Self-Adapting Fixed-End-Point Configurational-Bias Monte Carlo Method for the Regrowth of Interior Segments of Chain Molecules with Strong Intramolecular Interactions

Collin D. Wick and J. Ilja Siepmann*

Departments of Chemistry and of Chemical Engineering and Materials Science, University of Minnesota, 207 Pleasant St. SE, Minneapolis, Minnesota 55455-0431

Received January 31, 2000; Revised Manuscript Received June 16, 2000

ABSTRACT: An extension to the configurational-bias Monte Carlo method is presented which allows for the efficient conformational sampling of the interior segments of chain molecules whose interactions include strong bonded terms (governing bond stretching, bond angle bending, and dihedral angle rotation). The ability to regrow interior segments overcomes the limitations of conventional configurational-bias methods (where the regrowth is always directed to a free chain end) and now allows for the simulation of chain systems with low concentrations of chain ends, that is, higher molecular weights, networks, or cyclic structures. As previously proposed by Dijkstra et al. [*J. Chem. Phys.* **1994**, *101*, 3179] for lattice polymers and by Vendruscolo [*J. Chem. Phys.* **1997**, *106*, 2970] for freely jointed polymers, an additional biasing (closing) probability is used that guides the bead-by-bead configurational-bias regrowth of interior segments toward its desired fixed target. However, while the previous methods are limited to chain models for which the number of random walks that lead to closure is known or which rely on simpler and less efficient geometric considerations, the algorithm presented here allows for the simulation of chain molecules using force fields of arbitrary complexity for which the closing probability is not known a priori. It is important to note that the additional biasing probability used to guide the move does not necessarily have to be the true closing probability but that a good approximation thereof is essential to improve the sampling efficiency. To this extent, we obtain an intial guess of the biasing probability from a short presimulation or an earlier simulation of a related system or simply use a uniform biasing probability. A self-adapting scheme is then used to optimize the biasing probability during the course of the simulation for the system of interest. In addition to the conformational sampling of interior segments, the new algorithm also enables efficient particle insertions and removals of cyclic molecules (of moderate length) and thereby opens the door to simulations in the grand canonical and Gibbs ensembles. Simulation results are presented for linear, branched, and cyclic alkanes using the transferable potentials for phase equilibria (TraPPE) force field.

1. Introduction

The tremendous flexibility of the Monte Carlo method for particle simulations[1−4] has been exploited for the design of many efficient sampling schemes for systems composed of flexible chain molecules.[5,6] One such sampling scheme is the configurational-bias Monte Carlo (CBMC) algorithm,[7−10] which has been shown to be effective to simulate many different types of chain molecules.[11,12] However, since conventional CBMC methods rely on segment-by-segment regrowing (Rosenbluth growth)[13] of a part of the trial molecule starting from a randomly selected interior segment and moving toward a free end segment, CBMC has been found useful only for chain molecules of moderate length. The reason for this is that the CBMC attrition rate for the regrowth increases and the CBMC acceptance rate decreases dramatically with increasing number of segments, making the CBMC method unsuitable for long chain molecules at high packing fractions. Furthermore, for the special cases of cyclic molecules or cross-linked polymer networks, where there are either zero or only relatively few end segments, a conventional CBMC move cannot be used.

A solution for the problems mentioned above is to design moves that rearrange only interior segments of a chain molecule. The earliest such approach was the crankshaft move.[14−17] In the crankshaft move, one (or more) interior segment(s) of a chain is chosen at random and "cranked" around a "shaft" formed by the segments directly bonded to the selected interior segment(s), keeping the positions of the shaft segments and all bond lengths fixed. If one allows for flexible bond lengths (these bonds are just the connectors between adjacent segments and do not necessarily have to correspond to chemical bonds), then a single randomly selected interior segment can be moved to a new trial position (that is consistent with a model specific set of constraints). The bond-fluctuation method[18,19] is such a scheme for lattice polymers. These two methods have the advantage that their acceptance probabilities for an individual move are not chain-length-dependent, but the number of moves it takes to make any impact on the conformation of a large polymer becomes very high.

Over the past decade, a few methods have been developed to rearrange multiple interior segments of complex polymeric systems. Most noticeable are the concerted-rotation (CONROT) methods pioneered by Theodorou and co-workers[20−22] and recently extended to specific protein architectures by Deem and co-workers.[23−25] During a CONROT move the torsional angles of a 5−8 segment subunit of the polymer are sampled (while the bond lengths and angles are usually kept fixed) involving a rather sophisticated determination of all possible conformations of that subunit. The CONROT methods excel for molecular-mechanics type

* Corresponding author: siepmann@chem.umn.edu.

10.1021/ma000172g CCC: $19.00
Published on Web 09/01/2000

polymer models, and a recent extension exploits the polydispersity of most polymeric systems.[21,22]

In contrast, extensions of the CBMC method have been proposed for simplified lattice and freely jointed bead models of polymers. The extended continuum configurational-bias (ECCB) method developed by Escobedo and de Pablo[26] allows one to rearrange any number of interior segments of linear, branched, or cross-linked polymer chains or networks by discarding all growth attempts (setting their Boltzmann weight to zero) that would result in placing a given trial segment either too close (requiring the definition of a hard inner cutoff) or too far from the final destination of the growth (obviously considering the number of additional links that still need to be regrown). While this method works quite well for bead−spring or freely jointed polymer models, its biggest shortcoming is that for molecular-mechanics type models with full bonded and nonbonded interactions there are many growth attempts that do not get discarded despite that their final closure would require highly strained conformations (unfavorable bond lengths, bond angles, and dihedral angles). Neubauer et al.[27] have shown that use of the ECCB method is not advantageous for the simulation of systems such as cyclic alkanes.

For lattice polymers, Dijkstra, Frenkel, and Hansen[28] proposed a fixed-end-point CBMC algorithm that makes use of an additional biasing probability multiplied in the Boltzmann and Rosenbluth weights to guide each regrowth of interior segments toward its final destination. The guiding probability was calculated from the number of random lattice walks (of a given number of links) that can connect the trial site to the final closure point. This method was expanded upon by Vendruscolo,[29] who gave an expression for the number of accessible random walks for freely jointed polymers in continuum space. Nevertheless, while these schemes are very efficient for lattice and freely jointed polymers, for which a guiding probability can be determined from the number of accessible random walks, these schemes at present do not account for bonded interactions which are important for molecular-mechanics type models (favoring of certain bond lengths, bond angles, and dihedral angles).

The self-adapting fixed-end-point (SAFE) CBMC algorithm proposed here overcomes some of the limitations of the ECCB, FE-CBMC, and CONROT methods. SAFE-CBMC allows for the rearrangement of multiple interior segments for molecular-mechanics type models, but unlike the CONROT move it is possible to attempt a SAFE-CBMC move for any number of interior links and for branched polymers and polymer networks. Like FE-CBMC, SAFE-CBMC makes use of an additional guiding probability. However, in the latter case the guiding probability also accounts for special bonded interactions. In this case, an analytical solution for the number of closures is not known, and starting from an initial guess SAFE-CBMC self-adapts during the course of a simulation to arrive at suitable guiding biases.

The remainder of this paper is arranged as follows. In the next section a detailed description of the SAFE-CBMC algorithm is given, and it is shown that SAFE-CBMC satisfies the condition of superdetailed balance. The polymer model (force field) and simulation details are described in section 3. Results for SAFE-CBMC simulations in the canonical, isobaric−isothermal, and Gibbs ensembles are discussed in section 4.

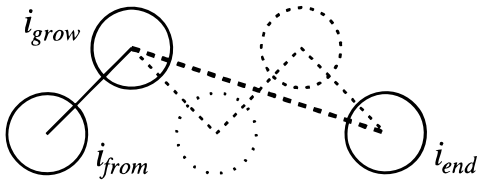

Figure 1. Growth scheme for a typical SAFE-CBMC regrowth attempt of a linear chain segment.

2. Self-Adapting Fixed-End-Point Configurational-Bias Monte Carlo

A. Statistical Mechanics. The SAFE-CBMC move follows closely the growth strategy of a regular CBMC move,[7−10] except that instead of generating the trial conformation by a segment-by-segment Rosenbluth growth that concludes with a free (unconstrained) end segment, the SAFE-CBMC trial move involves a Rosenbluth growth of interior segments that connects between one fixed starting site and one or more fixed end points. For simplicity, let us assume at this point that the SAFE-CBMC move involves the regrowth of a linear interior portion of the chain (i.e., in this case there is only one end point). Extensions to branched molecules and networks will be discussed in section 2.C.

The first step of a SAFE-CBMC trial move is the random selection of an interior segment of the chain that will be the starting site for the regrowth. (As with regular CBMC, this first selection can involve a uniform sampling probability or a predetermined nonuniform sampling probability that allows one to preferentially attempt moves of certain regions.) Thereafter, the direction (increasing or decreasing segment indices) of the Rosenbluth walk is randomly selected. The third step is the random selection of the number of segments n_{step} that need to be regrown. (Again use of uniform or predetermined nonuniform sampling probabilities is allowed.) Here we use the following indexing conventions (see also Figure 1): the fixed end point of the linear regrowth is called i_{end}; the (fixed) bead adjacent to i_{end} is denoted as i_{e+1} (this is done only if this bead is present, i.e., i_{end} is at least one bead away from the real end segment of the molecule); i_{grow} and $i_{\text{grow}+1}$ represent the segment that is presently being regrown and the segment that will be grown thereafter (if i_{grow} is not the n_{step} (last) segment of the growth), respectively; and i_{from} denotes the existing segment to which i_{grow} is directly connected. Bond length, bond angles, and dihedral angles are labeled as $L_{v/w}$, Θ_x, and $\Phi_{y/z}$, respectively, where v and w are the two adjacent segments defining the bond, x is the center segment of the three-segment group defining the bond angle, and y and z are the two segments connected by the central bond around which the dihedral rotation takes place.

Now the SAFE-CBMC Rosenbluth walk can be started. For every step n, we will consider nc_n trial sites, where a common nc_n can be used for all n, or sometimes use of different nc_n for different n enhances sampling efficiency.[30] The generation of these trial sites follows the coupled−decoupled CBMC strategy[12] that allows for the efficient sampling of the bonded interactions (bond stretch, angle bend, a dihedral rotation) that can be readily calculated. The coupled−decoupled CBMC strategy has been described in detail in a recent publication.[12] Thus, to keep the notation as simple as possible, the contributions from the bonded interactions are not included in the following description; i.e., the focus is on the nonbonded interactions and the special SAFE-

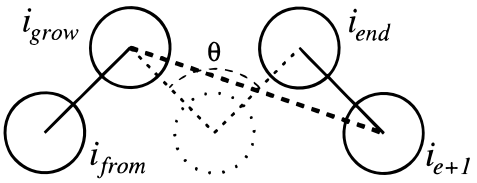

Figure 2. Growth scheme for the special case of placing the penultimate segment for a SAFE-CBMC regrowth attempt of a linear chain segment.

CBMC guiding probabilities. (A detailed description of our implementation of coupled−decoupled SAFE-CBMC is given in the Appendix.) For the biased selection of site i of the nc_n trial sites as the next growth segment, the usual Boltzmann weight of the nonbonded interactions *and* the SAFE-CBMC guiding probability are used as follows

$$P_{n,i} = \frac{\exp(-\beta u_{n,i}) g_{n,i}}{W_n} \tag{1}$$

where $\beta = (k_B T)^{-1}$ with k_B and T being Boltzmann's constant and the absolute temperature, respectively; $u_{n,i}$ and $g_{n,i}$ are the nonbonded potential energy and the guiding probability of site i at step n. As in conventional CBMC, the nonbonded potential energy u_i also includes terms arising from intramolecular interactions (say, van der Waals or Coulombic) with all beads that are present at step n, thus including nonbonded interactions with segments that have already been regrown and other intramolecular segments (such as i_{end}) that do not change position during the trial move. In our case, the guiding probabilities use only explicit distance dependencies, and methods for the determination of the guiding probabilities will be discussed in section 2.B. The Rosenbluth weight W_n at step n is given by

$$W_n = \frac{\sum_{j=1}^{nc_n} \exp(-\beta u_{n,j}) g_{n,j}}{nc_n} \tag{2}$$

Many Monte Carlo simulations using molecular-mechanics type force fields for chain molecules employ a semiflexible model where the bond lengths are fixed, while bond angles and dihedral angles are allowed to vary. In this case, special attention has to be paid when the trial segment i_{grow} is two segments away from i_{end}. Here, the placement of i_{grow} determines the bond angle $\theta_{i_{grow}+1}$, and the corresponding Boltzmann weight of its bending energy can be considered as its guiding probability (see Figure 2). To increase the acceptance probability, though, the distance dependence of i_{grow} on i_{e+1} (which would be a 1−4 or torsional interaction in molecular-mechanics type force fields) is also included as the guiding probability. To grow the last bead needed to connect a semiflexible chain, a crankshaft move is required. This crankshaft move is implemented simply by fixing i_{from} and i_{end} (having a distance of r_{crank}) and by rotating i_{grow} around the shaft keeping the bond lengths $L_{i_{from}/i_{grow}}$ and $L_{i_{grow}/i_{end}}$ fixed. Multiple (nc_n) trial sites are also used for this last step, and one trial site is selected according to the conventional CBMC rule (i.e., using eqs 1 and 2, but setting $g_i = 1$). However, there is also a Jacobian factor associated with this last crankshaft move; this factor is given by[26]

$$J = (L_{i_{from}/i_{grow}} L_{i_{grow}/i_{end}} r_{crank})^{-1} \tag{3}$$

Since the two bond lengths involved in the Jacobian are identical for the new and old conformation, only r_{crank} will contribute to the ratio of Jacobians required in the acceptance of the SAFE-CBMC move (see below). For each crankshaft attempt, two bending angles and four torsional angles can be altered.

As with any CBMC scheme, the Rosenbluth growth has to be carried out twice to generate the new (trial) conformation and to calculate a proper Rosenbluth weight for the old conformation. Once this has been accomplished, the acceptance of the SAFE-CBMC trial move can now be considered. Following Dijkstra et al.,[28] we can show that, to satisfy the detailed balance condition, the guiding probabilities corresponding to *only* the n_{step} segments of the new and old conformations (and not the contributions from the other $nc_n - 1$ trial sites for every step) have to be considered for the final acceptance probability. Let us consider a trial attempt that would change the conformation of n_{step} interior segments from state A (the old conformation) to state B (the trial conformation). The probability that the SAFE-CBMC growth results in conformation B is given by

$$P(B) = \prod_{n=1}^{n_{step}} \frac{\exp[-\beta u_{n,i}(B)] g_{n,i}(B)}{W_n(B)} \tag{4}$$

the corresponding Rosenbluth weight is

$$W(B) = \prod_{n=1}^{n_{step}} \frac{\sum_{j=1}^{nc_n} \exp[-\beta u_{n,j}(B)] g_{n,j}(B)}{nc_n} = \prod_{n=1}^{n_{step}} W_n(B) \tag{5}$$

and the potential energy of the n_{step} segments in conformation B (again, focusing solely on the nonbonded interactions considered as above) is

$$U(B) = \sum_{n=1}^{n_{step}} u_{n,i}(B) \tag{6}$$

Thus, the relevant term for the Metropolis acceptance rule using an asymmetric, biased underlying matrix of the Markov chain is[1,8]

$$\frac{P(B)}{\exp[-\beta U(B)]} = \prod_{n=1}^{n_{step}} \left[\frac{g_{n,i}(B)}{nc_n W_n(B)} \right] \tag{7}$$

Including also the Jacobian term (that arises from the last growth step for semiflexible chain models) and canceling the term $\prod_{n=1}^{n_{step}} nc_n$ (which is identical for the new and old conformations), we obtain for the SAFE-CBMC acceptance rule for the move from conformation A to conformation B

$$P_{acc}(A \rightarrow B) = \min\left[1; \left(\frac{\prod_{n=1}^{n_{step}} W_n(B)}{\prod_{n=1}^{n_{step}} W_n(A)}\right)\left(\frac{\prod_{n=1}^{n_{step}} g_{n,i}(A)}{\prod_{n=1}^{n_{step}} g_{n,i}(B)}\right)\left(\frac{J(B)}{J(A)}\right)\right] \tag{8}$$

It should be emphasized that since the contributions of the guiding probabilities are removed in the acceptance step of the SAFE-CBMC move, they act only as factors that change the underlying matrix of the Markov chain. Thus far, the guiding probabilities only influence the sampling efficiency but not the ensemble averages.

B. Determination of the Guiding Probabilities. As in the previous section, we discuss first the determination of the set of guiding probabilities that are used for the SAFE-CBMC regrowth of an interior part of a linear chain. The most appropriate guiding probability for position **r** of the trial segment i_{grow} would be the ensemble-averaged, normalized probability distribution to find i_{grow} at a specific separation from the target end segment i_{end} (or the intramolecular radial distribution function for the pair i_{grow} and i_{end}). However, this probability distribution is in all but the simplest cases not known a priori. Therefore, the following strategy is suggested here. First, an initial guess of the set of guiding probabilities is obtained either from a previous simulation of the same or a related polymer system (differing, say, in chain length, packing fraction (density), or temperature) or by performing a short presimulation for a single polymer molecule in a vacuum. This presimulation does employ only regular CBMC moves (sufficient to sample rather long polymers in a vacuum) and does not require much CPU time because it is done for only a single molecule. In our case, we use bins with a width of 0.1 Å to determine the ensemble averaged bead−bead distance distributions from one of the two options listed above. These distributions now have to be properly normalized to yield the corresponding guiding probabilities.

Second, the initial set of guiding probabilities is modified during the simulation for the system of interest. As discussed above, the SAFE-CBMC scheme satisfies the detailed balance condition irrespective of the set of guiding probabilities, but these are instrumental to improve the sampling efficiency. (For example, if one uses a uniform guiding probability for all possible distances between i_{grow} and i_{end}, the guiding probabilities in eq 8 cancel, and one reverts back to regular CBMC moves that will be inefficient in finding trial conformations that reach i_{end} but nevertheless sample from the correct ensemble distribution.) Now, the initial guess of the guiding probabilities will not be the optimal set for the system of interest because the conditions differ between presimulation (or previous simulation of a related system) and actual simulation. Thus, it will be prudent to optimize the set of guiding probabilities as more information about the system of interest becomes available. To this extent, we suggest to use a periodic self-adapting of the guiding probabilities. This self-adapting method is implemented by recording the intramolecular bead−bead distributions during the SAFE-CBMC simulation of the system of interest. After a predetermined number of cycles is reached, the set of bead−bead distributions is given a certain weight (related to the number of cycles over which it is calculated) and is used to modify the set of guiding probabilities. It should be emphasized here that the method of self-adapting the guiding probabilities is directly related to the common Monte Carlo practice of adjusting the maximum displacements for, say, translational moves. Both adjustments modify the underlying matrix of the Markov chain and can be safely used during the equilibration period of a simulation but should only be used with caution during the production period.

Let us briefly discuss some options for the set of guiding probabilities. For a homopolymer (say, with 100 segments), the simplest option would be to determine a set of guiding probabilities that only accounts for the number of bonds (links) between i_{grow} and i_{end}. Thus, we would use the same guiding probabilities for two five-bond connections associated with $i_{grow} = 1$ and $i_{end} = 6$ or $i_{grow} = 51$ and $i_{end} = 56$. In certain cases, however, the ensemble-averaged bead−bead distributions might be different for the same number of links either being close to the end of the homopolymer or being close toward its center. In this case, it will be advantageous to determine guiding probabilities that are specific for a given combination of i_{grow} and i_{end}, but this will dramatically increase the memory required for the set of guiding probabilities. However, as will be seen later, the decreasing acceptance probabilities with increasing n_{step} of SAFE-CBMC moves for dense polymer systems make it more practical to set a maximum for n_{step}; thus, the memory requirements still remain moderate for most chemical systems of interest. In the case of heteropolymers, guiding probabilities that are specific for a given combination of i_{grow} and i_{end} are always required, but simplifications are possible if the heteropolymer contains repeating bead sequences.

C. Extensions to Cyclic and Branched Molecules. The main advantage of using SAFE-CBMC moves (in addition to regular CBMC moves) for linear chain molecules is the increased sampling efficiency for interior segments, but the gains for cyclic molecules are far greater because regular CBMC cannot be used at all for molecules without free end segments. Crankshaft and CONROT moves might be an alternative for conformational changes of cyclic molecules,[27] but SAFE-CBMC also allows one to perform insertions and removals of (moderate-length) cyclic molecules as is required for simulations in the grand canonical and Gibbs ensembles.

For cyclic molecules without side branches, we can use the SAFE-CBMC algorithm as described in section 2.A without modification. However, the strategies that are used to determine the set of guiding probabilities have to be refined. There are three options to obtain an initial guess of these probabilities: (i) a presimulation using only crankshaft or CONROT moves might be used; (ii) a CBMC presimulation for a single, linear molecule, that is related to the cyclic molecules of interest, might be used; or (iii) one might simply start with uniform guiding probabilities. Any of these options has some shortcomings, such as the requirement of relatively long presimulations if only crankshaft moves are used or the fact that the guiding probabilities relevant for the connection of more than a few interior segments might be different between linear and cyclic molecules. Thus, we propose to use SAFE-CBMC of a single cyclic molecule to adapt the initial guess of the guiding probabilities obtained with any of the three options listed above, before starting the simulations of a many-polymer system.

There are more dramatic changes required for the extension of SAFE-CBMC to branched molecules and polymer networks. In these cases, a SAFE-CBMC growth might have to be targeted at multiple fixed end points. As for linear portions of a chain molecule, the first step of a SAFE-CBMC trial move is the random

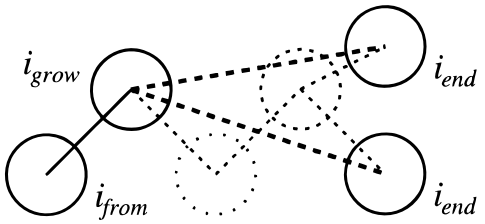

Figure 3. Growth scheme for a SAFE-CBMC reggrowth attempt that encounters a branch point leading to two fixed end points.

selection of a starting site for the regrowth, followed by the random selections of the direction of the Rosenbluth walk and of the number of growth segments n_{step}. Thereafter one can determine whether this regrowth will encounter one or more branch points in the chain. (A branch point is a segment that is directly connected to at least three other segments.) At every branch point, the SAFE-CBMC growth will continue with the regrowth of *all* remaining branches in a segment-by-segment fashion which might lead to a growth directed at multiple fixed end points (see Figure 3). For the growth of segments prior to a given branch point, the overall guiding probabilities are determined from the product of the individual guiding probabilities of the growing bead with *all* fixed end points of the growth.

If fully flexible models (including bond stretching) are used, then no further special considerations are required for branched molecules. For semiflexible models (fixed bond lengths as used here), special problems can arise because the last (n_{step}) segment of the growth might be a branch point, such as a tertiary or quaternary carbon. In this case, we cannot use a crankshaft move for the final closure. However, the situation simply becomes a geometrical problem. If the last unit needs to be connected to three segments (i_{from}, i^1_{end}, and i^2_{end}) with three bonds of given fixed lengths, there will usually be only two positions for the trial segment i_{grow} that satisfy all three bond constraints. (A special, but exceedingly unlikely, scenario arises when the three connection points lie exactly in a plane where there is only one satisfactory position for i_{grow}.) Since there are almost always two possible positions, the volume of phase space (Jacobian term) for the final closure is the same for the old and new conformation and cancels in the acceptance rate; see eq 8.

3. Models and Simulation Details

A. Force Field. All simulations described here were carried out using the transferable potentials for phase equilibria−united atom (TraPPE−UA) force field for linear and branched alkanes.[12,31] In this force field the individual atoms of methyl, methylene, and methine groups are merged into pseudoatoms. The bonded and nonbonded parameters for the TraPPE−UA force field are listed in Table 1. The TraPPE−UA force field uses a semiflexible representation of the alkanes, where all C−C bonds have the (fixed) length of 1.54 Å. A harmonic potential is used to govern bond bending[32]

$$u_{bend} = \frac{k_\theta}{2}(\theta - \theta_0)^2 \tag{9}$$

where k_θ and θ_0 are the bending force constant and the equilibrium bond angle. For the dihedral motion, the cosine series of the OPLS united-atom torsional potentials are employed[33]

$$u_{tors}(\text{C−C−C−C}) = c_0 + c_1[1 + \cos(\phi)] + c_2[1 - \cos(2\phi)] + c_3[1 + (3\phi)] \tag{10}$$

There are no additional 1−4 interactions. The intramolecular interactions of pseudoatoms separated by four or more bonds and all intermolecular interactions constitute the set of nonbonded interactions (as used in the SAFE-CBMC algorithm; see section 2.A and the Appendix). For the TraPPE−UA alkanes, only pairwise-additive Lennard-Jones 12−6 potentials (there are no electrostatic interactions of partial charges or dipoles, etc.) are used

$$u_{LJ}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{11}$$

where r_{ij}, ϵ_{ij}, and σ_{ij} are the separation, Lennard-Jones well depth, and Lennard-Jones diameter, respectively, for the pair of atoms i and j. For the determination of the Lennard-Jones parameters for unlike interactions, such as for the interaction of a methyl with a methylene pseudoatom, the standard Lorentz−Berthelot combining rules are used[34,35]

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \qquad \epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{12}$$

A spherical potential truncation is used for all bead−bead interactions, and analytical tail corrections (for internal energies, pressures, and chemical potentials) are added to account for the remainder of the interactions beyond the potential cutoff.[2]

Table 1. Bonded and Nonbonded Parameters for the TraPPE−UA Force Field

stretch	r_0 [Å]			
CH_x−CH_y	1.54			
bend	θ_0	k_θ/k_B [K]		
CH_x−(CH_2)−CH_y	114	62500		
CH_x−(CH)−CH_y	112	62500		
torsion	c_0/k_B [K]	c_1/k_B [K]	c_2/k_B [K]	c_3/k_B [K]
CH_x−(CH_2)−(CH_2)−CH_y	0	335.03	−68.19	791.32
CH_x−(CH_2)−(CH)−CH_y	−251.06	428.73	−111.85	441.27
pseudoatom	ϵ/k_B [K]	σ [Å]		
CH_3	98	3.75		
CH_2	46	3.95		
CH	10	4.68		

B. Simulations of Single Chain Molecules. Simulations using only regular coupled−decoupled CBMC moves[12] or using a combination of regular coupled−decoupled CBMC and coupled−decoupled SAFE-CBMC (selected randomly with equal probability) were carried out for different types of isolated alkane molecules. The following prototypical alkane molecules and temperatures were used: *n*-dodecane at $T = 500$, 1000, and 5000 K; 5-butylundecane at $T = 500$ K, and cyclododecane at $T = 500$ K. (Regular CBMC moves could of course not be used for this cyclic molecule.) Different options were used to obtain the initial guesses for the SAFE-CBMC guiding probabilities: (i) for *n*-dodecane and 5-butylundecane, a presimulation consisting of 10^5 regular CBMC moves was used; (ii) for cyclododecane using either a *n*-dodecane presimulation consisting of 10^5 regular CBMC moves or using a uniform distribution. In all of these cases, the total length of the simulations was 10^6 MC moves. No restriction was used

for n_{step} for both regular CBMC and SAFE-CBMC moves. The following set of choice parameters was used for the regular coupled−decoupled CBMC moves:[12] $n_{chLJ} = 20$, $n_{chtor} = 100$, and $n_{chbend} = 2000$. The corresponding parameters for the coupled−decoupled SAFE-CBMC moves were as follows: $nc_n = n_{chLJ} = 20$, $n_{chtor} = 100$, and $n_{chbend} = 2000$, for all i_{grow} that are separated by at least two beads from i_{end}; $nc_n = n_{chLJ} = 20$, $n_{chtor} = 100$, and $n_{chbend} = 2000$, for the last, but one bead; and $nc_n = n_{chLJ} = 20$ and $n_{chtor} = 200$ for the crankshaft closure of a linear segment. For closures at the ternary branch point, the two possible solutions were explored (see section 2.C).

C. Simulations in the Isobaric−Isothermal Ensemble. Isobaric−isothermal ensemble simulations[36] at $p = 101.3$ kPa were carried out for a system of 100 *n*-tetracosane molecules at $T = 420$ K and for a system of 120 cyclododecane molecules at $T = 353.15$ K. Again, two independent simulations were used for the linear molecule: one using solely regular coupled−decoupled CBMC conformational moves and the other using a combination of regular CBMC and SAFE-CBMC moves (selected with equal probabilities). The probabilities for the different types of Monte Carlo moves were 0.33, 0.33, 0.33, and 0.01 for translational, rotational, conformational, and volume moves, respectively. The maximum value of n_{step} was set to 20 for regular CBMC moves and to 10 for SAFE-CBMC. The initial guesses for the SAFE-CBMC guiding probabilities for these condensed-phase simulations were obtained (i) from a presimulation of a single *n*-tetracosane molecule consisting of 10^5 MC moves at $T = 420$ K for the *n*-tetracosane simulation and (ii) from the SAFE-CBMC simulation of an isolated cyclododecane molecule (see section 3.C). After lengthy equilibration periods, production runs of 5×10^4 MC cycles (one MC cycle consists of N MC moves where N is the number of molecules in the system) were performed for each case. The same sets of CBMC and SAFE-CBMC choice parameters were used as in the simulations of the isolated molecules (see section 3.B). To improve the computational efficiency of the condensed-phase simulations, we used an additional center-of-mass based cutoff, which avoids computing unneeded bead−bead distances,[31] and a split-energy cutoff ($r_{CBMC} = 5$ Å) for use during regular CBMC and SAFE-CBMC, which is then corrected to the full potential ($r_{cut} = 14$ Å) with tail corrections in the acceptance rule.[37]

D. Simulations in the Gibbs Ensemble. A combination of the Gibbs ensemble Monte Carlo[38−40] and SAFE-CBMC methods was used to determine the vapor−liquid coexistence properties of cyclododecane at $T = 500$ and 550 K. To swap a cyclic molecule, it has to be entirely regrown starting with bead 1 (as for linear and branched molecules, a multiple insertion method[41,42] with 20 trial sites was employed to find a favorable position for this first bead) and completely regrowing the cycle by reconnecting to bead 1. The total number of molecules was 200, and the volume of the two simulation boxes was adjusted so that the liquid phase contained approximately 160 molecules with the remainder in the vapor phase. The same set of CBMC and SAFE-CBMC choice parameters as in the simulations of the isolated cyclododecane (see section 3.B) were used, but the maximum value of n_{step} for conformational SAFE-CBMC moves was set to 4. Center-of-mass and split-energy cutoffs (see section 3.C) were employed. The equilibration and production periods consisted each of 10^4 and 5×10^4 Monte Carlo cycles, respectively, with the fractions of the different types of Monte Carlo moves being 0.3, 0.3, 0.3, 0.001, and 0.1 for translational, rotational, conformational, volume, and swap moves, respectively.

Table 2. Comparison of Average Potential Energies and Mean Square End-to-End Distances for Simulations Using Different Sampling Methods

molecule	sampling method	T [K]	$\langle U \rangle / k_B$ [K]	$\langle R_e \rangle$ [Å]
n-dodecane	CBMC	500	6133	113.16
		1000	10696	101.14
		5000	30873	86.65
	CBMC/SAFE-CBMC	500	6131	113.13
		1000	10701	101.48
		5000	30843	86.73
5-butylundecane	CBMC	500	7348	89.91
	CBMC/SAFE-CBMC	500	7355	89.93

5. Results and Discussion

A. Simulations of Single Chain Molecules. The purpose of the simulations for isolated *n*-dodecane and 5-butylundecane molecules was to confirm the correctness of our SAFE-CBMC computer code by comparing to simulations that employ only regular CBMC moves. The bond bending angle and dihedral angle distributions for both molecules (see Figures 4−7) show excellent agreement. It is important to note here that many CBMC methods developed previously fail to yield the correct conformational distributions for branched molecules[12] but that use of coupled−decoupled CBMC and coupled−decoupled SAFE-CBMC leads to the correct distributions. The corresponding average potential energies and mean-square end-to-end lengths (for 5-butylundecane, the separation of the two end segments of the undecane main chain are used) are listed in Table 2, and again agreement between regular CBMC and combined CBMC/SAFE-CBMC simulations is excellent.

The acceptance rates as a function of n_{step} for regular CBMC and SAFE-CBMC moves of *n*-dodecane and cyclodecane are compared in Figure 8. While the acceptance rates for conventional CBMC moves of an isolated linear molecule is fairly constant irrepective of the number of segments regrown, there is a strong dependence on the number of segments for the SAFE-CBMC regrowth of both the linear and cyclic molecules. An interesting feature is that the SAFE-CBMC acceptance probabilities for the linear molecule level off after a few segments. In contrast, the decrease in acceptance probabilities for the cyclic molecule remains exponential over the entire range.

To get a rough idea of the increased computational requirements for SAFE-CBMC moves versus regular CBMC moves, we compared the average times per move of a given n_{step}. Not counting the crankshaft moves ($n_{step} = 1$), it is observed that SAFE-CBMC takes approximately 30% longer. However, using the usual random selection of n_{step}, the average time needed per SAFE-CBMC move becomes actually smaller than that for regular CBMC moves which reflects the fact that on average SAFE-CBMC moves are performed for smaller n_{step} values. Finally, it is worth noting that these timing comparisons were carried out for single molecules of moderate length where the calculation of the bonded interactions dominates. Whereas the calculation of the nonbonded interaction is usually the computational

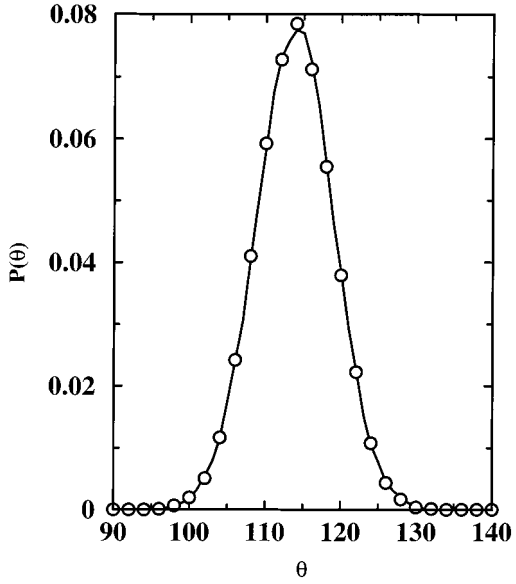

Figure 4. Comparison of the bond bending angle distributions of an isolated *n*-dodecane molecule (averaged over all 10 bending angles) obtained from simulations using solely regular CBMC moves (solid line) and using a combination of regular CBMC and SAFE-CBMC (circles) at $T = 500$ K.

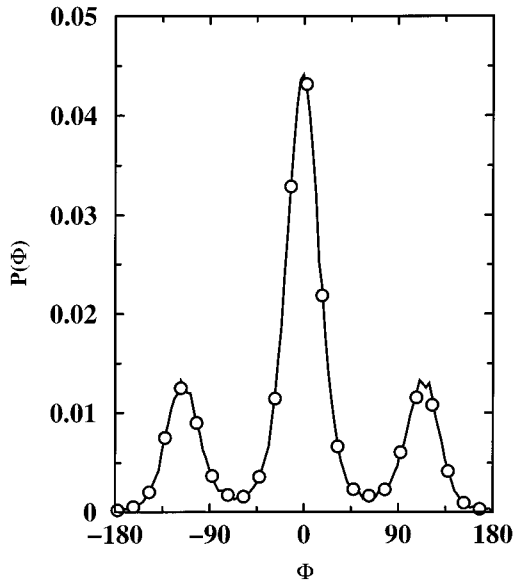

Figure 5. Comparison of the dihedral angle distributions of an isolated *n*-dodecane molecule (averaged over all nine dihedral angles) at $T = 500$ K. Line style and symbol as in Figure 4.

bottleneck for simulations of condensed-phase systems (or isolated high-molecular-weight polymers), and indeed using equal n_{step} we find little difference in the computational speeds for the simulations of the 100-molecule *n*-tetracosane system.

To test the effectiveness of the self-adapting scheme, we can compare the distributions of bead−bead separations (or guiding probabilities) that were obtained using a combined regular CBMC and SAFE-CBMC simulation started from a uniform distribution of guiding probabilities or using solely regular CBMC moves. As is shown in Figure 9, after just 10^5 moves the distribution of guiding probabilities has adapted quite well and is already a fair approximation of the final distribution. After 10^6 moves, the SAFE-CBMC guiding probabilities match the distribution of bead−bead distributions

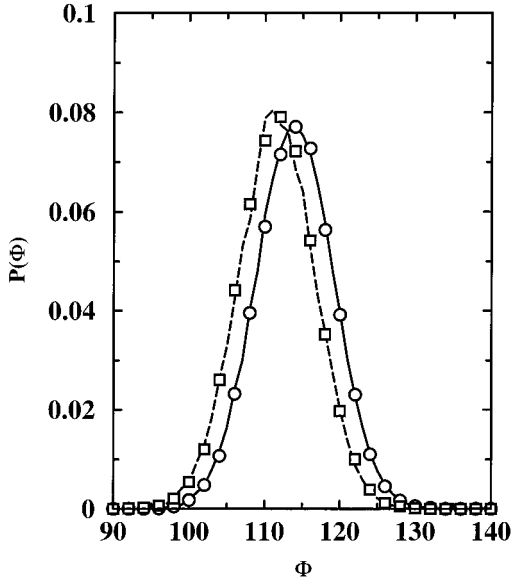

Figure 6. Comparison of the bond bending angle distributions of an isolated 5-butylundecane molecule at $T = 500$ K. The distributions obtained from simulations using solely regular CBMC moves are shown as solid line (averaged over the 11 bending angles centered at secondary carbons) and as dotted line (averaged over the three bending angles centered at the ternary carbon branch point). The corresponding distributions calculated from simulations using a combination of regular CBMC and SAFE-CBMC are depicted as circles and squares, respectively.

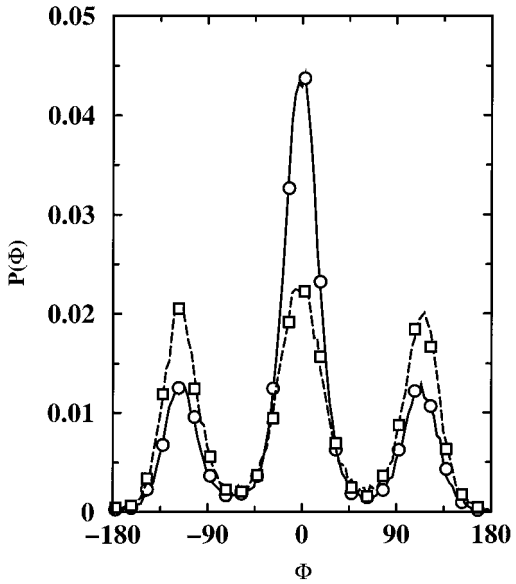

Figure 7. Comparison of the dihedral angle distributions of an isolated 5-butylundecane molecule (averaged over the eight dihedral angles that involve rotation around a bond formed by two secondary carbons and averaged over the three dihedral angles that involve rotation around a bond containing the ternary carbon branch point) at $T = 500$ K. Line styles and symbols as in Figure 6.

obtained from the regular CBMC simulation, and there is no trace of the uniform distribution from which the SAFE-CBMC simulation was started. Another encouraging feature is that the SAFE-CBMC acceptance rates (averaged for n_{step} ranging from 2 to 9) increases from 4.6% after 10^3 moves, over 9.9% after 10^4 moves, to 11.6% after 10^5 moves.

The evolutions of the guiding probabilities for the cyclododecane molecule are compared in Figures 10 and

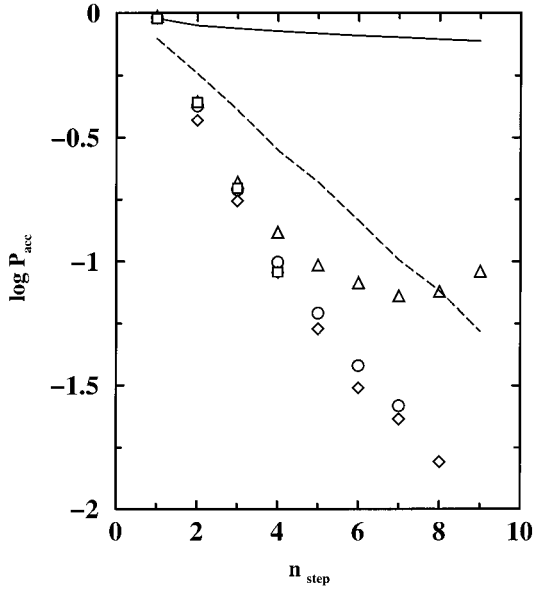

Figure 8. Comparison of acceptance rates for regular CBMC (lines) and SAFE-CBMC (symbols) moves: isolated *n*-dodecane at $T = 500$ K (solid line and triangles); *n*-tetracosane in the liquid phase at $T = 420$ K and $p = 101.3$ kPa (long-dashed line and diamonds); isolated cyclododecane at $T = 500$ K (circles); and cyclododecane in the liquid phase $T = 353$ K and $p = 101.3$ kPa (squares).

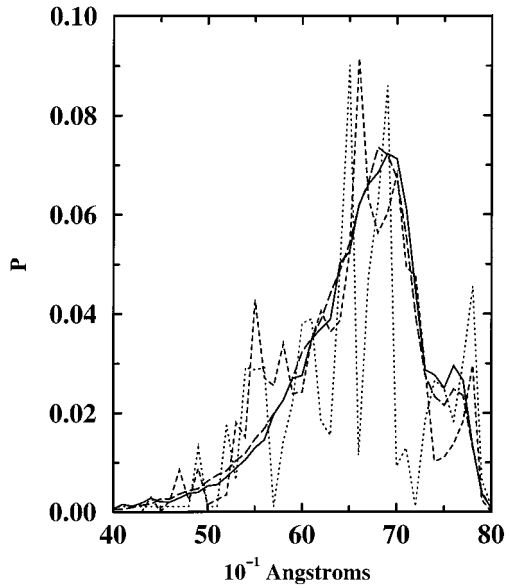

Figure 9. Guiding probabilities for the separation of beads 2 and 8 of the *n*-dodecane molecule at $T = 500$ K. Dotted, dashed, and solid lines depict the probabilities obtained using SAFE-CBMC started from a uniform distribution after 10^3, 10^4, and 10^5 moves, respectively. For comparison the outcome of a simulation using 10^6 regular CBMC is also shown as long-dashed line.

11 for SAFE-CBMC simulations starting either using the set of guiding probabilities obtained for *n*-dodecane or using a uniform distribution. In both cases, 10^5 SAFE-CBMC moves were not sufficient to converge to a satisfactory distribution. At this point the distributions are rather narrow and reflect to a large extent that of the initial cyclododecane conformation. (The distributions after 10^3 or 10^4 moves obtained form the simulation started from the set of *n*-dodecane guiding probabilties appears to be a little wider and is shifted a little bit to shorter distances.) This is quite different from the

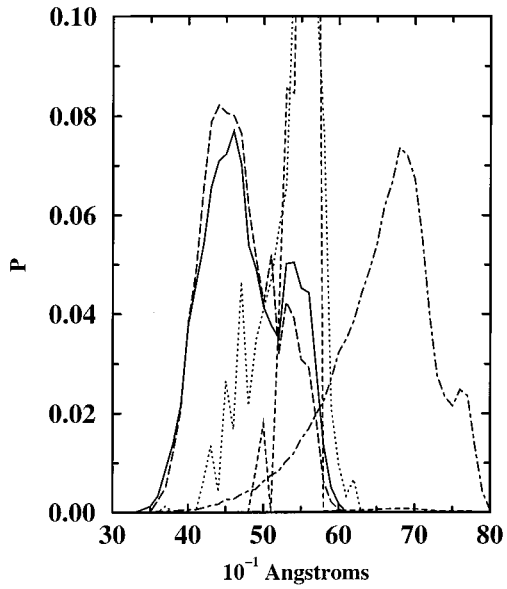

Figure 10. Guiding probabilities for the separation of beads 2 and 8 of the cyclododecane molecule at $T = 500$ K started from a *n*-dodecane presimulation (dot−dashed line) and afer self-adapting for 10^3 cycles (dashed line), 10^4 cycles (dotted line), 10^5 cycles (long-dashed line), and 2×10^5 cycles (solid line).

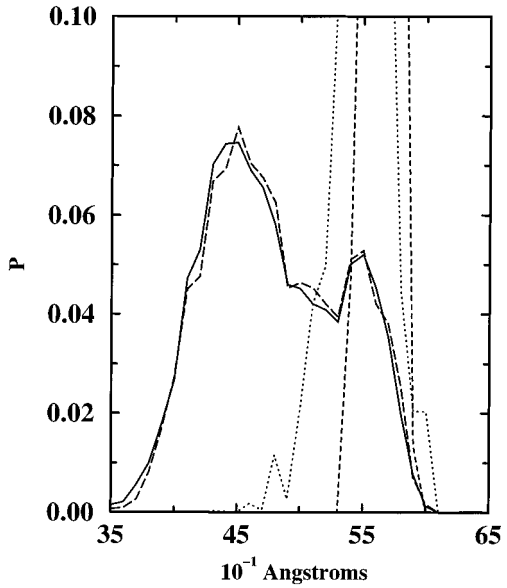

Figure 11. Guiding probabilities for the separation of beads 2 and 8 of the cyclododecane molecule at $T = 500$ K started using a uniform distribution. Line styles as in Figure 10.

n-dodecane simulations, most likely because of the lack of regular CBMC moves for cyclododecane. However, after 10^5 SAFE-CBMC moves the distribution of guiding probabilities shifted to significantly shorter distances and broadened considerably. Little change occurs thereafter. In contrast to the *n*-dodecane simulations, a significant increase of the SAFE-CBMC acceptance rate was not observed over the course of the simulation. However, it should be noted that given the similar acceptance rates, the wider distribution of guiding probabilities resulting from the self-adapting process allows a more efficient exploration of phase space.

B. Simulations in the Isobaric−Isothermal Ensemble. The simulations for the linear tetracosane molecule allow for a direct comparison of regular CBMC and SAFE-CBMC for condensed phases. *n*-Tetracosane

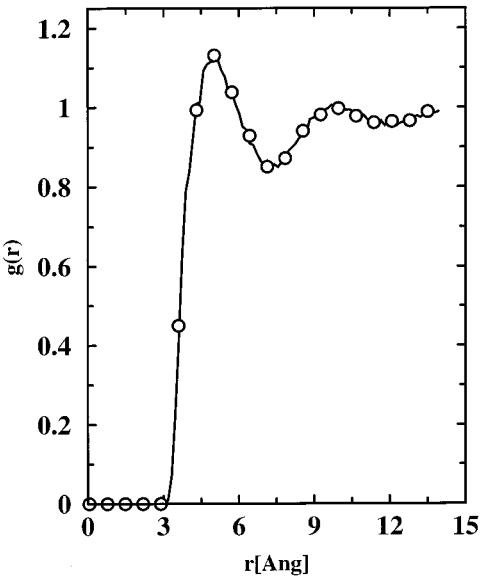

Figure 12. Intermolecular methylene−methylene radial distribution functions for liquid *n*-tetracosane at $T = 420$ K and $p = 101.3$ kPa calculated from the regular CBMC simulation (solid line) and from the combined CBMC/SAFE-CBMC simulation (circles).

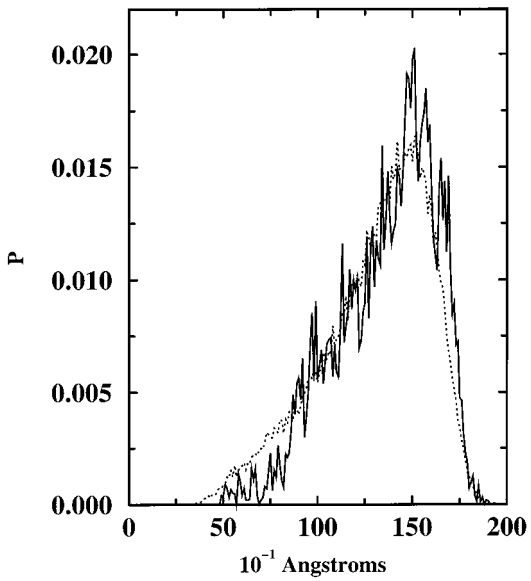

Figure 13. Guiding probabilities for the separation of beads 5 and 20 of the *n*-tetracosane molecule at $T = 420$ K. Dotted and solid lines depict the probabilities obtained from a CBMC presimulation of an isolated molecule and obtained at the end of the combined CBMC/SAFE-CBMC simulation, respectively.

is close to the maximum in chain length (molecular weight) that can be sampled reliably with regular CBMC moves at relatively high packing fractions (specific densities above 0.7 g/mL for alkanes). Regular CBMC simulations of longer chains become arduous because of the low acceptance rates for moves that regrow more than half of the chain which are required for sampling of the interior segments toward the middle of the chain.

The specific density and average internal energy calculated from the regular CBMC simulations are 0.727 ± 0.003 g/mL and 0.80 ± 0.05 kJ/mol; the corresponding values for the combined CBMC/SAFE-CBMC simulations are 0.728 ± 0.002 g/mL and 0.81 ± 0.06 kJ/mol. While there is excellent agreement between the two simulations, it should be mentioned that the experimental specific density of 0.719 g/mL differs by a little over 1% and lies outside the statistical uncertainties of the simulations. This deviation can be attributed to shortcomings of the simple united-atom force field. Radial distribution functions for methylene−methylene pairs were also calculated (see Figure 12), and agreement between the two simulations is excellent. The isobaric−isothermal simulation for cyclododecane at $T = 353.15$ K resulted in a specific density of 0.811 ± 0.004 g/mL, slightly lower than the corresponding experimental value of 0.820 g/mL.[44]

The distributions of guiding probabilities for the connection of beads 5 and 20 of *n*-tetracosane obtained from a CBMC presimulation of an isolated molecule and at the end of the combined CBMC/SAFE-CBMC of the bulk liquid are shown in Figure 13. While the distributions are similar, it is nevertheless clearly evident that the self-adapting procedure yielded an optimized set of guiding probabilities for the liquid system. Since its own liquid is an athermal solvent for *n*-tetracosane, the linear chains in the liquid are more extended than for the isolated molecule which collapses on to itself. The corresponding mean-square end-to-end lengths are 376 and 290 Å^2 for the liquid-phase and isolated molecules, respectively. The acceptance rates for CBMC and SAFE-CBMC moves are given in Figure 8. While the acceptance probabilities using conventional CBMC deteriorate dramatically when changing from an isolated molecule to one in a liquid environment, there is little difference between the isolated and the bulk-phase for the SAFE-CBMC regrowths of up to five segments for a linear molecule. Actually for cyclododecane, the acceptance rates for the isolated and liquid-phase environments are very similar over the entire range of regrowths.

C. Simulations in the Gibbs Ensemble. The purpose of the GEMC simulations was to test whether the SAFE-CBMC method can be used to facilitate the swap of the cyclic molecules between the two simulation boxes. In previous GEMC simulation studies for small cyclic alkanes,[27] the molecules were held rigid during the swap attempts, thereby greatly limiting the size of the cyclic alkanes that could be studied. Simulations for cyclododecane were carried out for two temperatures (500 and 550 K) in the proximity of the normal boiling temperature of 520 K determined experimentally.[44] The acceptance rates for swap moves at these temperatures were 2.5×10^{-4} and 8×10^{-4}, respectively. The SAFE-CBMC acceptance rates for conformational moves of cyclodocane in the GEMC simulations are rather satisfactory and similar to those values found in the isobaric−isothermal simulations. Finally, while no experimental data could be found for the orthobaric densities of cyclododecane (the results of the simulations are $\rho_{liq} = 0.705$ g/mL and $\rho_{vap} = 0.0088$ g/mL at $T = 500$ K and $\rho_{liq} = 0.655$ g/mL and $\rho_{vap} = 0.0116$ g/mL at $T = 550$ K), we can compare to the Antoine prediction of its vapor pressure.[44] The TraPPE-UA model for cyclodedecane yields vapor pressures that are substantially too high and the calculated normal boiling temperature of 470 K is too low by 50 K.

6. Conclusions and Future Work

The SAFE-CBMC method for molecules with strong intramolecular interactions greatly extends the range

of molecular systems that can be successfully simulated. The correctness of the SAFE-CBMC method has been confirmed for simulations of linear and branched molecules, and it is demonstrated that the SAFE-CBMC method also allows for efficient conformational moves and particle insertion/removal moves for cyclic molecules. For condensed phases, a rapid decrease in the SAFE-CBMC acceptance rates with increasing number of segments involved in the move is observed; thus, it is prudent to specify an upper limit for the number of segments for which a SAFE-CBMC regrowth is attempted. The main advantage of SAFE-CBMC method over the related ECCB[26] and FE-CBMC[28,29] approaches is that it can be readily applied to any type of molecular-mechanics force field. Its main advantages over CONROT schemes are the general applicability to branched chains and networks and the ability to perform efficient insertions/removals of cyclic molecules as required for simulations in open ensembles. It should be emphasized here, however, that we have not done a direct comparison of the computational efficiencies of the CONROT and SAFE-CBMC algorithms (see below).

Future work is planned to calculate the coexistence curves and critical properties of high-molecular-weight cyclic alkanes using SAFE-CBMC in the Gibbs and grand canonical ensembles.[45] A useful extension will be the SAFE-CBMC switch move[46] which allows for the identity switch of two related molecules, such as *n*-hexadecane and 8-methylpentadecane, by transferring a few interior segments from one molecule to the other. Finally, it will be worthwhile to explore a combination of look-ahead CBMC[47] with SAFE-CBMC and to use SAFE-CBMC for end-bridging moves[22] in polymeric systems.

While this article was pending review, a paper describing a very much related algorithm, called internal configuration bias Monte Carlo (ICBMC), was published.[48] Uhlherr's ICBMC algorithm also allows for the sampling of an arbitrary number of interior segments using a combination of a conventional CBMC growth and a CONROT closure for the last three segments.[48] A set of finite extendable nonlinear elastic (FENE) springs is used for the guiding probabilities of the CBMC growth part. Uhlherr applied the ICBMC algorithm to the simulation of isolated linear and cyclic alkanes with 16−100 united atoms (connected by bonds with fixed length and bending angle) and compared extensively to CONROT Monte Carlo and conventional molecular dynamics simulations. His conclusions are that ICBMC is computationally more efficient than CONROT for medium-size cyclic alkanes where the restricted conformational space requires cooperative transitions of multiple dihedral angles, whereas CONROT was found to be more efficient for larger, less constrained rings.[48] In addition, Uhlherr reported that ICBMC is about 2 orders of magnitude more efficient than a molecular dynamics simulation for the cyclic alkanes. A direct comparison of ICBMC and SAFE-CBMC is not straightforward because different molecular models were used for the alkanes. One might infer that self-adapted guiding probabilities should improve the efficiency over generalized FENE springs, particularly for cases such as constrained rings, where the distributions of bead−bead distances can exhibit multiple maxima (see Figure 11). To confirm this, an additional SAFE-CBMC calculation (using the standard set of coupled−decoupled CBMC parameters) was carried out for an isolated cyclotriacontane molecule at $T = 300$ K. The fractions of accepted SAFE-CBMC moves (and completed SAFE-CBMC regrowth) for 10- and 20-segment rearrangements were 4.4×10^{-3} (0.62) and 1.4×10^{-3} (0.35), respectively, while Uhlherr found lower acceptance ratios for ICBMC moves on the same system (6×10^{-4} and 3×10^{-5}, respectively).[48]

Acknowledgment. We thank Don Truhlar and Marcus Martin for stimulating discussions and the reviewers for helpful comments and for alerting us to ref 48. Financial support from the National Science Foundation (CTS-9813601 and CHE-9816328), a McKnight/Land-Grant Fellowship, and an Alfred P. Sloan Research Fellowship is gratefully acknowledged. Part of the computer resources were provided by the Minnesota Supercomputing Institute.

Appendix: Details of the Coupled−Decoupled SAFE-CBMC Growth

Coupled−decoupled CBMC regrowths for linear and branched molecules with strong intramolecular interactions are generated by coupling the Lennard-Jones and torsional selections and using decoupled bond bending selections.[12] SAFE-CBMC follows the same route, except for the crankshaft closure, in which torsional and bending selections cannot be carried out separately, and are both coupled to the Lennard-Jones selections. Since the calculations of the SAFE-CBMC guilding probabilities, g_n, is very fast, these probabilities are included in the torsional selections (whereas Uhlherr's ICBMC scheme places them in the Lennard-Jones part). Remember that the guiding probabilities for a regrowth that passes through a branch point are the products of the individual guiding probabilities to *all* the end points that the segment being regrown at present can reconnect with.

The procedure for a coupled−decoupled SAFE-CBMC regrowth is as follows. First, select a molecule (of the desired type for mixtures) at random. Second, select a site q on this molecule at random which is used as the starting site for the regrowth (if this is an insertion/removal move, then the site q is selected from multiple first bead insertions).[41,42] Third, the directions of the Rosenbluth walk is determined by picking a "previous" site p (connected by one bond to q) at random and fixing its position. In some special cases, such as when the first end point is an end segment of the chain or during insertion/removal moves, there is no "previous" site for the first step, and one of the sites to be regrown is designated as p site (see below). Fourth, the maximum number of segments between site q and any end point of the regrowth is selected. Thereafter, the architecture of the chain is searched following all branches encountered to determine all sites n_{step} that need to be regrown and all end points. If this is an insertion/removal move or if q is an end segment, then all sites on the molecule are regrown during the coupled−decoupled SAFE-CBMC process.

The probability for generating a configuration that involves the regrowth of n_{step} interior segments is given in eqs A1−A6.

$$P_{\text{gen}} = \left\{ \prod_{n=1}^{n_{\text{step}}-n_{\text{cr}}} \left[\frac{\exp(-\beta u_{n,i}^{\text{LJ}}) W_{n,i}^{\text{T}}}{W_n^{\text{L}}} \right] \left[\frac{\exp(-\beta u_{n,i,j}^{\text{tors}}) g_{n,i,j}}{W_{n,i}^{\text{T}}} \right] B_n \right\} P_{\text{cr}} \tag{A1}$$

$$B_n = \left\{\prod_{a=st(n)}^{n_{\mathrm{grow}(n)}}\left[\frac{\exp(-\beta u_{n,k}^{\mathrm{bend}[a]})}{W_n^{P[a]}}\right]\right\}\left\{\prod_{b=st(n)+1}^{n_{\mathrm{grow}(n)}}\left[\frac{\exp(-\beta u_{n,k}^{\mathrm{bend}[b]})}{W_n^{P[b]}}\right]\right\} \tag{A2}$$

$$W_n^{\mathrm{L}} = \sum_{i=1}^{n_{\mathrm{chLJ}}} \exp(-\beta u_{n,i}^{\mathrm{LJ}}) W_{n,i}^{\mathrm{T}} \tag{A3}$$

$$W_{n,i}^{\mathrm{T}} = \sum_{j=1}^{n_{\mathrm{chtor}}} \exp(-\beta u_{n,i,j}^{\mathrm{tors}}) g_{n,i,j} \tag{A4}$$

$$W_n^{P[x]} = \sum_{k=1}^{n_{\mathrm{chbend}}} \exp(-\beta u_{n,k}^{\mathrm{bend}[x]}) \tag{A5}$$

$$W_n^{\mathrm{B}} = \left(\prod_{a=st(n)}^{n_{\mathrm{grow}}(n)} W_n^{P[a]}\right)\left(\prod_{b=st(n)+1}^{n_{\mathrm{grow}}(n)} W_n^{P[b]}\right) \tag{A6}$$

where n, n_{step}, and n_{cr} are the running index for the bead presently grown, the total number of segments to be regrown, and the number of segments that are regrown by the crankshaft procedure. P_{cr} is the corresponding generating probability for the crankshaft part. There are three different contributions to the internal energy of a bead: u^{bend}, u^{tors}, and u^{LJ} (see eqs 9−11). The running indices i, j, and k are used for the Lennard-Jones, torsional, and bending selections. For each regrowth step (n), $n_{\mathrm{grow}}(n)$ is the number of sites that need to regrown simultaneously, that is, all segments not yet present and connected to the site being regrown from.

If this is the first regrowth step ($n = 1$) and we are regrowing all of the neighboring sites (that is one of the special regrowth of the entire molecule mentioned above), then one of those sites is selected at random, designated as the "previous" site, and $st(n) = 2$. In this case, n_{chLJ} orientations of the bond vector between the present and "previous" sites are selected randomly from a unit sphere. Otherwise, there must be a neighboring site $i_{\mathrm{from}-1}$ that is already present and $st(n) = 1$. The bond bending angle selections are performed in two stages: [a] for all angles involving sites $i_{\mathrm{from}-1}$, i_{from}, and one of the regrowth sites, and [b] for all angles centered at i_{from} and involving two of the regrowth sites.[12]

Once all the bond bending angles are determined, the coupled selection involving the nonbonded (LJ) and torsional terms is carried out. For each of the n_{chLJ} selections, n_{chtor} random orientations on a cone with axis $i_{\mathrm{from}} - i_{\mathrm{from}-1}$ are chosen uniformly on $(0;2\pi)$ for the bond vector from i_{from} to site $st(n)$. The energy u^{tors} is calculated from the sum of *all* relevant dihedral angles which are those that contain i_{from} and $i_{\mathrm{from}-1}$ as the central bond, except in the special case where the previous site $i_{\mathrm{from}-1}$ is one of the sites that is also regrown in this step. One of the n_{chtor} torsional choices is selected in the usual biased CBMC fashion, that is, with probabilities determined by u^{tors}. The second step of the coupled selection is to compute the nonbonded energy for each of the n_{chLJ} orientations (summing over all n_{grow} segments that are grown at this step). One of these orientations is then selected in a coupled biased fashion based upon the Boltzmann weight of u^{LJ} and the torsional Rosenbluth weights W^{T}.

For the special cases, where the "previous" site is one of the sites that is grown at this step, then there are no torsional angles that need to be considered. In that case, a decoupled biased selection is made for n_{chLJ} random orientations (taking i_{from} as the origin) of the rigid structure involving all n_{grow} segments (with the bond angles from the decoupled first step).

The specific values of the coupled−decoupled CBMC parameters (n_{chLJ}, n_{chtor}, and n_{chbend}) used for the simulations described in this paper are listed in section 3.

For the final SAFE-CBMC closure steps, crankshaft moves are generated with a probability as follows

$$P_{\mathrm{cr}} = \prod_{n=n_{\mathrm{cr}}}^{n_{\mathrm{step}}}\left\{\frac{\exp(-\beta u_{n,i}^{\mathrm{LJ}}) W_{n,i}^{\mathrm{TC}}}{W_n^{\mathrm{LC}}}\right\} \times \left\{\frac{\exp(-\beta[(\sum_{a=1}^{n_{\mathrm{tors}}} u_{n,i,j}^{\mathrm{tors}[a]}) + (\sum_{b=1}^{n_{\mathrm{bend}}} u_{n,i,j}^{\mathrm{bend}[b]})])}{W_{n,i}^{\mathrm{TC}}}\right\} \tag{A7}$$

$$W_n^{\mathrm{LC}} = \sum_{i=1}^{n_{\mathrm{chLJ}}} \exp(-\beta u_{n,i}^{\mathrm{LJ}}) W_{n,i}^{\mathrm{TC}} \tag{A8}$$

$$W_{n,i}^{\mathrm{TC}} = \sum_{j=1}^{n_{\mathrm{chtor}}} \{\exp(-\beta[(\sum_{a=1}^{n_{\mathrm{tors}}} u_{n,i,j}^{\mathrm{tors}[a]}) + (\sum_{b=1}^{n_{\mathrm{bend}}} u_{n,i,j}^{\mathrm{bend}[b]})])\} \tag{A9}$$

The torsional selections are made along a line connecting i_{from} and i_{end} (as shown in Figure 3) from a cone chosen uniformly on $(0; 2\pi)$. For each torsional selection, up to two bending angles (n_{bend}) and four torsional angles (n_{tor}) need to be calculated for a linear chain part.

After the coupled−decoupled regrowth of the trial conformation is successfully completed, the process is repeated tracing the old configuration. The acceptance probability for a change from configuration A to B is

$$P_{\mathrm{acc}}(A \rightarrow B) = \min\left[1; \left(\frac{\prod_{n=1}^{n_{\mathrm{step}}-n_{\mathrm{cr}}} W_n^{\mathrm{L}}(B)\, W_n^{\mathrm{B}}(B)}{\prod_{n=1}^{n_{\mathrm{step}}-n_{\mathrm{cr}}} W_n^{\mathrm{L}}(A)\, W_n^{\mathrm{B}}(A)}\right) \times \left(\frac{\prod_{n=1}^{n_{\mathrm{step}}-n_{\mathrm{cr}}} g_{n,i}(A)}{\prod_{n=1}^{n_{\mathrm{step}}-n_{\mathrm{cr}}} g_{n,i}(B)}\right)\left(\frac{\prod_{n=n_{\mathrm{cr}}}^{n_{\mathrm{step}}} W_n^{\mathrm{LC}}(B)\, J(B)}{\prod_{n=n_{\mathrm{cr}}}^{n_{\mathrm{step}}} W_n^{\mathrm{LC}}(A)\, J(A)}\right)\right] \tag{A10}$$

where $J(A)$ and $J(B)$ are the products of the Jacobian factors for *all* crankshaft moves (see eq 3) that are used during the complete SAFE-CBMC regrowth procedure.

References and Notes

(1) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem Phys.* **1953**, *21*, 1087.
(2) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.
(3) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: New York, 1996.

(4) Ferguson, D. M.; Siepmann, J. I.; Truhlar, D. G. *Monte Carlo Methods in Chemical Physics*; Adv. Chem. Phys. Vol. 105; Wiley: New York, 1999.
(5) Kremer, K.; Binder, K. *Comput. Phys. Rep.* **1988**, *7*, 259.
(6) Leontidis, E.; de Pablo, J. J.; Laso, M.; Suter, U. W. In *Advances in Polymer Science*; Springer-Verlag: New York, 1994; Vol. 116, p 185.
(7) Siepmann, J. I. *Mol. Phys.* **1990**, *70*, 1145.
(8) Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59.
(9) Frenkel, D.; Mooij, G. C. A. M.; Smit, B. *J. Phys.: Condens. Matter* **1992**, *4*, 3053.
(10) de Pablo J. J.; Laso, M.; Suter, U. W. *J. Chem. Phys.* **1992**, *96*, 2695.
(11) Macedonia, M. D.; Maginn, E. J. *Mol. Phys.* **1999**, *96*, 1375.
(12) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 4508.
(13) Rosenbluth, M. N.; Rosenbluth, A. W. *J. Chem. Phys.* **1955**, *23*, 356.
(14) Verdier, P. H.; Stockmayer, W. H. *J. Chem. Phys.* **1962**, *36*, 227.
(15) Hilhorst, H. J.; Deutch, J. M. *J. Chem. Phys.* **1975**, *63*, 5153.
(16) Kumar, S. K.; Vacatello, M.; Yoon, D. Y. *J. Chem. Phys.* **1988**, *89*, 5206.
(17) Li, X.; Chiew, Y. C. *J. Chem. Phys.* **1994**, *101*, 2522.
(18) Carmesin, I.; Kremer, K. *Macromolecules* **1988**, *21*, 2819.
(19) Shaffer, J. S. *J. Chem. Phys.* **1994**, *101*, 4205.
(20) Dodd, L. R.; Boone, T.; Theodorou, D. N. *Mol. Phys.* **1993**, *78*, 961.
(21) Pant, P. V. K.; Theodorou, D. N. *Macromolecules* **1995**, *28*, 7224.
(22) Mavrantzas, V. G.; Boone, T. D.; Zervopoulou, E.; Theodorou, D. N. *Macromolecules* **1999**, *32*, 5072.
(23) Deem, M. W.; Bader, J. S. *Mol. Phys.* **1996**, *87*, 1245.
(24) Wu, M. H. G.; Deem, M. W. *Mol. Phys.* **1999**, *97*, 559.
(25) Wu, M. H. G.; Deem, M. W. *J. Chem. Phys.* **1999**, *111*, 6625.
(26) Escobedo, F. A.; de Pablo, J. J. *J. Chem. Phys.* **1995**, *102*, 2636.
(27) Neubauer, B.; Boutin, A.; Tavitian, B.; Fuchs, A. H. *Mol. Phys.* **1999**, *97*, 769.
(28) Dijkstra, M.; Frenkel, D.; Hansen, J. *J. Chem. Phys.* **1994**, *101*, 3179.
(29) Vendruscolo, M. *J. Chem. Phys.* **1997**, 106, 2970.
(30) Mooij, G. C. A. M.; Frenkel, D. J. *Chem. Phys.* **1992**, *97*, 5113.
(31) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569.
(32) van der Ploeg, P.; Berendsen, H. J. C. *J. Chem. Phys.* **1982**, *94*, 3271.
(33) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 813.
(34) Lorentz, H. A. *Ann. Phys.* **1881**, *12*, 127.
(35) Berthelot, D. *C. R. Hebd. Seances Acad. Sci., Paris* **1898**, *126*, 1703.
(36) McDonald, I. R. *Mol. Phys.* **1972**, *23*, 41.
(37) Vlugt, T. J. H.; Martin, M. G.; Smit, B.; Siepmann, J. I.; Krishna, R. *Mol. Phys.* **1998**, *94*, 727.
(38) Panagiotopoulos, A. Z. *Mol. Phys.* **1987**, *61*, 813.
(39) Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. J. *Mol. Phys.* **1988**, *63*, 527.
(40) Smit, B.; de Smedt, P.; Frenkel, D. *Mol. Phys.* **1989**, *68*, 931.
(41) Esselink, K.; Loyens, L. D. J. C.; Smit, B. *Phys. Rev. E* **1995**, *51*, 1560.
(42) Mackie, A. D.; Tavitian, B.; Boutin, A.; Fuchs, A. H. *Mol. Simul.* **1997**, *19*, 1.
(43) Errington, J. R.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **1999**, *103*, 6314.
(44) Stephenson, R. M.; Stanislaw, M. *Handbook of the Thermodynamics of Organic Compounds*; Elsevier: New York, 1987; p 373.
(45) Wick, C. D.; Potoff, J. J.; Siepmann, J. I., unpublished results.
(46) Martin, M. G.; Siepmann, J. I. *J. Am. Chem. Soc.* **1997**, *119*, 8921.
(47) Consta, S.; Vlugt, T. J. H.; Hoeth, J. W.; Smit, B.; Frenkel, D. *Mol. Phys.* **1999**, *97*, 1243.
(48) Uhlherr, A. *Macromolecules* **2000**, *33*, 1351.

MA000172G