Aggregation-volume-bias Monte Carlo simulations of vapor-liquid nucleation barriers for Lennard-Jonesium

Bin Chen[a)]
Center for Molecular Modeling and Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, and Departments of Chemistry and of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455-0431

J. Ilja Siepmann
Departments of Chemistry and of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455-0431

Kwang J. Oh and Michael L. Klein
Center for Molecular Modeling and Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

(Received 9 August 2001; accepted 20 September 2001)

A combination of the aggregation-volume-bias Monte Carlo algorithm and the umbrella sampling technique is applied to investigate homogeneous vapor–liquid nucleation. This combined approach is simple, general, and robust. Its efficiency is demonstrated for nucleation of Lennard-Jonesium, for which the precise calculation of the nucleation barriers takes only a few minutes at higher supersaturations to a few hours at lower supersaturations. Comparison of the simulation results to the classical nucleation theory (CNT) shows that CNT overestimates the barrier heights by a value nearly independent of the supersaturation, but provides a reasonable description of the critical cluster sizes. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1417536]

I. INTRODUCTION

A supersaturated vapor is a metastable phase and the spontaneous formation of liquid droplets (nucleation) initiates the relaxation process to the more stable liquid phase. Being of both fundamental and practical importance, nucleation has been a popular scientific topic for a long time. The theoretical studies of nucleation phenomena began with the pioneering research by Volmer and Weber[1] and by Berker and Döring.[2] More recently, the development of several sophisticated experimental techniques has brought new momentum to this field.[3–5]

It is well known that vapor–liquid nucleation is a thermally activated process characterized by a nucleation free-energy (NFE) barrier at the critical cluster which separates the liquid phase from the vapor. The classical nucleation theory (CNT)[6] predicts the free-energy barrier and the critical cluster size in terms of macroscopic properties that are accessible by experiments, including the liquid density and the surface tension. However, it is assumed that the cluster behaves as an infinite bulk liquid phase with regard to density and surface tension, which contradicts the fact that clusters are finite objects. In addition, both experimental[7] and density functional studies of nucleation[8,9] have shown that a temperature-dependent factor is needed for the CNT to yield nucleation rates close to experiments.

During the past 10 years, significant progress has been made in improving the molecular-level understanding of nucleation phenomena using molecular simulations.[5,10–23] These simulations are useful in that they allow to provide rich microscopic information for the model used, test the assumptions made by the CNT, and also compare directly to density functional theory for a consistent model. However, while the accuracy of the simulation results depends upon the approximations that are made in the force field used to mimic the true physical system, the precision of the simulation results depends solely on the extent to which the phase space of the system is sampled throughout the simulation.[24] In particular, the precise determination of the NFE barrier via molecular simulation is a challenging task. In general, one needs to perform an expensive calculation to determine the NFE curve as function of cluster size (up to a size slightly larger than the critical cluster size), or the equivalent equilibrium cluster distribution. Conventional simulations are hampered by the following: (i) the instability of the supersaturated vapor phase (metastable in nature); (ii) the extremely low probabilities for the occurrence of clusters near the critical nucleus size; and (iii) the extremely slow transfer (low exchange rates) of particles between clusters of different sizes due to the large separation between clusters.

To overcome the first two problems, umbrella-sampling techniques[25] are often employed which make use of an additional biasing potential to favor formation of certain clusters and/or of several simulations that sample only a limited ''window'' of cluster sizes.[5,15–17,23] In addition, different maximum displacements for free particles in the vapor and for those in clusters can be used in Monte Carlo simulations to speed up the diffusion of vapor particles, thereby increasing the rates of cluster formation and destructions.[5,22] It should be noted that the ''small-system'' grand canonical ensemble[15–17] appears to avoid the exchange problem since

a)Author to whom correspondence should be addressed. Electronic mail: binchen@cmm.chem.upenn.edu

the whole system is assumed to be one single cluster (not including the ideal gas box) and the sampling of cluster sizes is simply done by particle insertions and removals. Thus it is a very efficient scheme, in particular, for strongly associating fluids. For nonpolar fluids, however, many of the insertions and removals (randomly inside an arbitrary volume) are rejected because they fail to maintain the cluster-size criterion.[23]

Recently, the aggregation-volume-bias Monte Carlo (AVBMC) algorithm has been developed to enhance the efficiency of sampling particle exchanges between free monomers and clusters and between two clusters.[26,27] The AVBMC algorithm is implemented here for various types of ensembles used in nucleation simulations. In the next section, some theoretical background on nucleation is summarized. The novel simulation approach obtained by combining the AVBMC algorithm with the umbrella sampling technique is presented in Sec. III. Thereafter, the simulation details of this study are briefly described. The simulation results for Lennard-Jonesium are presented and compared to the CNT and to earlier simulation results in Sec. V, and Sec. VI provides concluding remarks.

II. THEORETICAL BACKGROUND ON NUCLEATION

A consistent definition of clusters is required for all nucleation studies. Most often an explicit criterion is used, which needs to be satisfied by all particles that belong to a given cluster, e.g., the Stillinger-type cluster criterion defines a cluster as a group of molecules of which every molecule has at least one neighbor in the group within a distance of r_{cl}.[13,14,23,28] For simplicity, this definition is also used in this work and r_{cl} was set to 1.5 σ, a typical value for Lennard-Jonesium. After selecting the cluster criterion, the partition function of an i-mer (composed by n monatomic particles), q_i, can be defined as follows:[18,23]

$$q_i=\frac{1}{i!\Lambda^{3i}}\int_V H(\mathbf{r}^i)\exp[-\beta U_i(\mathbf{r}^i)]d\mathbf{r}^i, \tag{1}$$

where Λ is the thermal de Broglie wavelength, $\beta=1/k_BT$ is the inverse temperature, U_i stands for the cluster potential energy specified by the $3i$ position coordinates $\mathbf{r}^i$, and $H(\mathbf{r}^i)$ is a unit step function which is equal to 1 when the cluster criterion is satisfied, or 0 otherwise. Considering that the molecules only move within a small region v around the center-of-mass (COM) of the cluster due to the cluster criterion, Eq. (1) is most often rewritten through a coordinate transformation from the laboratory coordinate system to a COM coordinate system as follows[18,23]

$$q_i=\frac{i^3V}{i!\Lambda^{3i}}\int_v H(\mathbf{r}'^{i-1})\exp[-\beta U_i(\mathbf{r}'^{i-1})]d\mathbf{r}'^{i-1}. \tag{2}$$

The extra terms of i^3 and V arise from the Jacobian of the coordinate transformation and from the integration over the COM coordinates [i.e., the dimensionality of the configurational integral in Eq. (2) is reduced by three position coordinates], respectively.

The canonical partition function for a supersaturated vapor is given by

$$Q=\frac{1}{N!\Lambda^{3N}}\int_V \exp[-\beta U_N(\mathbf{r})]d\mathbf{r}, \tag{3}$$

where N is the total number of molecules and U_N is the total system energy which typically is a sum over all the pair interactions. It has been shown by Oh and Zeng[23] that from the viewpoint of physical cluster theory this equation can be rewritten in terms of the above i-mer partition function as follows

$$Q=\sum_{\{n_i\}}\left(\Pi_{i=1}^{i_{\max}}\frac{q_i^{n_i}}{n_i!}\right)\times\left[1+\sum_i^{i_{\max}}(n_i^2-n_i)F_{ii}+\sum_i^{i_{\max}}\sum_{j\neq i}^{i_{\max}}n_in_jF_{ij}+\cdots\right], \tag{4}$$

where $i_{\max}$ is the upper bound of the cluster size (a constraint) to maintain the supersaturated vapor in metastable equilibrium, the first term is the ideal gas term for a mixture of clusters with size distribution $\{n_i\}$, and F_{ij} is the first-order correction to the ideal gas partition function to account for the intercluster interactions involving two clusters. Higher order terms contributed by the intercluster interactions involving more than two clusters are not shown in Eq. (4).

The most probable cluster-size distribution is then obtained by maximizing Q in Eq. (4) or minimizing the corresponding Helmholtz free energy A of the system with respect to $\{n_i\}$ under the constraint $N=\Sigma_i{}^{i_{\max}}i\times n_i$ using the method of Lagrange undetermined multipliers. This process also yields the chemical potential of an i-mer at its partial pressure p_i,

$$\mu_i(p_i)=\left(\frac{\partial A}{\partial n_i}\right)_{T,V,n_{j\neq i}}=k_BT\ln\frac{n_i}{q_i}-k_BT(2n_i-1)F_{ii}-k_BT\sum_{j\neq i}^{i_{\max}}n_jF_{ij}=i\mu_1(p_1). \tag{5}$$

The constant-pressure NFE of an i-mer, ΔG_i, can be defined in two different ways. By viewing the formation of an i-mer from i monomers as a reaction, it can be defined as the reversible work in this process at the partial pressure of monomer p_1,[11,23]

$$\Delta G_i=\mu_i(p_i)-i\mu_1(p_1). \tag{6}$$

In addition, ΔG_i can also be defined as the chemical potential difference of an i-mer at partial pressure p_i and p_1[23]

$$\Delta G_i=\mu_i(p_1)-\mu_i(p_i). \tag{7}$$

These two definitions are equivalent since $\mu_i(p_i)=i\mu_1(p_1)$ as shown above [see Eq. (5)]. Applying the ideal gas approximation to $\mu_i(p_1)$

$$\mu_i(p_1)\approx\mu_i(p_i)+k_BT\ln\frac{n_1}{n_i} \tag{8}$$

we then obtain

$$\Delta G_i = k_B T \ln \frac{n_1}{n_i}. \tag{9}$$

It should be mentioned here that the interactions between the clusters are neglected in the grand canonical ensemble nucleation simulations. This is only valid for very low density phases where the ideal gas approximation holds. For those systems where the interactions between the clusters are important, those interactions contribute to μ_i and correspondingly ΔG_i as well. Such contributions can be approximated in a mean-field way (in particular, for the dominant monomer-cluster interactions) as shown by Oh and Zeng.[23]

III. SIMULATION METHODS

A. Umbrella sampling (Ref. 25)

As an equilibrium property, the cluster distribution, $\{n_i\}$, and the corresponding free energies of clusters, ΔG_i, can be computed from both Monte Carlo and molecular dynamics simulations.[5] However, due to the low probabilities of the clusters near the barrier height [$\mathcal{O}(10^{-30})$ for typical experimental conditions], without any bias it would require an exceedingly long simulation to achieve sufficient sampling of these clusters, which is unfortunately required for the convergence of the cluster distributions. This problem can be solved by the umbrella sampling scheme,[25] in which a biasing potential is used to enhance the frequencies of the clusters and the bias can be removed later in the analysis of the cluster distributions. ten Wolde and Frenkel[5] have employed the umbrella-sampling technique to study the vapor–liquid nucleation for a Lennard-Jones system, in which the biasing potential is chosen to be a harmonic function of the size of the largest cluster, n, present in the system

$$W_{\text{bias}} = \tfrac{1}{2} k (n - n_0)^2 \tag{10}$$

each choice of k and n_0 limits the sampling to a ''window'' of cluster sizes and multiple windows are required to compute the NFE barrier (15 windows were used in Ref. 5).

It should be noted that one can reduce the number of windows (simulation runs) to only one by enhancing the frequencies of all cluster sizes of interest in the same window. For example, for those simulations in which our attention is only on one cluster (the concept used by the grand canonical ensemble, which is also transferable to other ensembles, see Sec. III B), this can be achieved by using a biasing potential that is the negative of the NFE of this cluster, $-\Delta G(i)$, to allow for equal probabilities of all sizes of interest sampled for this cluster by the simulations. However, the input $\Delta G(i)$ is not known beforehand and in fact is what needs to be determined in the simulation. Therefore, the biasing potential, $f_{\text{bias}}(i)$, is solved iteratively using a self-adapting procedure.

First, simulations were carried out without any biasing potential [or $f_{\text{bias}}(i) = 0$] at a higher supersaturation (S_{high}) where the NFEs and, in particular, the NFE barrier heights are lower. The preliminary nucleation free energy curves were then used as an input to construct the initial $f_{\text{bias}}(i)$ applied for a lower supersaturation (S_{low}) as follows:

$$f_{\text{bias}}^{S_{\text{low}}}(i) = -\Delta G^{S_{\text{high}}}(i) - (i-1) \times k_B T \ln(S_{\text{high}}/S_{\text{low}}) \tag{11}$$

in which it is assumed that the configuration integral (or the Helmholtz free energy) of a cluster is independent of the supersaturation (or the chemical potential) of the gas phase.[15] Thereafter $f_{\text{bias}}(i)$ was allowed to self-adapt using relatively short simulations until convergence.

Finally, the bias can be removed in the calculation of the NFE as follows:

$$\Delta G(i) = -k_B T \ln \frac{n_i}{n_1} - f_{\text{bias}}(i), \tag{12}$$

where n_i and n_1 represent the number of times that the cluster we focus on appears as an i-mer or a monomer, respectively, in the whole simulation.

B. Aggregation-volume-bias Monte Carlo (Refs. 26 and 27)

Nucleation systems are microheterogeneous, characterized by the presence of a spectrum of microphase regions (e.g., monomer and clusters). These microphase regions may differ to a great extent on both energetic and entropic factors. In contrast, the transition probabilities used in the conventional Metropolis scheme[29] lack the balance of these two factors: (i) for cluster formation/growth, the transition probability is dominated by the entropic term (the low probability to attempt a random move that places a particle in the vicinity of another particle); and (ii) for cluster destruction/shrinkage, it is dominated by the energetic term (the energy penalty encountered when a particle leaves a cluster). Therefore, conventional Metropolis Monte Carlo simulations might either miss forming clusters or get trapped once a cluster configuration is found. This problem is alleviated by the AVBMC algorithm[26,27] in which both Boltzmann weight (energetic factor) and volume (entropic factor) terms are taken into account in the transition probability. In order to enhance the efficient particle hopping between the clusters and the vapor phase, a biased swap move is used by the AVBMC algorithm. Additionally the space surrounding a molecule is explicitly divided into in (bonded) and out (nonbonded) regions, which allows for other molecules to directly hop between these two regions via the swap moves and to achieve more efficient sampling of all relevant clusters.

First, we describe the aggregation-volume-bias Monte Carlo technique used for the conventional canonical and isobaric–isothermal ensemble nucleation simulations. Starting from a given nuclear configuration, state $\mathcal{A}$, an AVBMC swap (in this case, intrabox swap) move proceeds as follows: (i) randomly select a particle j and another particle k with $j \neq k$ and whose in (bonded) regions do not overlap (otherwise, randomly select particle k again until this condition is satisfied); (ii.a) with a probability of P_{bias}, randomly select one particle either from the in region of particle k (with a probability of P_{select}) or from the out (nonbonded) region of the particle j (with a probability of $1 - P_{\text{select}}$), and move it to the in region of particle j; or (ii.b) with a probability of $1 - P_{\text{bias}}$ select one particle randomly inside the in region of particle j and move it either to the in region of particle k

(with a probability of P_{select}) or to the out region of particle j (with a probability of $1-P_{\text{select}}$); (iii) calculate the potential energy difference between the initial state $\mathcal{A}$ and the trial state $\mathcal{B}$, $\Delta E=E_{\mathcal{B}}-E_{\mathcal{A}}$, and (iv) accept this move with the following set of acceptance probabilities:

(a) If it is an out $\rightarrow$ in move, then the following acceptance rule is used:

$$\text{acc}(\mathcal{A}_{\text{out}}\rightarrow\mathcal{B}_{\text{in}}) = \min\left[1,\frac{(1-P_{\text{bias}})\times V_{\text{in}}\times N_{\text{out}}\times\exp(-\Delta E/k_BT)}{P_{\text{bias}}\times V_{\text{out}}\times(N_{\text{in}}+1)}\right], \tag{13a}$$

where V_{in} and V_{out} are the volumes of the bonded and nonbonded regions, respectively, which sum to V. N_{in} and N_{out} are the number of molecules in the bonded and nonbonded regions, respectively, which sum to $N-1$ (j is excluded).

(b) If it is an in $\rightarrow$ out move, then the following acceptance rule is used:

$$\text{acc}(\mathcal{A}_{\text{in}}\rightarrow\mathcal{B}_{\text{out}}) = \min\left[1,\frac{P_{\text{bias}}\times V_{\text{out}}\times N_{\text{in}}\times\exp(-\Delta E/k_BT)}{(1-P_{\text{bias}})\times V_{\text{in}}\times(N_{\text{out}}+1)}\right]. \tag{13b}$$

(c) If it is an in $\rightarrow$ in move, e.g., move one particle from the in region of particle j to the in region of particle k, the following acceptance rule is used:

$$\text{acc}(\mathcal{A}_{\text{in}}\rightarrow\mathcal{B}_{\text{in}}) = \min\left[1,\frac{P_{\text{bias}}\times V_{\text{in}}^{k}\times N_{\text{in}}^{j}\times\exp(-\Delta E/k_BT)}{(1-P_{\text{bias}})\times V_{\text{in}}^{j}\times(N_{\text{in}}^{k}+1)}\right]. \tag{13c}$$

(d) If it is a particle swap from the in region of particle k to the in region of particle j, the indices j and k as well as the biasing probabilities P_{bias} and $1-P_{\text{bias}}$ in Eq. (13c) need to be switched.

The grand canonical ensemble was first used by Kusaka *et al.*[15–17] for nucleation simulations. Unlike the canonical ensemble or the isobaric–isothermal ensemble where both vapor particles and clusters coexist in the same phase, the system (excluding the fictitious ideal gas phase) contains only one physical cluster. Therefore, this method requires a much smaller system size and is computationally very efficient. Another advantage of this method is that the cluster size distribution is sampled by particle insertions or removals with the conventional grand canonical acceptance rules, which take both energetic (cluster energy) and entropic (ideal gas chemical potential) factors into account. Therefore, it has partially overcome the problems encountered in conventional nucleation simulations in the canonical ensemble or the isobaric–isothermal ensemble (i.e., the low probabilities of formations and destructions of the clusters). However, since all density fluctuations contribute to the free energy of a cluster in this scheme, the fluctuations of the center of mass of the cluster with respect to the simulation cell need to be analyzed carefully or completely suppressed to avoid redundant counting of configuration space.[18,23,30] In addition, the simulation results can be cell-size dependent and thus care should be taken in the choice of the size of the simulation cell.[23] Certain cluster configurations (i.e., the spatially extended ones) may be excluded for the smaller cell sizes. Whereas, use of too large cell sizes may greatly decrease the acceptance rates for particle insertions (and correspondingly also for particle removals) due to the lower probabilities to find a position that is energetically favorable (for the original implementation of this method on water) or a position that satisfies the cluster criterion (e.g., for Lennard-Jonesium).[15,23,31] It has been demonstrated that this method is particularly efficient for associating fluids, such as water.[15] However, it is less efficient for non-polar molecules which do not associate.[23]

Since the AVBMC algorithm[26,27] allows the swap move to directly target at a certain region (e.g., the bonded region), it can avoid the arbitrary choice of the cell size for the grand canonical ensemble method. The following implementation of the AVBMC algorithm employs an interbox swap move (between the fictitious ideal gas phase and the cluster) sequence. Starting from a given configuration of the cluster (say N-mer, with $N\geqslant 1$), state $\mathcal{A}$, an AVBMC swap (interbox swap) move proceeds as follows: (i) randomly select the move type as an insertion or a deletion (with equal probabilities); (ii) randomly select a particle j that acts as the target for the swap move; (iii) if it is selected as an insertion move, then a particle (chosen from the ideal gas phase) is placed randomly inside the in region of particle j, or otherwise, if it is a deletion, then a particle to be deleted is chosen randomly from the in region of particle j; (iv) calculate the potential energy difference between the initial state $\mathcal{A}$ and the trial state $\mathcal{B}$, $\Delta E=E_{\mathcal{B}}-E_{\mathcal{A}}$; and (v) accept this move with the following set of acceptance probabilities:

(a) If it is an insertion, then the following acceptance rule is used:

$$\text{acc}(\mathcal{A}\rightarrow\mathcal{B}) = \min\left[1,\frac{V_{\text{in}}\times N\times\exp(\mu/k_BT)\times\exp(-\Delta E/k_BT)}{(N_{\text{in}}+1)\times(N+1)}\right], \tag{14a}$$

where μ is the chemical potential of the fictitious ideal gas phase.

(b) If it is a deletion, then the following acceptance rule is used:

$$\text{acc}(\mathcal{A}\rightarrow\mathcal{B}) = \min\left[1,\frac{N_{\text{in}}\times N\times\exp(-\mu/k_BT)\times\exp(-\Delta E/k_BT)}{V_{\text{in}}\times(N-1)}\right]. \tag{14b}$$

It is evident that the underlying transition matrix of the Markov chain for both versions of the AVBMC algorithm is asymmetric and to ensure the detail balance condition, the acceptance rules have taken into account the different transition probabilities for different types of swap moves. By its

construction, the AVBMC algorithm is able to balance the different energetic and entropic factors exhibited by the microphase regions (i.e., bonded and nonbonded regions) by (i) attempting more frequently moves that lead to lower entropy and energy (cluster formation/growth); and by (ii) enhancing the acceptance rates for moves that lead to higher entropy and energy (cluster destruction/shrinkage).[26,27] Therefore, the AVBMC algorithm allows for more efficient sampling of all the relevant microphase regions (monomers and clusters). It should be noted that the grand canonical version of the AVBMC algorithm is similar to the approach used by ten Wolde *et al.*[18] to study the nucleation of polar fluids, in which, particle insertions and removals are targeted at a spherical volume around the center of mass of the cluster (also notice the similarities in the acceptance rules). Both schemes avoid the arbitrary choice of the cell volume and thus allow to overcome the cell-size dependence problem in the original grand canonical method. The scheme of ten Wolde *et al.* avoids redundant counting of configuration space by targeting the move at a volume around the center-of-mass, whereas the AVBMC method solves this problem by targeting the move at a volume centered on a randomly chosen particle (and thus the COM of both monomers and clusters have the same amount of accessible volume). The difference between them is that the AVBMC scheme further restricts the volume to the associating region and higher acceptance probabilities are expected. In particular, the in region can be chosen to conform to the cluster criterion, and thus checking of the cluster criterion for the particle insertion is no longer required (note that the system in the grand canonical system should always satisfy the cluster criterion for both particle insertions and deletions, see Sec. IV on how to maintain the cluster criterion).

Finally, it should be pointed out that the second AVBMC scheme allows a straightforward transplantation of the underlying concept of the grand canonical ensemble to the canonical and isobaric–isothermal ensembles for the NFE calculations. In this case, the two systems (the cluster and the gas phase) which are separated in the grand canonical ensemble are brought into contact with each other. An obvious advantage of combining the two systems together is that the interactions between the clusters, which are neglected in the grand canonical ensemble simulations, are properly taken into account in the calculations of the NFE using the second AVBMC scheme. In contrast to conventional nucleation simulations in the canonical or the isobaric–isothermal ensembles, however, only *one* cluster is the central focus in these AVBMC simulations, and information about the whole system (such as cluster-size distributions) is deduced solely from the configurations obtained for the central cluster (such as the fluctuations of the cluster size).

Similar to the grand canonical ensemble, efficient sampling of the fluctuations in the size of the central cluster is facilitated by the AVBMC moves. The AVBMC move sequence is identical to the one used in the grand canonical ensemble except that the acceptance rules need to be changed since the ideal gas phase used in the grand canonical ensemble is replaced by a real gas phase. There are two types of AVBMC moves: particle ''insertion'' and ''deletion.'' In the ''insertion'' move, one particle is randomly selected from the volume outside the cluster and positioned into the bonded volume of a target particle randomly selected from the cluster. Therefore, the chemical potential term is substituted by $(N-n)/V$ for the acceptance rule shown in Eq. 14(a). In the ''deletion'' move, one particle is randomly selected from the bonded volume of a particle randomly selected from the cluster and positioned into the volume outside the cluster. Accordingly, for the acceptance rule shown in Eq. 14(b), the chemical potential term is substituted by $V/(N-n+1)$. Notice that the volume outside the cluster is approximated by the volume V of the whole system; a very minor approximation considering the very small fraction of the total volume that is occupied by the cluster (the exact determination of the cluster volume is cumbersome and also unnecessary). In addition, after any deletion, the remaining particles in this cluster should still satisfy the cluster criterion and the cluster size decreases by one. To ensure the reversibility between the deletion and insertion moves, each insertion can only increment the cluster size by one. Thus, if an insertion would result in placing a particle inside the bonding region of a gas phase particle in addition to the bonding region of the chosen target particle belonging previously to the cluster, then it should be rejected.

IV. SIMULATION DETAILS

All simulations described here were performed for simple Lennard-Jonesium system. For all simulations, the bonded V_{in} region for the AVBMC swap moves was bounded by a radius of 1.5σ, the same distance used for the Stillinger cluster criterion. Both P_{bias} and P_{select} were set to 0.5 for the one simulation in the canonical ensemble.

A. Nucleation free-energy barrier as function of cluster size

Three sets of simulations were carried out. First, an AVBMC simulation in the canonical ensemble was carried out at $T^*=0.67$ and $\rho^*=0.023$. This state condition was also used by Oh and Zeng[22] in a canonical ensemble Monte Carlo simulation. Since we will directly compare to their simulation results, we followed their prescription of using a maximum cluster-size cutoff of $i_{\max}=30$ and a potential truncation at 4.5σ. Only particle swap moves were employed to sample the phase space. To enforce that no cluster in the system has a size larger than $i_{\max}=30$, a cluster analysis (a calculation of the cluster sizes present in the system) was performed for every swap move which has been accepted using the normal acceptance rules [see Eq. (13)]. If a particle swap move leads to a cluster with a size larger than $i_{\max}=30$, it is then rejected. The system consisted of 864 particles and the simulation length for the production period was 10^4 Monte Carlo cycles.

The second set of simulations were carried out using both AVBMC and umbrella sampling techniques in the grand canonical ensemble at $T^*=0.7$ and $n_v=5.75\times10^{-3}$ (the chemical potential is specified in terms of n_v, the number density of the ideal gas phase), $T^*=0.75$ and $n_v=8.2\times10^{-3}$, and $T^*=0.8$ and $n_v=1.1\times10^{-2}$. The full LJ potential was used. Only particle swap moves were employed to

sample the phase space and they belong to either particle insertions or deletions. Since the particle insertion is targeted at a position within 1.5σ away from one of the particles in the original cluster, the Stillinger cluster criterion is maintained. Therefore, each insertion increases the cluster size by one and the calculation of the biasing potential is straightforward. However, the cluster criterion may not be satisfied after the particle deletion. Thus, checking of the cluster criterion is required but only applied to those particle deletions which have been accepted using the regular acceptance rules [see Eq. (14)]. If the remaining system does not satisfy the cluster criterion, then the particle deletion is rejected. Otherwise, this particle deletion decreases the cluster size by one and the calculation of the biasing potential is also simple. The configurational-bias Monte Carlo technique[32–34] with multiple insertions[35,36] of the first bead ($n_{\text{choice}}^{\text{first}}=10$) was used to enhance the acceptance rates for the AVBMC swap moves. A total of 600 particles (combining the fictitious ideal gas phase and the cluster) and a production length of about 5×10^4 Monte Carlo cycles were used for this set of simulations.

The last set of simulations were carried out at $T^*=0.741$ in the isobaric–isothermal ensemble using a combination of the AVBMC algorithm and the umbrella sampling technique. Simulations were performed at $p^*=0.012\,02$, $0.014\,06$, and $0.016\,11$. A temperature of $T^*=0.741$ was also used by ten Wolde and Frenkel[5] in a nucleation study employing the isobaric–isothermal ensemble and the umbrella sampling scheme. Following their prescription, the LJ potential was truncated and shifted at 2.5σ. Unlike their study, the concept used in the grand canonical ensemble was brought to these isobaric–isothermal ensemble simulations. Thus we traced only one cluster in the whole system and the biasing potential was chosen as a function of the size of this cluster. Analogous to the grand canonical ensemble simulations, the efficient sampling of the size distribution of this cluster was achieved by the second type of the AVBMC moves (described in Sec. III) and the cluster criterion is satisfied in the same way. The type of Monte Carlo move was selected at random according to the following probabilities: 0.001 (volume move), 0.499 (swap), and 0.5 (translation). For translations and volume moves, the maximum displacements were adjusted to yield acceptance probabilities of 50%. The system consisted of 864 particles and the simulation length of the production period was 5×10^4 Monte Carlo cycles.

B. Gibbs ensemble simulations

The Gibbs ensemble simulations[37–39] were carried out to compute the equilibrium chemical potential at the coexistence point at $T^*=0.7$, 0.75, and 0.8 using a cutoff distance of 3.75σ and tail corrections.[24] All systems consisted of 500 particles. Three different kinds of Monte Carlo moves were used to sample phase space: translations, volume exchanges between the two boxes, and particle swaps between the two boxes. The maximum displacements used for the translational and volume moves were adjusted to yield acceptance rates of 50%, where different maximum translational displacements were used for the vapor and liquid boxes. The frequencies of swap and volume moves were adjusted for the equilibration runs to yield approximately one accepted swap and one accepted volume move per 10 Monte Carlo cycles. The configurational-bias Monte Carlo technique[32–34] with multiple insertions[35,36] of the first bead ($n_{\text{choice}}^{\text{first}}=10$) was used to enhance the acceptance rates for the swap moves. The production run consisted of 4×10^6 Monte Carlo cycles and the statistical uncertainties were estimated by dividing the simulations into 5 blocks.

C. Surface tension

Simulations in the canonical ensemble were performed to compute the surface tension at $T^*=0.7$, 0.75, and 0.8 using a cutoff distance of 5σ. The surface tension is calculated as follows:[40,41]

$$\gamma=\frac{1}{2A}\langle 2V_{zz}-V_{xx}-V_{yy}\rangle=\frac{1}{2A}\left\langle \sum_{i<j}\left(r_{ij}-\frac{3z_{ij}^2}{r_{ij}}\right)u'(r_{ij})\right\rangle, \tag{15}$$

where $A=2L_xL_y$ is the total surface area, V_{xx}, V_{yy}, and V_{zz} are the virial tensors, r_{ij} and z_{ij} are the distance and its z-axis component between particle i and j, respectively, $u'(r_{ij})$ is the first derivative of the potential energy between i and j over their distance (or equivalently the force between them), and the angle brackets denote an ensemble average. The tail corrections for the surface tension were approximated by[42]

$$\gamma_{\text{tail}}=\frac{\pi}{4}\int_{-L/2}^{L/2}dz_\alpha\int_{-L/2}^{L/2}dz_\beta\int_{p_{\min}}^{\infty}p\times\rho(z_\alpha)\times\rho(z_\beta)\times r\times u'(r)\times\left[1-\frac{3(z_\alpha-z_\beta)^2}{r^2}\right]dp, \tag{16}$$

where L and ρ are the box dimension and the density profile along the direction normal to the interface (z axis), respectively, $p_{\min}=\{\max[0,r_{\text{cut}}^2-(z_\alpha-z_\beta)^2]\}^{1/2}$ with r_{cut} as the potential cutoff distance, and $r^2=p^2+(z_\alpha-z_\beta)^2$. The integral over p can be evaluated analytically and the integrals over the z axis can be evaluated numerically by applying the trapezoid rule using the density profile computed from the simulations. As pointed out by Mercke *et al.*,[43] the surface tension of a full Lennard-Jones potential can be estimated very accurately by simulations with a cutoff radius of at least 5σ (shorter cutoffs would lead to lower liquid densities and correspondingly lower surface tensions) and a tail correction term.

The system consisted of 2000 particles and the simulations were started from a bulk liquid phase with the saturated liquid density (obtained from the Gibbs ensemble simulations). After equilibration for more than 10^4 Monte Carlo cycles, the simulation box was elongated along the z axis to create the vapor–liquid interfaces. Then the system was equilibrated for another 10^4 Monte Carlo cycles. The production period consisted of 5×10^5 Monte Carlo cycles and the statistical uncertainties were estimated by dividing the simulations into five blocks.

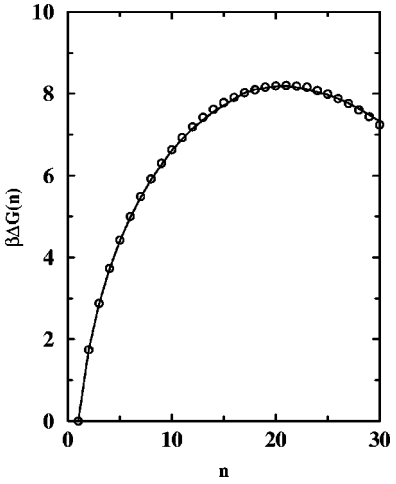

FIG. 1. Gibbs free-energy barrier as function of cluster size calculated from canonical ensemble simulations at $T^*=0.67$ and $\rho^*=0.023$. The solid line and circles represent the data from this work and by Oh and Zeng (Ref. 22), respectively.

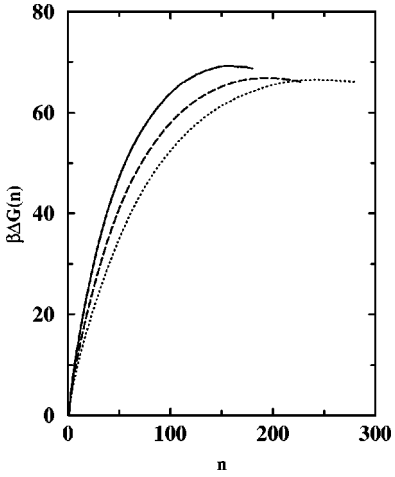

FIG. 2. Gibbs free-energy barrier as function of cluster size calculated from the grand canonical ensemble simulations using both AVBMC algorithm and umbrella sampling technique at $T^*=0.7$ and $n_v=5.75\times10^{-3}$ (solid line), $T^*=0.75$ and $n_v=8.2\times10^{-3}$ (dashed line), and $T^*=0.8$ and $n_v=1.1\times10^{-2}$ (dotted line).

V. RESULTS AND DISCUSSION

A. Nucleation free-energy barrier as function of cluster size

Plotted in Fig. 1 are the NFEs as function of cluster size from an AVBMC simulation in the canonical ensemble at $T^*=0.67$ and $\rho^*=0.023$. For comparison, the results obtained from a canonical ensemble Monte Carlo simulation by Oh and Zeng[22] at the same condition are also shown. The very good agreement between the two simulations demonstrates that the AVBMC algorithm indeed yields the correct ensemble averages. Approximately 10^7 moves were used in the AVBMC simulation, while 10^8 Monte Carlo moves were used in the simulation by Oh and Zeng.[22]

In the second set of simulations, both AVBMC algorithm and umbrella sampling technique were used. The NFE curves obtained from the grand canonical ensemble simulations at $T^*=0.7$ and $n_v=5.75\times10^{-3}$, $T^*=0.75$ and $n_v=8.2\times10^{-3}$, and $T^*=0.8$ and $n_v=1.1\times10^{-2}$ are plotted in Fig. 2. It is obvious that the critical droplet size at similar barrier heights increases with increasing temperature. With the combination of the AVBMC algorithm[26,27] and the umbrella sampling technique,[25] we are able to study nucleation systems at very low supersaturations with barrier heights larger than $70k_BT$ corresponding to a probability of $\mathcal{O}(10^{-30})$, i.e., in the range of typical conditions investigated by nucleation experiments.

It has been demonstrated (for water nucleation) by Kusaka *et al.*[15] that the NFE values at one n_v can be used to determine those values at other n_v values by applying the following expression:

$$\Delta G_{n_{v1}}(i)=\Delta G_{n_{v0}}(i)+(i-1)\times k_BT\ \ln(n_{v1}/n_{v0}) \quad (17)$$

in which, it is assumed that the Helmholtz free energy of a cluster is supersaturation independent (notice that this relation is also used in the construction of the biasing potential, see Sec. III). Actually, this is also implied by the fact that the Helmholtz free energy of an i-mer A_i is equal to $-k_BT\ln q_i$ if all the inter-cluster interactions are neglected[5,18] and both A_i and q_i [see Eq. (2)] are independent of the environment. We carried out some additional grand canonical simulations for $T^*=0.7$ at higher n_v values. 10^4 Monte Carlo cycles were used for each of these simulations and they took from 10 minutes at $n_v=1.3\times10^{-2}$ to 17 minutes at $n_v=7.5\times10^{-3}$ for a single node of a Pentium III at 866 MHz computer. The NFEs for these cases are plotted as function of the cluster size in Fig. 3. Using Eq. (17), the NFEs at $n_v=5.75\times10^{-3}$ were computed from the results obtained at the higher n_v values, and are also shown in Fig. 3. The good agreement among them and with the data obtained from the direct simulation at $n_v=5.75\times10^{-3}$, supports that the approximation of the supersaturation independence of the Helmholtz free energy of a cluster also holds for

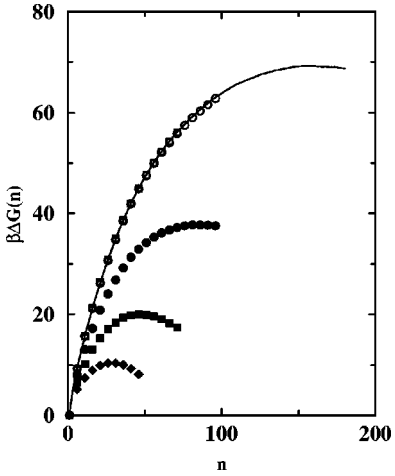

FIG. 3. Gibbs free-energy barrier as function of cluster size calculated from the grand canonical ensemble simulations using both AVBMC algorithm and umbrella sampling technique at $T^*=0.7$ and at $n_v=5.75\times10^{-3}$ (solid line), $n_v=7.5\times10^{-3}$ (filled circles), $n_v=1\times10^{-2}$ (filled squares), and $n_v=1.3\times10^{-2}$ (filled diamonds). Open circles, squares, and diamonds represent the results obtained for $n_v=5.75\times10^{-3}$ from Eq. (17) using simulation data at $n_v=7.5\times10^{-3}$, $n_v=1\times10^{-2}$, and $n_v=1.3\times10^{-2}$, respectively.

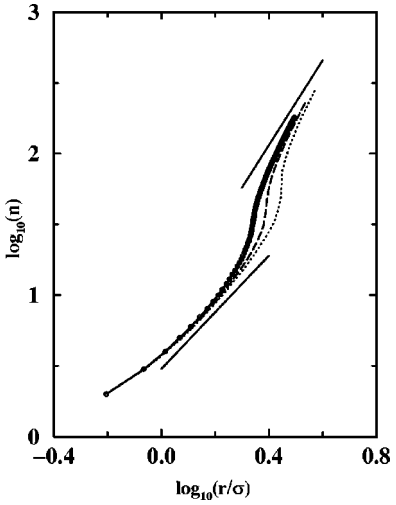

FIG. 4. Log–log plot of cluster size vs radius of gyration for $T^*=0.7$ and $n_v=5.75\times10^{-3}$ (solid line) and $n_v=7.5\times10^{-3}$ (circles), $T^*=0.75$ and $n_v=8.2\times10^{-3}$ (dashed line), and $T^*=0.8$ and $n_v=1.1\times10^{-2}$ (dotted line). The two straight lines are drawn as a guide to the eye with slopes of 2 (lower) and 3 (upper).

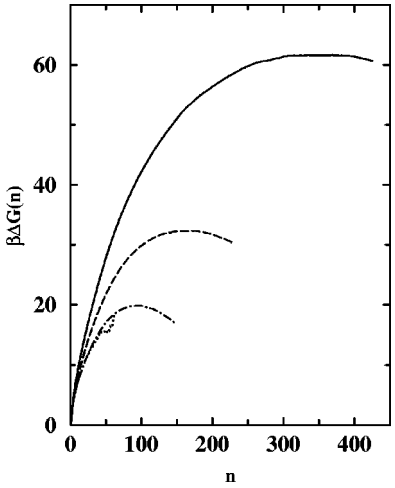

FIG. 5. Gibbs free-energy barrier as function of cluster size calculated from the isobaric–isothermal ensemble simulations using both AVBMC algorithm and umbrella sampling technique at $p^*=0.012\,02$ (solid line), 0.014 06 (dashed line), and 0.016 11 (dashed–dotted line). The data from the simulation at $p^*=0.016\,11$ without biasing potential is shown as dotted lines.

nonassociating molecules. In addition, it appears that 5×10^4 Monte Carlo cycles (a few hours of CPU time) are sufficient to achieve a converged evaluation of the NFE curve, and that one single simulation is sufficient to investigate the NFE barriers at different supersaturations.

To analyze the fractal nature of clusters smaller than the critical size, the average radii of gyration of the clusters were calculated. A log–log plot of the aggregation number versus the radius (see Fig. 4) reveals some interesting aspects of the Lennard-Jones clusters. First of all, the data obtained for different supersaturations at the same temperature agree very well. This might explain as to why a single simulation at one supersaturation is sufficient to calculate the NFE for other supersaturations at the same temperature. Second, all curves show a pronounced S-shape, i.e., the fractal dimension of the clusters does not approach smoothly the limiting value of 3. In fact, there appears to be a narrow transition region where the aggregation number increases more steeply than the third power, i.e., in this transition region the addition of further particles causes a change in the cluster structures to more compact objects (very different from a diffusion-limited growth mechanism that leads to fractal objects). The transition region shifts to larger cluster sizes with increasing temperature, similarly as the critical cluster sizes shift with increasing temperature (see Fig. 2). However, the transition region is found for aggregation numbers that are about a factor of 5 smaller than the critical cluster sizes.

The NFE data obtained from the set of isobaric–isothermal simulations at $p^*=0.012\,02$, $0.014\,06$, and $0.016\,11$ are shown in Fig. 5. For $p^*=0.01611$, the results obtained from another AVBMC simulation without the umbrella sampling technique are also included in Fig. 5. The NFEs given by this simulation without biasing appear to deviate from the converged answer for cluster sizes above 30, whose NFE is greater than 13 k_BT [which corresponds to a probability of $\mathcal{O}(10^{-6})$]. Therefore, the umbrella sampling technique with appropriate biasing potentials helps to reach a converged answer and the conventional nucleation simulations without biasing, in general, are not suitable to study the nucleation cases characterized by barrier heights much larger than $10k_BT$. It should be noted that Oh and Zeng[23] carried out a brute-force simulation (10^9 Monte Carlo moves) without umbrella sampling for one nucleation system with a barrier height close to $17k_BT$.

The nucleation barrier heights at $p^*=0.012\,02$, 0.014 06, and 0.016 11 were found to be 62, 32, and 20 k_BT, respectively, which are slightly higher than the results obtained by ten Wolde and Frenkel[5] using a combination of the umbrella sampling technique and a thermodynamic integration scheme (their values were around 57, 29, and $18k_BT$). However, a direct comparison is difficult because different cluster criteria were used and their results at $p^*=0.014\,06$ and 0.016 11 were derived from the barrier height calculated at $p^*=0.012\,02$ using a thermodynamic integration scheme. Only 5×10^4 Monte Carlo cycles and only one window were used here, which took less than 5 hours for a Pentium III at 866 MHz computer, whereas a total of 15 windows with each window of 2.5×10^5 Monte Carlo cycles were used by ten Wolde and Frenkel for their calculations at $p^*=0.012\,02$.

B. Nucleation free-energy barrier as function of supersaturation

Various definitions of the supersaturation have appeared in the literature. ten Wolde and Frenkel[5] used a supersaturation definition based on the pressure, that is, the supersaturation is given by $S=p/p_{\mathrm{coex}}$. This definition is appealing for isobaric–isothermal ensemble simulations since the pressure is specified for the simulation system. For canonical simulations, Oh and Zeng[22] used a definition of $S=\sum_{i=1}^{i_{\max}}\rho_i/\rho_{\mathrm{coex}}$, where ρ_i and ρ_{coex} are the number densities of i-mer and the saturated vapor phase, respectively. In the CNT, a chemical potential definition of the supersaturation is used: $S=\exp[-\beta(\mu_v(p)-\mu_l(p))]$, where μ_v and μ_l are the chemical potentials for the vapor and liquid phases, respec-

TABLE I. Vapor-liquid coexistence properties (ρ, μ, and n_v^{coex} are the saturated number density, chemical potential and the ideal gas-phase number density at coexistence, respectively).

T^*	ρ_{liq}^*	ρ_{vap}^*	μ_{liq}^*	μ_{vap}^*	n_v^{coex}
0.7	$0.842\,75_6$	$0.002\,032_8$	-4.370_9	-4.368_2	0.001 95
0.75	0.8213_1	$0.003\,67_3$	-4.25_1	-4.260_5	0.003 41
0.8	$0.799\,11_5$	$0.006\,10_3$	-4.152_5	-4.156_3	0.005 54

tively, at pressure p. Since the liquid density is relatively insensitive to the very small variations of the pressure above the saturation pressure, $\mu_{\mathrm{l}}(p)$ is often approximated by $\mu_{\mathrm{l}}(p_{\mathrm{coex}})=\mu_v(p_{\mathrm{coex}})$. Thus, the supersaturation can be also defined by $S=n_v/n_v^{\mathrm{coex}}$, which is straightforward for grand canonical ensemble simulations. In the following, we will focus only on the results from the grand canonical simulations and compare them to those predicted by the CNT. Therefore, the last supersaturation definition is used.

In order to calculate the supersaturation using the equation $S=n_v/n_v^{\mathrm{coex}}$, n_v^{coex} or the equivalent chemical potential at coexistence needs to be determined. Extensive Gibbs ensemble[37–39] Monte Carlo simulations were performed to calculate this quantity at $T^*=0.7$, 0.75, and 0.8, and the results are listed in Table I. It should be noted here that each of these Gibbs ensemble calculations is about 10 times more expensive than a grand canonical nucleation simulation used to determine a single NFE curve, but a precise n_v^{coex} is important to determine accurately S for comparison with the CNT. The NFE barrier height, ΔG^* is plotted versus the supersaturation in Fig. 6. With increasing temperature, the curve shifts to lower supersaturations and becomes steeper.

C. Comparison with the classical nucleation theory

The CNT[6] predicts the free-energy barrier and also the critical cluster size in terms of macroscopic properties as follows:

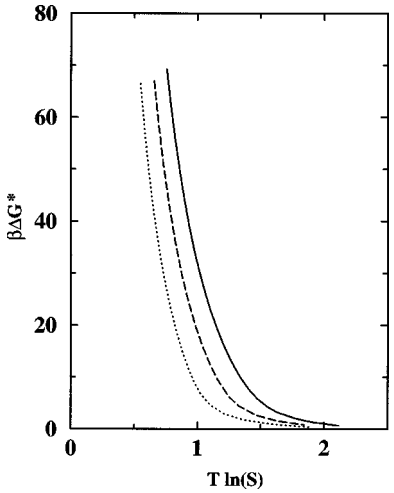

FIG. 6. Nucleation free-energy barrier height as function of supersaturation for $T^*=0.7$ (solid line), 0.75 (dashed line), and 0.8 (dotted line).

TABLE II. Surface tension data [r_c^*, γ^*, and γ_{tail}^* are the potential cutoff distance, the surface tension and the tail correction calculated from Eqs. (15) and (16), respectively].

T^*	r_c^*	γ^*	γ_{tail}^*	$\gamma^*+\gamma_{\mathrm{tail}}^*$
0.7[a]	5.0	1.032_9	0.116	1.147
0.7[b]	5.0	1.022_{19}	0.118	1.139
0.7[b]	6.5	1.075_{19}	0.070	1.145
0.75[a]	5.0	0.944_9	0.108	1.052
0.8[a]	5.0	0.816_{11}	0.103	0.918
0.8[c]	5.5	0.854	0.086	0.940

[a]This work.
[b]Results from Ref. 43.
[c]Results from Ref. 44, but γ_{tail}^* was computed using Eq. (16) and the density profile of this work.

$$\Delta G^*=\frac{16\pi\gamma_\infty^3}{3(\rho_{\mathrm{l}}k_BT\ln S)^2}, \tag{18a}$$

$$n^*=\frac{32\pi\gamma_\infty^3}{3\rho_{\mathrm{l}}^2(k_BT\ln S)^3}, \tag{18b}$$

where γ_∞ is the surface tension of the infinite planar interface and ρ_{l} is the bulk liquid density at coexistence. Both ΔG^* and n^* are proportional to the third power of the surface tension, thus a very precise determination of the surface tension is required to apply the CNT to the prediction of ΔG^* and n^*. To this respect, several long simulations were carried out to compute this property and the results are listed in Table II together with some literature data.[43,44] In general, good agreement was found between the different simulations.

Using the surface tension data calculated from our simulations, the NFE barrier height predicted by the CNT was determined from Eq. (18a) and plotted as function of $(T\ln S)^{-2}$ in Fig. 7. As predicted by the CNT, the NFE barrier height is proportional to $(T\ln S)^{-2}$ and thus the CNT curves in Fig. 7 are straight lines with the slope governed by the surface tension and the bulk liquid density at

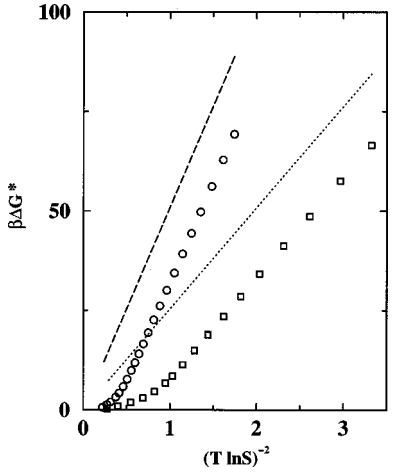

FIG. 7. Nucleation free-energy barrier height as function of $(T\ln S)^{-2}$ for $T^*=0.7$ and 0.8. The simulation data are depicted as circles ($T^*=0.7$) and squares (0.8), and the CNT predictions are represented by dashed (0.7) and dotted (0.8) lines.

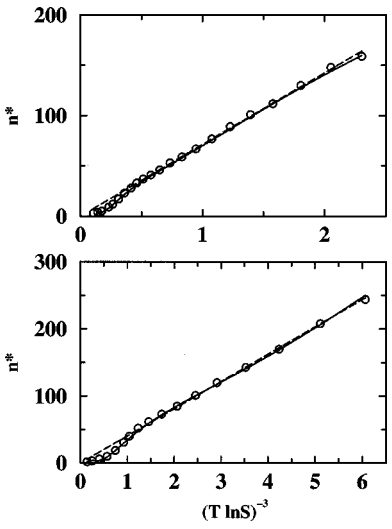

FIG. 8. Comparison of the excess number of particles (solid lines) and the critical cluster size (circles) calculated from the simulations to the critical cluster size predicted by the CNT (dashed lines) as function of $(T \ln S)^{-3}$ for $T^*=0.7$ (top) and 0.8 (bottom).

coexistence. In agreement with the CNT prediction, the curves determined from the nucleation simulations are also straight (see Fig. 7) at lower supersaturations but show a sharp turning point at higher supersaturations. The reason behind the deviations at higher supersaturations is that the critical cluster sizes become very small and, as discussed earlier, these small clusters are far from spherical. The slopes of the straight parts of the simulation curves are very close to the CNT predictions. However, there is a significant offset between the two curves, with the barrier heights determined from the simulations consistently lower than the CNT predictions for all supersaturations and all temperatures. Similar observations were also reported from density functional calculations[8,9] and other simulations.[5] In particular, using a nonuniform spherical droplet model, McGraw and Laaksonen[9] explained that the barrier height displacement is caused by the interfacial curvature free energy as follows:

$$\Delta G^* - \Delta G^*_{\mathrm{CNT}} = 4\pi k_s, \tag{19}$$

where $4\pi k_s$ is the interfacial curvature free energy with k_s as the rigidity coefficient. From density functional calculations, McGraw and Laaksonen[9] estimated $k_s=-1.576k_BT$ at $T^*=0.8$ which corresponds to an offset of $19.8k_BT$. In comparison, our simulation data yield an offset of about $17.7k_BT$ averaged over the lower supersaturations. In addition, Fig. 7 also shows that this offset shifts to larger values at lower temperatures (about $19.1k_BT$ at $T^*=0.7$), which again agrees with McGraw–Laaksonen predictions.

As stated in the nucleation theorem,[45–51] the excess number of molecules Δn^* of the critical nucleus is equal to the negative derivative of ΔG^* with respect to μ_v. Thus the Δn^* values at various supersaturations were determined from the slopes of the curves shown in Fig. 6 (through a polynomial interpolation of the curves). In Fig. 8 these Δn^* values are compared to the critical nucleus sizes predicted by the CNT and also the critical cluster sizes estimated directly from the NFE curves (these numbers should be close to each other at low vapor densities). In contrast to the NFE barrier heights, very good agreement was found between the excess number of particles in the critical nucleus and the critical cluster size determined from the simulations and the critical cluster size predicted by the CNT at lower supersaturations (the deviations at higher supersaturations are again caused by the small critical clusters). These observations agree with the predictions by McGraw and Laaksonen[9] and the simulation results by ten Wolde and Frenkel.[5]

VI. CONCLUSIONS

A novel Monte Carlo method is presented for the efficient simulation of vapor–liquid nucleation. This method is developed by integrating the aggregation-volume-bias Monte Carlo (AVBMC) algorithm into the other nucleation simulation approaches to enhance the sampling efficiency of cluster formations and destructions. To achieve efficient sampling of all relevant clusters, a new type of move (AVBMC swap) is introduced which allows particles to move in or out of the same region used to define the cluster criterion. Compared to the unbiased translations used in the conventional Metropolis scheme, the AVBMC swap move is able to balance the different energetic and entropic factors exhibited by monomer and clusters since both Boltzmann weight (energetic factor) and accessible volume (entropic factor) are taken into account in the transition probability of this move. In addition, the AVBMC swap move allows a straightforward extension of the concept used by the grand canonical ensemble to the canonical and isobaric–isothermal ensembles for nucleation barrier calculations. Here, only one cluster is singled out for the AVBMC swap moves and the umbrella biasing, and the cluster-size distribution of the whole system can be obtained from the size distribution of this cluster. To enhance the frequencies of all cluster sizes of interest, a biasing potential that is a function of all relevant cluster sizes is introduced which allows us to reduce the number of umbrella-sampling windows to only one.

The implementation of the combined AVBMC/umbrella sampling method is described for nucleation simulations in the canonical, isobaric–isothermal, and grand canonical ensembles. For all cases studied, one single simulation with a moderate length (on the order of 10^4 Monte Carlo cycles) is sufficient to calculate the entire NFE curve. In particular, for the single-cluster version of the AVBMC/umbrella sampling method (used in both grand canonical and isobaric–isothermal ensembles), it is demonstrated that the precise calculation of the nucleation barriers takes only a few minutes at higher supersaturations to a few hours at lower supersaturations, many orders of magnitude more efficient than other nucleation simulation approaches. In addition, it was found that the supersaturation independence of the Helmholtz free energy of a cluster is obeyed for the Lennard-Jones system, similar to the water system studied by Kusaka *et al.*[15] Thus, the NFE barriers at different supersaturations can be obtained from one single simulation.

Comparison of the simulation results to the CNT shows that the latter systematically overestimates the barrier heights by a value nearly independent of the supersaturation, but dependent on the temperature. However, the agreement be-

tween the simulation data and the CNT predictions for the critical nucleus size is excellent for lower supersaturations. These observations are consistent with recent density functional calculations,[8,9] and Monte Carlo simulations by ten Wolde and Frenkel.[5,18] In particular, the offset on the nucleation barrier height increases with increasing temperature with values close to those predicted by McGraw and Laaksonen.[9] This may help to shed light on the observation that the CNT needs an empirical temperature-dependent scaling factor to yield the correct nucleation rates for experimental systems.

Finally, the AVBMC/umbrella sampling approach is by no means limited to Lennard-Jonesium systems. In particular, configurational-bias Monte Carlo strategies can be employed during the AVBMC moves,[26,27] thereby allowing for the straightforward application to more complex molecules. To further pursue this part, nucleation simulations for alkanes are currently in progress.

ACKNOWLEDGMENTS

The authors thank Richard McClurg and Ken Leopold for many stimulating discussions. Financial support from the National Science Foundation CTS-9813601 (J.I.S.) and CHE-9623017 (M.L.K.), an Alfred P. Sloan Fellowship (J.I.S.), and a Doctoral Dissertation Fellowship (B.C.) is gratefully acknowledged. Part of the computer resources were provided by the Minnesota Supercomputing Institute.

[1] M. Volmer and A. Weber, Z. Phys. Chem., Stöchiom. Verwandtschaftsl. **119**, 227 (1926).
[2] R. Berker and W. Döring, Ann. Phys. (Leipzig) **24**, 719 (1935).
[3] Y. Viisanen, R. Strey, and H. Reiss, J. Chem. Phys. **99**, 4680 (1993).
[4] K. N. H. Looijmans, C. C. M. Luitjen, and M. E. H. van Dongen, J. Chem. Phys. **103**, 1714 (1995).
[5] P. R. ten Wolde and D. Frenkel, J. Chem. Phys. **109**, 9901 (1998).
[6] F. F. Abraham, *Homogeneous Nucleation Theory* (Academic, New York, 1974).
[7] M. M. Rudek, J. A. Fisk, V. M. Chakarov, and J. L. Katz, J. Chem. Phys. **105**, 4707 (1996).
[8] C. Seok and D. W. Oxtoby, J. Chem. Phys. **109**, 7982 (1998).
[9] R. McGraw and A. Laaksonen, J. Chem. Phys. **106**, 5284 (1997).
[10] C. L. Weakliem and H. Reiss, J. Chem. Phys. **99**, 5374 (1993).
[11] K. J. Oh, X. C. Zeng, and H. Reiss, J. Chem. Phys. **107**, 1242 (1997).
[12] P. Schaaf, B. Senger, and H. Reiss, J. Phys. Chem. B **101**, 8740 (1997).
[13] B. Senger, P. Schaaf, D. S. Corti, R. Bowles, J.-C. Voegel, and H. Reiss, J. Chem. Phys. **110**, 6421 (1999).
[14] B. Senger, P. Schaaf, D. S. Corti, R. Bowles, D. Pointu, J.-C. Voegel, and H. Reiss, J. Chem. Phys. **110**, 6438 (1999).
[15] I. Kusaka, Z.-G. Wang, and J. H. Seinfeld, J. Chem. Phys. **108**, 3416 (1998).
[16] I. Kusaka and D. W. Oxtoby, J. Chem. Phys. **110**, 5249 (1999).
[17] I. Kusaka, D. W. Oxtoby, and Z.-G. Wang, J. Chem. Phys. **111**, 9958 (1999).
[18] P. R. ten Wolde, D. W. Oxtoby, and D. Frenkel, J. Chem. Phys. **111**, 4762 (1999).
[19] G. T. Gao, K. J. Oh, and X. C. Zeng, J. Chem. Phys. **110**, 2533 (1999).
[20] K. Yasuoka and M. Matsumoto, J. Chem. Phys. **109**, 8451 (1998).
[21] K. Yasuoka and M. Matsumoto, J. Chem. Phys. **109**, 8463 (1998).
[22] K. J. Oh and X. C. Zeng, J. Chem. Phys. **110**, 4471 (1999).
[23] K. J. Oh and X. C. Zeng, J. Chem. Phys. **112**, 294 (2000).
[24] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Oxford University Press, Oxford, 1987).
[25] G. M. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974).
[26] B. Chen and J. I. Siepmann, J. Phys. Chem. B **104**, 8725 (2000).
[27] B. Chen and J. I. Siepmann, J. Phys. Chem. B **105**, 11275 (2001).
[28] F. H. Stillinger, J. Chem. Phys. **38**, 1486 (1963).
[29] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[30] H. Reiss, W. K. Kegel, and J. L. Katz, J. Phys. Chem. A **102**, 8548 (1998).
[31] K. J. Oh and X. C. Zeng, J. Chem. Phys. **1114**, 2681 (2001).
[32] J. I. Siepmann and D. Frenkel, Mol. Phys. **75**, 59 (1992).
[33] T. J. H. Vlugt, M. G. Martin, B. Smit, J. I. Siepmann, and R. Krishna, Mol. Phys. **94**, 727 (1998).
[34] M. G. Martin and J. I. Siepmann, J. Phys. Chem. B **103**, 4508 (1999).
[35] K. Esselink, L. D. J. C. Loyens, and B. Smit, Phys. Rev. E **51**, 1560 (1995).
[36] A. D. Mackie, B. Tavitian, A. Boutin, and A. H. Fuchs, Mol. Simul. **19**, 1 (1997).
[37] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813 (1987).
[38] A. Z. Panagiotopoulos, N. Quirke, M. Stapleton, and D. J. Tildesley, Mol. Phys. **63**, 527 (1988).
[39] B. Smit, P. de Smedt, and D. Frenkel, Mol. Phys. **68**, 931 (1989).
[40] J. S. Rowlinson and B. Widom, *Molecular Theory of Capillarity* (Clarendon, Oxford, 1982).
[41] E. Salomons and M. Mareschal, J. Phys.: Condens. Matter **3**, 3645 (1991).
[42] J. G. Harris, J. Phys. Chem. **96**, 5077 (1992).
[43] M. Mecke, J. Winkelmann, and J. Fischer, J. Chem. Phys. **107**, 9264 (1997).
[44] A. Trokhymchuk and J. Alejandre, J. Chem. Phys. **111**, 8510 (1999).
[45] Y. Viisanen, R. Strey, and H. Reiss, J. Chem. Phys. **99**, 4680 (1993).
[46] R. Strey and Y. Viisanen, J. Chem. Phys. **99**, 4693 (1993).
[47] Y. Viisanen, R. Strey, A. Laaksonen, and M. Kulmala, J. Chem. Phys. **100**, 6062 (1994).
[48] P. Strey, Y. Viisanen, and P. E. Wagner, J. Chem. Phys. **103**, 4333 (1995).
[49] D. Kashchiev, J. Chem. Phys. **76**, 5098 (1984).
[50] D. W. Oxtoby and D. Kashchiev, J. Chem. Phys. **100**, 7665 (1994).
[51] I. J. Ford, J. Chem. Phys. **105**, 8324 (1996).