delta\alpha/2$ to $+\Delta\alpha/2$.

In all the Metropolis Monte Carlo work published from this laboratory,[6,7,15] the rotation axis ξ has been sampled according to the Barker and Watts procedure. We have recently also implemented the Euler angle treatment as described above. Our exploratory simulations on liquid water, as shown in Table I, show no convergence differences between the two sampling schemes. Thus, all our Metropolis Monte Carlo results to be discussed in this paper refer to the Barker and Watts sampling for the angular space.

In a typical Monte Carlo computer simulation on a molecular liquid in the (T, V, N) ensemble, the system consists of a simulation cell containing N molecules in a volume V determined by N/ρ, where ρ is the experimental density at the system temperature T. The system is presented with a condensed phase environment by means of periodic boundary conditions, with the central cell surrounded at each face, edge and vertex by a self-image. The configurational energy of the system is computed by means of analytical potential functions. Calculations from this laboratory use mainly simple cubic or face centered cubic boundary conditions and spherical and/or minimum image cutoff criterions for the potential functions. The initial segment of the calculation is an equilibration phase, and is discarded in the formation of ensemble averages.

The reference calculation for the convergence acceleration procedures described herein is a long realization of the 4400 K configurations on liquid water carried out in this laboratory[15] using the standard Metropolis method and a potential function representative of quantum mechanical calculations of the intermolecular interaction energy developed by Matsuoka, Clementi, and Yoshimine. The calculated radial distribution function $g(R)$ for the molecular centers-of-mass obtained from this simulation is presented along with the corresponding experimental $g(R)$[16] in Fig. 1. The calculated position of the successive hydration shells, 3.3, 5.7, and 8.0 Å, serves as an index of comparison for the convergence studies described in this paper.

The convergence profile from this study is shown in Fig. 2. Note particularly the expanded scale chosen for the energy ordinate here and in subsequent analogous figures; the convergence characteristics here are discussed in tenths of kcal/mol. The calculation achieves a mean energy of -8.57 kcal/mol after 200 K configurations. The mean energy oscillates within 0.1 kcal/mol of this value for the next 1000 K configurations. The convergence profile indicates a mean energy value of -8.56 ± 0.03 kcal/mol for this section of the run. At $N = 1200$–1400 K, the control function shows a sharp decline of 0.13 kcal/mol in energy to -8.78 kcal/mol at $N = 1400$ K, and the onset of a region of 1600 K configurations with a mean energy -8.75 ± 0.02 kcal/mol. Concomitant with this decline is a sharp increase in heat capacity. At $N = 3000$ K, the control function rises again, and at termination is oscillating about -8.64 ± 0.03 kcal/mol. The heat capacity is relatively constant from $N = 2000$ K on. The general appearance of the control function suggests that the high frequency oscillations in the control function are sumperimposed on a grand oscillatory cycle with an amplitude of 0.2 kcal in the realization, of which our calculation covers 1 and 1/2 cycles. Similar behavior has also been discussed by Pangali *et al*.[20] The cumulative mean energy is -8.65 kcal/mol, with a heat capacity of 14.1 cal/mol deg, which after the kinetic energy correction results in -6.87 kcal/mol, and 20.1 cal/mol deg for mean energy and heat capacity, respectively.

Based on these results, a hierarchy of calculated properties of the system, with respect to computational effort, was provisionally established as follows: mean energy and orientationally averaged radial distribution functions are the quantities most accessible to calculation by computer simulation. Fluctuation properties such as heat capacity require at least twice as many configurations to be calculated with the same relative precision. The calculations of dipole correlation functions and other quantities not based on fully orientational averaging can be expected to remain problematic at this time.

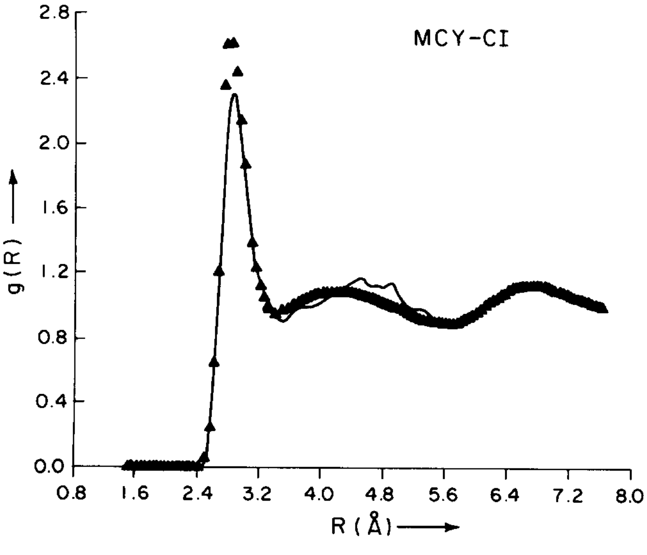

FIG. 1. Comparison between the calculated (Ref. 15) (▲) and the experimental (Ref. 16) (—) $g(R)$'s, where R is the distance between the center-of-masses, for liquid water.

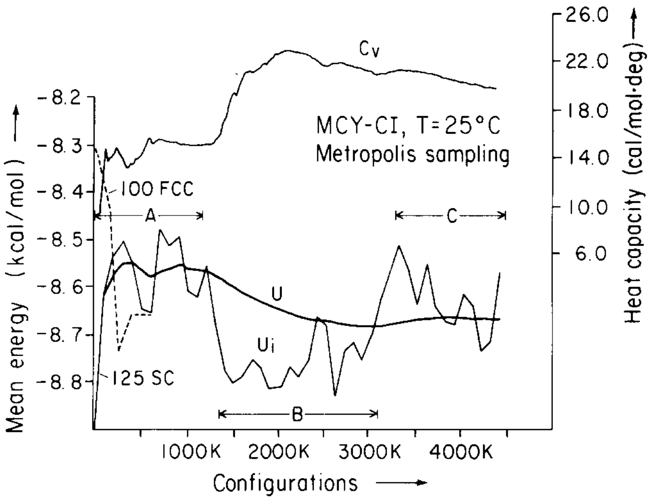

FIG. 2. Convergence curves (Ref. 15) for the Metropolis Monte Carlo simulations on liquid water based on the MCY-CI potential.

III. CALCULATIONS, RESULTS AND DISCUSSIONS

The convergence acceleration procedures under consideration here are the force bias method, applicable to both pure liquids and solutions, and the method of preferential sampling, designed for use in simulations on solutions. The force bias results in this section are presented in the context of convergence studies on liquid water, $[H_2O]_l$, at 25 °C. The liquid water simulation is subsequently analyzed as an aqueous solution problem of one "solute" water in the remaining "solvent" waters, denoted as $[H_2O]_{aq}$, and used to demonstrate in prototype some significant convergence problems of aqueous solution simulations. Results on convergence acceleration in Monte Carlo simulations on $[H_2O]_{aq}$ by both force bias and preferential sampling methods are then presented. Finally, the force bias and preferential sampling techniques are applied to the aqueous solution problem $[CH_4]_{aq}$ at 25 °C, studied previously in this laboratory by the standard Metropolis method.

A. Studies on the pure liquid case

1. Force bias procedures

In the force bias sampling,[20] the particle moves are biased in the direction of forces and torques on the molecule selected for the move. The elements q_{kl} of the transition probability matrix for force bias sampling are given by the expressions

$$q_{kl} = Q^{-1}(\mathbf{X}_k^N) \exp\{\lambda(\mathbf{F}_m(\mathbf{X}_k^N) \cdot \delta\mathbf{r} + \mathbf{N}_m(\mathbf{X}_k^N) \cdot \delta\boldsymbol{\alpha})/kT\} ,$$

$$\mathbf{X}_l^N \in \mathbf{X}_k^N + D \tag{19}$$

and

$$q_{kl}=0\ , \quad \mathbf{X}_l^N \notin \mathbf{X}_k^N + D\ . \tag{20}$$

Here $\delta\mathbf{r}=(\delta x_{\mathrm{c.m.}}, \delta y_{\mathrm{c.m.}}, \delta z_{\mathrm{c.m.}})$, $\delta\boldsymbol{\alpha}=\hat{\xi}\delta\alpha$, where $\hat{\xi}$ is a unit vector along the rotation axis, $Q(\mathbf{X}_k^N)$ is a normalization constant, and λ is a parameter to be optimized during the sampling. The quantities $\mathbf{F}_m(\mathbf{X}_k^N)$ and $\mathbf{N}_m(\mathbf{X}_k^N)$ are the forces and torques in the state k on the particle to be moved. Note that $q_{lk} \neq q_{kl}$. In our force bias calculations, λ is set equal to 0.5 following Rossky *et al.*[25]

As in the Metropolis case discussed in the preceding section, we have implemented both the Barker and Watts method and the Euler angle treatment for sampling angular coordinates. The Barker and Watts sampling was used in the first paper on the force bias sampling by Pangali, Rao, and Berne.[19] These authors subsequently reported a force bias sampling version incorporating the Euler angle method.[20] Intuitively, it appears that Euler angle treatment in conjunction with force bias sampling should lead to an improved sampling, because the selection of the rotation axis is also biased in the direction of the torque. Our exploratory calculations on liquid water summarized in the Table I, suggest that the Euler angle treatment does not lead to significant improvements here. Since the additional computational overhead in implementation of the Euler angle method is significant, we adopted the Barker and Watts procedure for the calculations presented herein.

For an absolute comparison between Metropolis sampling and force bias sampling, the magnitudes of the step-size parameters Δr and $\Delta\alpha$ and other setup characteristics must be optimized as fully as possible in an initial short segment of the realization. One useful criterion for this optimization, based on translational and angular diffusion of particles, was reported by Pangali *et al.*[20] A different criterion related to particle diffusion was also considered. Using the analogy of random walks, Kalos[36] has proposed the quantity $\Delta r\langle A\rangle^2$ as an index of sampling efficiency, where A denotes the acceptance rate. To include the cage effect to first order, we have modified this index to $\Delta r\langle A\rangle^2 \times(1+\langle\cos(\theta)\rangle)$,[37] where θ is the angle between the successive accepted moves of a molecule. The best positional and angular displacement for the force bias method, according to the new criterion, were found to be neary double of the respective displacement in the standard Metropolis method.

2. *Force bias results*

Convergence acceleration studies using force bias were carried out on MCY-CI water at 25 °C and analyzed in terms of particle diffusion, "rate" of equilibration and the evolution of internal energy and heat capacity during the Monte Carlo realization. The particle diffusion rate was found to be four times greater in the force bias computation, indicative of the increased sampling of the configuration space for individual particles. The equilibration of the computation in the initial segment of the realization was found to be 2–3 times faster with force bias. The convergence profile, Fig. 3, shows that both internal energy and heat capacity are well settled down after 1000 K configurations, an improvement of a factor of 2 to 3 over the standard Metropolis results shown in Fig. 2. Preliminary studies of the energy autocorrelation function for the force bias and standard Metropolis method show that in the latter, significant correlations persist even after some 2000 K configurations of sampling. In the force bias simulations the energy correlations are greatly reduced after mere 500 K configurations. These findings on MCY-CI water are in general accord with the force bias results of Pangali *et al.*[20] on ST2 water.

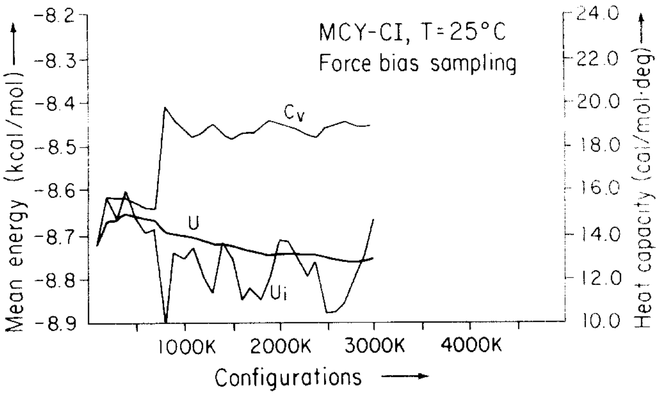

FIG. 3. Convergence curves for the force bias Monte Carlo simulation on liquid water based on the MCY-CI potential. Notation follows that of Fig. 2.

While the convergence acceleration achieved by force bias on MCY-CI water is quite significant, the overhead incurred in a force bias simulation is also large. The extra computation time for the forces involves an additional factor of ~1.2. For medium size (N= 100–300) systems, there is a further expense. In the standard Metropolis method, the recomputation of the N pair energies at every accepted step can economically be avoided by storing all pair energies in an $N(N-1)/2$ matrix in the main memory. For the force bias algorithms, pair forces (a three component vector), pair torques (a six component vector) and virial sum contribution (for the pressure calculation) must also be stored requiring 11 times as much core storage. This exceeds reasonable limits in most computer facilities, so in force bias the pair energies, forces, torques, and virial sum must be recomputed at every accepted step, increasing the computation time by a further factor of (1+ acceptance rate). (For calculations where the pressure is required, the force calculations are necessary anyway, but if the standard Metropolis method is used, the forces for pressure calculations can be computed periodically, e.g., at every 2*N step, adding only a 1.3 factor to the Metropolis method.) Thus, in a 125 molecule liquid water simulation, the economies realized in the acceleration of convergence are to some extent offset by the computational overhead. In larger systems, where the pair energy storage algorithm is no longer applicable, the factor of 2–3 improvement applies overall. The essential result is that convergence is significantly accelerated by the force bias procedure while the cost of the calculation in terms of computer time at worst stays about the same or at best goes down by a factor of 2–3. For $[H_2O]_l$, these results indicate that force bias sampling is a definite improvement over the standard Metropolis method, and is a preferred procedure for doing the simulation.

The extent to which this conclusion may be generalized remains to be seen. Each new system and really each new potential function is a new problem, and broad experience will be necessary to evaluate the overall improvement to be expected for molecular liquids and solutions from the force bias sampling algorithm. However, water is certainly a representative case of associated liquids, and we feel the convergence acceleration by force bias for this system is quite impressive.

B. Studies on the aqueous solutions case

We turn now to a consideration of the use of the force bias and the preferential sampling as a means of accelerating the convergence of Metropolis Monte Carlo calculations on aqueous solutions. In the preceding calculations on pure water, configurational averages are computed by averaging over all configurations and, since all particles are identical, averaging also over all of the N particles in the system. In a commensurate simulation on aqueous solutions, there is one solute and there are N-1 solvent particles. In calculating configurational averages with respect to the solute such as solute–water radial distribution functions and related quantities, the statistics are *a priori* reduced by a factor of $1/N$ with respect to the pure liquid case. Thus, to attain an equivalent convergence stability, the Monte Carlo realization would be N times as long, requiring an unreasonably long computation time.

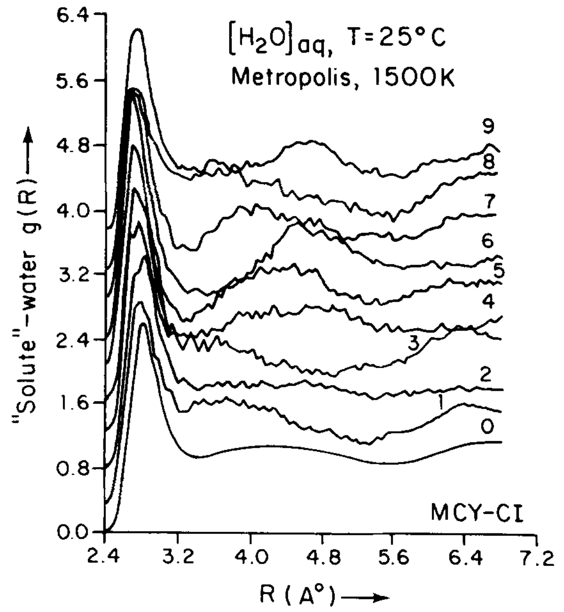

FIG. 4. The solute–water $g(R)$'s for the nine water molecules randomly chosen from the 125 water molecule Metropolis Monte Carlo simulation on MCY-CI liquid water. The total number of configurations involved in computations of $g(R)$'s is 1500 K, and constitute the 900–2600 K interval of the 4400 K run. The curve o refers to the pure liquid water $g(R)$ obtained by averaging over all the 125 water molecules. Note that the origins of the various $g(R)$'s are shifted for clarity.

TABLE II. Analysis of the Metropolis Monte Carlo simulation on MCY-CI water. The number of configurations involved in statistical averaging is 1500 K.

Property	Minimum and maximum values observed among 125 solute waters		Pure liquid water (averaged over 125 waters)
	minimum	maximum	
$g(R_{max})$	2.15	3.12	2.67
$\bar{K}$	3.81	4.73	4.30
x_c (3)	0.06	0.32	0.15
x_c (4)	0.30	0.67	0.47
x_c (5)	0.14	0.43	0.29

Current aqueous solution work from this laboratory[38–41] and elsewhere[11–13,23,44] has typically relied on roughly the same number of configurations as for simulations on the pure liquid. As a result, the reported error-bars on solute–solvent properties were found to be an order of magnitude larger than those observed for the solvent–solvent properties in pure liquid simulations.[38–41] While these simulations yield the gross structural and energetic features of the aqueous solutions, the large errors become of particular concern when studying small energetic and structural changes of current interest, such as structure-making and structure-breaking, temperature dependence of aqueous solution composition, and the effect of solvent on the small conformational energy changes in the solute. For such studies, a higher degree of computational precision is desired.

The severity of the problem may be quantified by once again focusing on the system of pure water, and treating it as an aqueous solution where one of the waters is arbitrarily chosen as solute and the rest are treated as solvent in a manner identical to that used in a study of an infinitely dilute solution simulation. Here we have the known results for the pure liquid available for quantitative comparison, so effectively we know the answer. Within the computer simulation results on the pure liquid, there are N solute–water problems, and the solute–water $g(R)$'s and related properties can be examined individually. In calculations fully converged at a level appropriate for aqueous solution studies, all solute–water $g(R)$'s and other equilibrium properties should be identical to each other and to the $g(R)$ and corresponding properties calculated for liquid water (averaged over all the N particles of the system).

To examine this, we first considered a 2400 K contigous section of the 4400 K realization. The first 900 K configurations were discarded and configurational averages were formed from the remaining 1500 K of the section. The results are presented in Fig. 4 and Table II. In Fig. 4, the solute–water $g(R)$'s [referred as solute $g(R)$'s] for nine water molecules randomly selected from the set of 125 water molecules are presented. Note that R is the distance between the center-of-mass (c.m.), of the solute water and the c.m. of the solvent water, and N-1 distances (between 1 solute and N-1 solvent molecules) are involved in the computation for

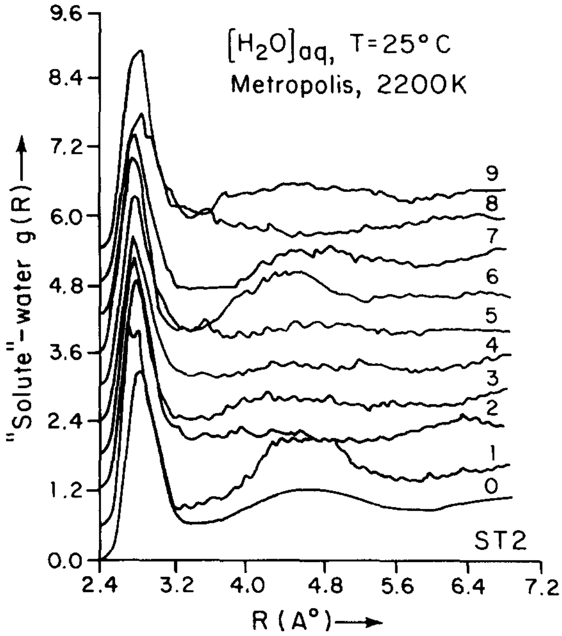

FIG. 5. The solute–water $g(R)$'s for the nine water molecules randomly selected from the 216 water molecule Metropolis Monte Carlo simulation on ST2 water. The total number of configurations involved in computations of $g(R)$'s is the last 2200 K configurations of the 5000 K realization. The curve o refers to the pure liquid water $g(R)$ obtained by averaging over all the 216 water molecules. Note that the origins of the various $g(R)$'s are shifted for clarity.

each N-particle configuration. The differences among the nine solute $g(R)$'s in Fig. 4 are clearly significant in both the first hydration shell and beyond. We have in fact examined the solute $g(R)$ for all the 125 water molecules, and similar differences are observed among most of them. The clear implication is that realizations on solutions are not well converged in runs of lengths sufficient to stabilize corresponding results on a pure liquid.

In order to further appreciate the magnitude of differences present in the first shell of the $g(R)$'s, we have compiled, in Table II, the minimum and maximum values (observed among the 125 water molecules) of a few properties dependent on the water molecules inside the first shell. For example, the height of the first peak, as reflected in the first maximum of $g(R)$, ranges from 2.15–3.25. The coordination number (computed for a cutoff of 3.3 Å) varies from 3.81 to 4.73. Similar differences are also observed among individual quasicomponents of the coordination number. Further examination indicates that all the aforementioned differences persist even when all the 3500 K post-equilibration configurations are included in the statistical averaging.

TABLE III. Analysis of the Metropolis Monte Carlo simulation on ST2 water. The number of configurations involved in statistical averaging is 2200 K.

Property	Minimum and maximum values observed among 216 solute waters		Pure liquid water (averaged over 216 waters)
	minimum	maximum	
$g(R_{max})$	2.67	4.15	3.36
$\bar{K}$	3.38	6.36	4.69
$x_c(4)$	0.03	0.98	0.42
x_c (5)	0.01	0.59	0.33
x_c (6)	0.00	0.47	0.15

TABLE IV. Analysis of the force bias Monte Carlo simulation on MCY-CI water. The number of configurations involved in statistical averaging is 1500 K.

Property	Minimum and maximum values observed among 125 solute waters		Pure liquid water (averaged over 125 waters)
	minimum	maximum	
$g(R_{max})$	2.35	2.89	2.61
$\bar{K}$	4.01	4.46	4.23
x_c (3)	0.11	0.22	0.16
x_c (4)	0.36	0.56	0.46
x_c (5)	0.19	0.36	0.28

We have also carried out an analogous analysis for liquid water simulations based on ST2 water. The Metropolis Monte Carlo simulation on ST2 water employed 216 water molecules, and the last 2200 K configurations of 5000 K run published from this laboratory[15] were used in statistical averaging. The solute $g(R)$'s for the nine randomly selected water molecules are presented in Fig. 5, and the range of values observed for the height of the first peak of $g(R)$, coordination number, etc. are compiled in Table III. In comparison to the MCY-CI water, the differences among individual water properties are in fact larger for the ST2 water. In any case, this indicates that the convergence problems on aqueous solution simulation case arise largely independent of the choice of the potential functions.

1. *Force bias results*

In order to examine the effect of force bias sampling in improving the convergence of solute properties in an aqueous solution case, we have reanalyzed our force bias liquid water simulation discussed in Sec. III A, viewing it as a prototype aqueous solution problem. The single water analysis is identical to one discussed in the preceding section. The nine solute–water $g(R)$'s computed over the 1500 K configurations are presented in Fig. 6, and the results for coordination number, etc. are presented in Table IV.

The results show that differences among individual water properties are reduced by approximately half in the force bias simulation compared to standard Metropolis simulations of the same length. According to $N^{1/2}$ law of statistics, this amounts to a convergence improvement of approximately four in the force bias simulation, which is in accord with a factor of 2–3 discussed in the pure liquid section. Despite this improvement, sizable

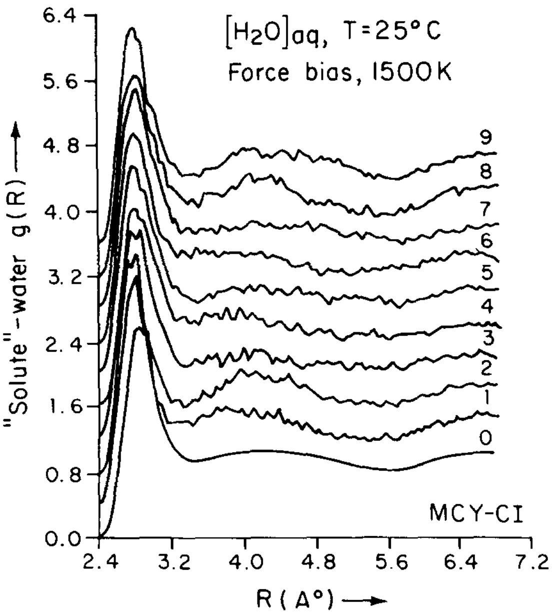

FIG. 6. The solute–water $g(R)$'s for the nine water molecules randomly selected from the 125 water molecule force bias Monte Carlo simulation on MCY-CI water. The total number of configurations involved in statistical averaging is 1500 K. The curve o refers to the pure liquid water $g(R)$ obtained by averaging over all the 125 water molecules. Note that the origins of the various $g(R)$'s are shifted for clarity.

significant differences among individual water properties still exist, indicating that more configurations are required in the averaging before the differences among individual water properties disappear.

In a recent study on the potential of mean force calculations of two apolar solutes in 214 water molecules, Pangali, Rao, and Berne[42] in order to justify the adequate convergence for the solute–solvent properties compared a solute–water $g(R)$ (computed over the two water molecules instead of one), with the $g(R)$ computed over all 216 water molecules for ST2 water. Force bias sampling was used, and the number of configurations involved in the statistical averaging were O(1100 K). They reported a rather good agreement between the solute and pure liquid water $g(R)$'s. In these more extensive studies we do find that there are a few water molecules whose solute $g(R)$'s are similar to the pure liquid water $g(R)$, but these are only a small fraction of the total. Kincaid and Scheraga[27] also arrive at conclusions similar to ours, in their analogous studies on liquid water.

2. *Preferential sampling procedures*

We subsequently set out to learn whether the convergence of solute properties can be improved by using preferential sampling. In both Metropolis and force bias samplings on dilute solutions, all the solvent molecules are sampled (selected for configuration perturbation) with equal probability. On the other hand, many of the interesting properties of the dilute solutions are primarily determined by the solute–solvent and solvent–solvent interactions near the solute. This suggests that the computational efficiency for the convergence of many solution properties could be expedited with sampling concentrated primarily in the neighborhood of the solute. Clearly the solvent molecules far away form the solute should also be sampled, but not necessarily as frequently as those near the solute. This strategy is the basis for the method of preferential sampling first proposed by Owicki and Scheraga for the Monte Carlo calculation on $[CH_4]_{aq}$. Owicki subsequently presented a generalized version of the preferential sampling methodology. The two preferential sampling schemes have recently been applied by Bigot and Jorgensen[10] to study the conformational equilibria of a n-butane in carbon tetra-chloride liquid.

Owicki's sampling scheme[24] can be described as follows. Let $w(R)$ be a weighting function which is nonnegative and decreasing function of R, where R is dependent on solute–solvent distances. For each N-particle configuration X_k^N define a probability distribution function $W(X_k^N)$,

$$W(X_k^N) = \{W_2(X_k^N),\ W_3(X_k^N), \ldots ,\ W_N(X_k^N)\} , \tag{21}$$

where the subscript refers to the molecule number and the solute is labeled as the molecule 1. Furthermore,

$$W_m(X_k^N) = w(R_m) \Big/ \sum_{i=2}^{N} w(R_i) . \tag{22}$$

In preferential sampling, the solvent molecules are selected for the moves by sampling from the probability distribution $W(X_k^N)$. If m is the solvent molecule thus selected in a single-particle move, then the elements q_{kl} are

$$q_{kl} = \text{const.} * W_m(X_k^N) . \tag{23}$$

Clearly, q_{kl} is not equal to q_{lk}, and

$$q_{lk} = \text{const.} * W_m(X_l^N) . \tag{24}$$

The selection of the solute for the move follows a different treatment. The preferential sampling scheme, in principle, allows that the solute can be picked with an arbitrary frequency, thus,

$$q_{kl} = q_{lk} = \text{const} . \tag{25}$$

In the limits, the solute is either never moved or is always selected for move. As Owicki points out, there are definite statistical advantages if the solute is also moved, because a solute move leads to perturbations in N-1 solute–solvent pairs, whereas a solvent move causes only one solute–solvent perturbation. On the other hand, if the solute moves are attempted too frequently there is little room for the solvent molecules to properly relax. Clearly, a compromised approach is desired. Instead of carrying out a complete optimization requiring several long test runs, the solute molecule, in present studies, is perturbed with the same frequency as its neighboring solvent molecules.

The preferential sampling can be combined with the Metropolis sampling as well as with the force bias sampling. Then the elements q_{kl} are the product of the preferential sampling q_{kl} with the Metropolis of the force bias q_{kl}, as the case may be. For the simple weighting function as used in our studies, the computational overhead due to the preferential sampling is negligible.

Preferential sampling simulations on $[H_2O]_{aq}$ in conjunction with both the Metropolis sampling and the force bias sampling were carried out as follows. Starting from an equilibrated configuration from our 4400 K Monte Carlo simulation, we arbitrarily label one water as the solute. A weighting function of $1/R$, where R is the distance between the center-of-mass of the solute water and the center-of-mass of the solvent water was employed. The step-sizes used for the solute water are same as for the solvent water. In comparison to a standard Metropolis or force bias simulation on $[H_2O]_{aq}$, the solute and its neighboring solvent molcules are perturbed 2.5–3.5 times more frequently in our preferential sampling simulation. This frequency can be further increased if weighting functions of the form $1/R^2$[10,24] or $1/R^4$,[10] are used instead of $1/R$.

3. *Preferential sampling results*

The results of preferential sampling used in conjunction with the Metropolis sampling are shown in Fig. 7. The 4.2 Å peak in the $g(R)$ is chosen as our index of comparison, since it is both a sensitive test and an essential element in the description of tetrahedrally coordinated network of molecules in a waterlike system. The solid line in each case is the pure liquid $g(R)$ (averaged over all particles) obtained in the previous 4400 K simulation [cf. Fig. 1(a)] and can be considered the desired result. The various dotted lines are the solute $g(R)$'s resulting from simulations on $[H_2O]_{aq}$. The results for standard Metropolis sampling after 1600 K configurations and 3200 K configurations, curves 7(a) and 7(b), respectively, do not accurately reproduce the shell structure of the liquid. With preferential sampling, curve 7(c), there is a marked improvement after 1800 K configurations, but clear discrepancies with the desired result remain.

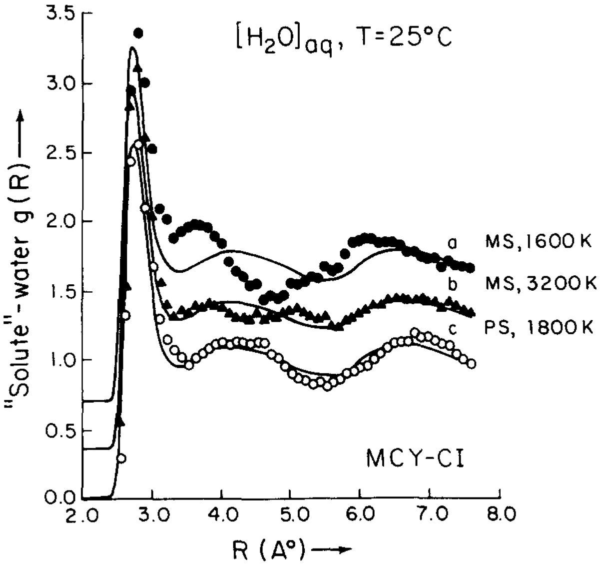

FIG. 7. The solute–water $g(R)$'s for the Metropolis Monte Carlo MCY-CI liquid water simulations. All $g(R)$'s refer to the same water molecule. The same initial configuration was used in all the simulations. (a) 1600 K Metropolis (MS) Monte Carlo simulation; (b) 3200 K Metropolis Monte Carlo simulation; and (c) 1800 K preferential sampling (PS) Monte Carlo simulation. The solid line is the reference pure liquid water $g(R)$ and the origins of various curves are shifted for clarity.

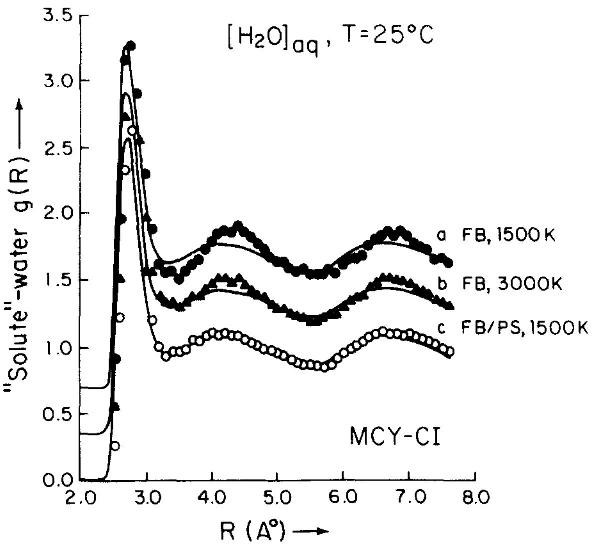

FIG. 8. The solute–water $g(R)$'s for the force bias Monte Carlo simulation. The initial configuration and the water molecule selected as the solute are same as in Fig. 7. (a) 1500 K force bias (FB) Monte Carlo simulation:, (b) 2500 K force bias Monte Carlo simulation; and (c) 1500 K preferential sampling in conjunction with force bias (FB/PS) Monte Carlo simulation. The solid line is the reference pure liquid water $g(R)$ and the origins of various curves are shifted for clarity.

4. *Results from force bias and preferential sampling*

The effect of force bias sampling used alone and in conjunction with preferential sampling on $[H_2O]_{aq}$ is shown in Fig. 8. The results of force bias sampling after 1500 and 2500 K configurations are shown as curves 8(a) and 8(b) against the desired result (solid line). Clear discrepancies between the computed and desired result remain. In the last simulation, curve 8(c), we have combined the force bias and preferential sampling procedures, and have carried out a realization of 1500 K configurations. Here the result is a curve almost exactly coincident with that desired. This indicates that the Metropolis sampling enhanced with a combination of force bias and preferential sampling procedures has been successful at converging the simulation results on an aqueous solution case using runs of similar order of length used for the liquid water problem.

The conclusion is supported by results from a simulation taking a different water as the reference solute. In Fig. 9, the solute $g(R)$ averaged over the 2000 K configurations of the combined force bias-preferential sampling simulation is compared with the solute $g(R)$ obtained from the 2500 K configurations of the force bias simulation. The solid line once again represents the pure liquid water $g(R)$. Although there are small differences between the solute $g(R)$ and pure liquid water $g(R)$ for the combined simulation, these dif-

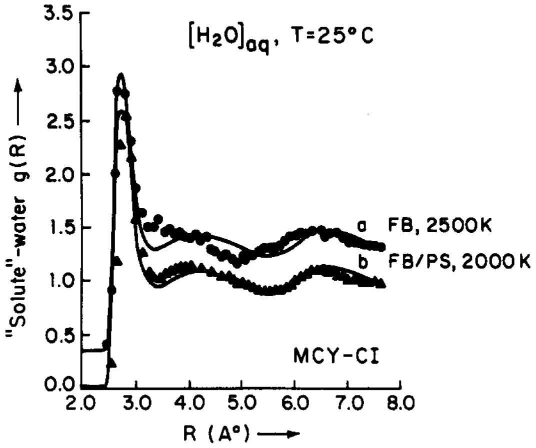

FIG. 9. The solute–water $g(R)$'s for the force bias Monte Carlo simulation. The initial configuration is same as used in derivation of Fig. 9, but the water molecule labeled as solute is different. (a) 2500 K force bias (FB) Monte Carlo simulation; and (b) 2000 K preferential sampling in conjunction with force bias (FB/PS) Monte Carlo simulation. The solid line is the reference pure liquid water $g(R)$ and the origins of various curves are shifted for clarity.

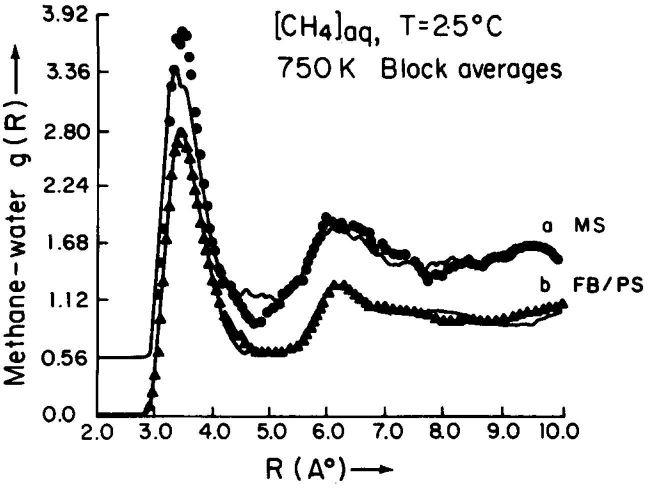

FIG. 11. The methane–water $g(R)$'s for the standard Metropolis Monte Carlo and for the force bias-preferential sampling Monte Carlo simulations on $[CH_4]_{aq}$ (a) Metropolis sampling (MS). ●, $g(R)$ computed over the first 750 K block, —$g(R)$ computed over the second 750 K block. (b) same as (a) except the use of force bias-preferential sampling (FB/PS).

ferences are negligible in comparison to large differences observed between the force bias solute $g(R)$ and the reference liquid water $g(R)$.

One remaining point to consider is whether or not preferential sampling degrades the precision of simulation results on solvent properties. The radial distribution function computed for the solvent waters in $[H_2O]_{aq}$ are shown in Fig. 10, where the solvent $g(R)$ is compared with the reference $g(R)$ result for pure liquid water. There are no noticeable differences between the two results, indicating that the statistics for solvent properties are not degraded by the preferential sampling procedure referenced to the solute. This is not surprising because from simple geometric considerations it follows that the solvent molecules more frequently moved are relatively small in number. Thus, the decrease in the sampling frequency of the solvent molecules far way from the solute is relatively small, and the solvent statistics are not significantly degraded.

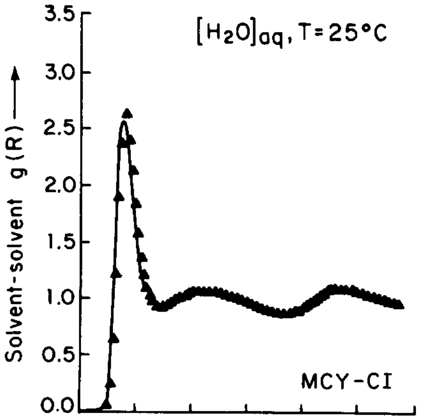

FIG. 10. The solvent–solvent $g(R)$ for the force bias-preferential sampling simulation on MCY-CI water. The number of configurations used in statistical averaging is 1500 K. The solid line is the reference pure liquid water $g(R)$.

5. Results of the force bias and preferential sampling on $[CH_4]_{aq}$

In order to further document the improvements in convergence observed for the force bias-preferential sampling simulation over the standard Metropolis sampling, Monte Carlo calculations were also carried out on $[CH_4]_{aq}$ at 25 °C. Starting from an equilibrated configuration of an earlier Metropolis sampling simulation on this system,[43] a realization of 2000 K configurations was generated employing the combined force bias-preferential sampling procedure. The potential function and other setup characteristics such as T, V, N and potential cutoffs are kept the same as in the standard Metropolis run to provide for a consistent comparison. The step-sizes, weighting function and other parameters for the combined sampling simulation on $[CH_4]_{aq}$ are identical to ones used for the $[H_2O]_{aq}$ simulation discussed earlier.

The first 500 K configurations of the 2000 K configurations realization were discarded. The remaining 1500 K configurations have been further divided into two nonoverlapping blocks of 750 K each. The solute–solvent $g(R)$ between the center-of-mass of the CH_4 and the center-of-mass of the water, computed over the two blocks are presented in Fig. 11(a). The results of a parallel analysis for the Metropolis sampling simulations are presented in Fig. 11(b) for comparison. The results for the Metropolis sampling show the differences between the two block averages are significant. In the combined simulation, the differences are negligible (especially in the first and second shells). This result reinforces the conclusions on the efficacy of a combination of the force bias and the preferential sampling inferred from the prototype problem.

IV. SUMMARY AND CONCLUSIONS

The convergence characteristics of standard Metropolis Monte Carlo calculations on liquid water and aqueous solutions are described, and documentation of the need for convergence acceleration procedures was presented. Comparative studies on liquid water at $T = 25\,^{\circ}\mathrm{C}$ using the force bias method show that convergence is improved by a factor of 2–3 over standard Metropolis results. For aqueous solutions, force bias and preferential sampling procedures used separately and together were studied on a prototype aqueous solution problem based on the liquid water simulation. Neither method alone was found to be adequate for describing structural characteristics of solutions in realizations of the order presently used for simulations on pure liquids. A combination of the force bias and preferential sampling methods was found to be quite successful, and makes aqueous solutions accessible to simulation studies at levels of rigor commensurate with that obtained for pure liquids. Preliminary convergence acceleration results on $[\mathrm{CH_4}]_{\mathrm{aq}}$ using the force bias-preferential sampling acceleration methods further support the conclusion.

ACKNOWLEDGMENTS

This research was supported by Grant GM-24149 from the National Institutes of Health. Studies on the temperature dependence of $[\mathrm{CH_4}]_{\mathrm{aq}}$ using the standard Metropolis sampling by G. Ravi-Shanker in our laboratory provided the major stimulus for the present work. The authors are grateful to G. Ravi-Shanker. T. R. Vasu, F. T. Marchese, R. H. Kincaid, W. J. Peer, H. A. Scheraga, and W. L. Jorgensen for helpful discussions.

[1] W. W. Wood and J. J. Erpenbeck, Annu. Rev. Phys. Chem. **27**, 319 (1979).

[2] J. P. Valleu and S. G. Whittington, *Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1977), Vol. 5, Chap. 5.

[3] J. E. Erpenbeck and W. W. Wood, Ref. 2, Vol. 6, Chap. 1.

[4] J. Kushick and B. J. Berne, Ref. 2, Vol. 6, Chap. 2.

[5] F. H. Stillinger, Science **209**, 451 (1980).

[6] D. L. Beveridge, M. Mezei, S. Swaminathan, and S. W. Harrison, *Computer Modeling of Matter*, edited by P. G. Lykos (American Chemical Society, Washington, D. C., 1978), Chap. 16.

[7] D. L. Beveridge, M. Mezei, P. K. Mehrotra, F. T. Marchase, V. Thirumali, and G. Ravi-Shanker, *Quantum Chemistry for Biomedical Sciences* (Annals New York Academy of Sciences, New York, 1981), Vol. 367, p. 108.

[8] G. Nemethy, W. J. Peer, and H. A. Scheraga, Annu. Rev. Biophys. Bioeng. **10**, 459 (1981).

[9] W. L. Jorgensen, J. Am. Chem. Soc. **103**, 4721 (1981).

[10] B. Bigot and W. L. Jorgensen, J. Chem. Phys. **75**, 1944 (1981).

[11] A. T. Hagler and J. Moult, Nature (London) **272**, 222 (1978).

[12] P. J. Rossky and M. Karplus, J. Am. Chem. Soc. **101**, 1913 (1979).

[13] E. Clementi, *Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry* (Springer, New York, 1980), Vol. 19.

[14] J. Hermans and M. Vacatello, *Water in Biopolymers*, edited by S. P. Rowland (American Chemical Society Symposium Series, Washington, D.C., 1979), Chap. 11.

[15] M. Mezei, S. Swaminathan, and D. L. Beveridge, J. Chem. Phys. **71**, 3166 (1979).

[16] A. H. Narten and H. A. Levy, J. Chem. Phys. **55**, 2263 (1971).

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

[18] D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977).

[19] C. Pangali, M. Rao, and B. J. Berne, Chem. Phys. Lett. **55**, 413 (1978).

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

[21] M. Rao and B. J. Berne, J. Chem. Phys. **71**, 129 (1979).

[22] J. C. Owicki and H. A. Scheraga, Chem. Phys. Lett. **47**, 600 (1979).

[23] J. C. Owicki and H. A. Scheraga, J. Am. Chem. Soc. **99**, 7413 (1977).

[24] J. C. Owicki, *Computer Modeling of Matter*, edited by P. G. Lykos (American Chemical Society, Washington, D. C. 1978).

[25] P. J. Rossky, J. D. Doll, and H. L. Friedman, J. Chem. Phys. **69**, 4628 (1978).

[26] P. K. Mehrotra, M. Mezei, and D. L. Beveridge, Computer Simulations on Organic and Biological Molecules, NRCC workshop, Lawrence Berkeley Laboratory, University of California, 1981.

[27] R. Kincaid and H. A. Scheraga, J. Comp. Chem. (to be published).

[28] M. Mezei, Mol. Phys. **40**, 901 (1980).

[29] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, New York, 1976).

[30] A. Ben-Naim, *Water and Aqueous Solutions* (Plenum, New York, 1974).

[31] W. K. Hastings, Biometrika **57**, 1 (1970).

[32] P. H. Peskun, Biometrika **60**, 3 (1973).

[33] S. Swaminathan and D. L. Beveridge, J. Am. Chem. Soc. **99**, 8392 (1977).

[34] J. A. Barker and R. O. Watts, Chem. Phys. Lett. **3**, 144 (1969).

[35] J. C. Owicki and H. A. Scheraga, J. Am. Chem. Soc. **99**, 7403 (1977).

[36] M. H. Kalos (private communication).

[37] M. Mezei, P. K. Mehrotra, and D. L. Beveridge (to be published).

[38] S. Swaminathan, S. W. Harrison, and D. L. Beveridge, J. Am. Chem. Soc. **100**, 5705 (1978).

[39] P. K. Mehrotra and D. L. Beveridge, J. Am. Chem. Soc. 102, 4287 (1980).

[40] P. K. Mehrotra, F. T. Marchese, and D. L. Beveridge, J. Am. Chem. Soc. **103**, 672 (1981).

[41] M. Mezei and D. L. Beveridge, J. Chem. Phys. **74**, 6902 (1981).

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

[43] G. Ravi-Shanker and D. L. Beveridge (to be published).

[44] A. Geiger, A. Rahman, and F. H. Stillinger, J. Chem. Phys. 70, 263 (1979).