Fast-Growth Thermodynamic Integration: Calculating Excess Chemical Potentials of Additive Molecules in Polymer Microstructures

Berk Hess,* Christine Peter, Tugba Ozal, and Nico F. A. van der Vegt

Max-Planck Institute for Polymer Research, Ackermannweg 10, D-55128 Mainz, Germany

Received September 14, 2007; Revised Manuscript Received December 12, 2007

ABSTRACT: We have calculated excess chemical potentials of additive molecules in a bisphenol A-polycarbonate matrix using fast-growth thermodynamic integration. It is shown that this method, which is based on Jarzynski's nonequilibrium work theorem, is ideally suited to overcome some of the typical sampling problems one meets when inserting large additive molecules into a dense polymer matrix. The method provides a direct link between the calculated excess chemical potential and the atomic-scale environment preferred by the inserted molecule. We discuss for which types of free-energy landscapes fast-growth thermodynamic integration is well suited and which are the optimal parameters to use.

I. Introduction

Excess chemical potentials of additive molecules in dense, atomistically detailed polymer microstructures can in principle be calculated with statistical mechanics based methods. In particular, the Widom test-particle-insertion method has frequently been used to calculate excess chemical potentials of small, gaseous molecules such as argon, methane, and carbon dioxide.[1−9] The Widom method is particularly well suited for these gases because any amorphous polymer matrix contains pre-existing, empty cavities large enough to accommodate them. Configurational bias Monte Carlo (CBMC) simulations have proven to be very efficient to extending the range of particle sizes to molecules with flexible internal (torsion) degrees of freedom such as linear alkanes in flexible polyolefin or siloxane-based polymers.[10−12] These computational approaches are however less suitable for bulky and internally rigid additive molecules (e.g., chloroform, phenol) in high-performance polymers such as polycarbonates, polyimides, etc. Because large penetrant molecules occupy large cavities that statistically occur only rarely in the penetrant-free system, computational methods using particle insertions will not yield converged values of the excess chemical potential (convergence can be improved by using excluded volume map sampling).[6,13] Thermodynamic integration (TI) methods may alternatively be used (for applications to polymer-penetrant systems, see refs 14−16). In TI, a coupling parameter slowly introduces the attractive- and repulsive interactions between the penetrant and the polymer matrix thereby "forcing" the formation of an appropriate cavity as well as the required local relaxations of polymer chains. The corresponding thermodynamic coupling work (discussed in more detail in the next section) provides the penetrant's excess chemical potential provided that the coupling process has been performed sufficiently slow (reversible) or the system has been fully equilibrated at each discretely chosen value of the coupling parameter connecting the fully coupled and fully uncoupled states. In practice, however, this can never be achieved. The major problem to overcome is related to the sampling of the fully coupled end state. Because the free energy landscape of the fully coupled penetrant typically contains many minima of varying depth separated by barriers significantly larger than the thermal energy (k_BT), typically only few minima are sampled leading to poor estimates of the excess chemical potential. A way to overcome this problem is to perform many independent TI calculations, which however would quickly make the calculation by far too expensive. The goal of the present paper is to examine the recently developed fast-growth thermodynamic integration method,[17] which we apply to couple a single penetrant molecule to a realistic example of a stiff-chain, all-atom polymer microstructure simulated by molecular dynamics. The method requires several independent fast-growth TI runs based on which a nonequilibrium work distribution is accumulated which by virtue of Jarzynski's nonequilibrium work theorem[18] allows one to obtain the excess chemical potential. We illustrate that this method is ideally suited to sample the full end-state ensemble of complex polymer−penetrant systems (see Figure 1). We moreover discuss how to minimize errors by optimizing the balance between the number of fast-growth TI runs and the time of a single run. On the basis of this we illustrate that excess chemical potentials of propane, chloroform, and dimethyl sulfoxide (as model penetrants) in a bisphenol A-polycarbonate matrix can be calculated within an error bar of 1 kJ/mol.

II. Theoretical Background

There are two major standard approaches to compute free energy differences between two states A and B.[1] One is the thermodynamic integration (TI) method,[19] which relies on a nonphysical path and a coupling parameter λ through which the two states are connected. Standard TI assumes that one obtains full equilibration along this λ-path which allows the calculation of the free energy difference by integrating the derivative of the Hamiltonian with respect to λ.

$$\Delta F_{AB} = F(\lambda_B = 1) - F(\lambda_A = 0) = \int_0^1 \left\langle \frac{\partial H_\lambda}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (1)$$

The other most prominent method to compute free-energy differences relies on thermodynamic perturbation theory originally developed by Zwanzig,[20] which states that free-energy differences can be determined by exponentially/Boltzmann reweighting (with the energy difference between state A and B) of the conformations sampled only at state A in order to obtain the properly weighted energy distribution at state B.

$$\Delta F_{AB} = -k_BT \log\langle e^{-\beta \Delta H_{AB}} \rangle_A \quad (2)$$

* Corresponding author.

10.1021/ma702070n CCC: $40.75
Published on Web 02/29/2008

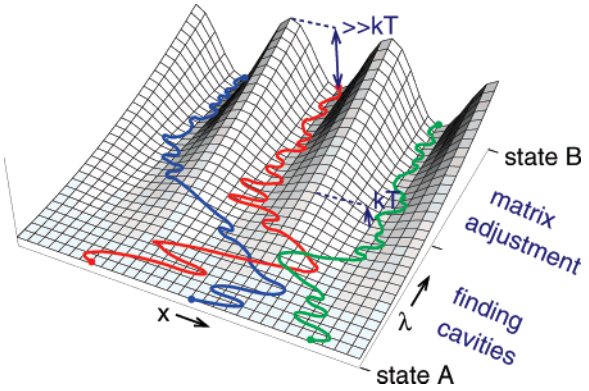

Figure 1. A schematic one-dimensional representation of a free-energy landscape for an additive molecule (x is a component of its center of mass) in a polymer matrix as a function of the coupling parameter λ. During a fast-growth simulation, a molecule is forced along λ from state A to B, while sampling in x; three example paths are given. With fast-growth TI, it is possible to explore the multiple minima in the free energy landscape by multiple trajectories starting from a flat free-energy landscape (a solute not interacting with the matrix).

Related to these two basic approaches, there are a huge variety of methods to calculate free energy differences[21−25] and potentials of mean force,[26−29] for recent reviews on this see refs 30 and 31.

In this article, we discuss the rather recent approach to determine free-energy differences based on the nonequilibrium work theorem introduced by Jarzynski. Jarzynski relates the work W_{AB} that is being performed on a system when going from state A to state B (along a coordinate λ) with the free energy change, irrespective of the reversibility of the sampling. He finds that the exponentially averaged work of repeated sampling of the path from A to B is equal to the free energy difference between A and B[18]

$$\Delta F_{AB} = -k_B T \log \langle e^{-\beta W_{AB}} \rangle_A \qquad (3)$$

where the angular bracket denotes averaging over an canonical ensemble of the initial state A. Proofs of this equation and the conditions under which it is valid have been given by many authors and will not be discussed here.[17,18,32−36]

The impact of the nonequilibrium work theorem has been enormous both for experiment and simulation. Jarzynski's equation has been used to analyze single molecule experiments,[17,37] and in the field of molecular simulation the equation has triggered a wave of developments to compute free energies. On the basis of eq 3, the fast-growth TI method was developed,[38,39] which is a direct application of the irreversible work identity. Instead of one very long TI of the system, one conducts many very fast ones and exponentially averages the work performed on the system.

The Jarzynski formalism and the resulting fast-growth methods bridge the gap between TI and perturbation methods, since eqs 1 and 2 can be viewed as the infinitely slow and the infinitely fast limits of eq 3. TI relies on exhaustively sampling phase space at each intermediate λ point to meet the requirement of being always in equilibrium and fully reversible, a requirement which is difficult to meet in complex systems. In free energy perturbation, the sampling problem is a different one, and it lies in the condition that the distributions of starting and end states have to overlap. If this overlap is insufficient, the B state is not well represented by the ensemble generated by sampling the A state, which leads to a bad sampling of low energies in the B state and therefore a prediction of the free energy difference which is too high. In the case of fast-growth methods, one has to fully sample the work distribution functions, ergo all the (relevant) paths between the two states A and B. This sampling problem is in spirit similar to the typical perturbation sampling problems: in order to replace one very long equilibrium experiment to determine a free energy difference with many fast nonequilibrium experiments, the number of fast experiments required to sample the (exponentially rarely occurring) tails of the work distribution function grows exponentially with the amount of *irreversible* work performed during the experiment.[40] This means it gets very difficult to sample the relevant (favorable) paths between A and B.

For more complex systems, two reasons for missing out on very favorable paths between A and B can be identified. The first problem occurs if the underlying free energy landscape is very complex, with potentially different paths between A and B, with different transition states. In this case it is difficult to decide whether the relevant paths (i.e., the relevant transition states) were sampled at all. The folding of biomolecules is particularly problematic, since it is notoriously difficult to identify "good" transition paths/reaction coordinates. This problem is related to the condition that one needs to "fully" sample state A, namely, the relevant regions that then lead to the various transition paths. The second reason for missing out on relevant paths is that the time scale of the experiment or simulation is faster than the relevant relaxation times of the system of interest. If the transition between states A and B is too fast so that the system is not allowed to relax, the irreversible work conducted on the system is too high and more favorable paths are not sampled. An example of this phenomenon is given below.

Many people have investigated and compared extensively the various methods to compute free energies, frequently on model systems such as an ideal gas and a piston, the growing and charging of particles, the growing and annihilation of molecules which is common in biomolecular and pharmaceutical simulations, and many more.[40−44] In general, fast growth gives good results with a reasonable effort but cannot be identified as clearly superior to other methods. On the other hand, the method has been appealing for biomolecular systems where a potential of mean force, i.e., the free energy along a certain reaction coordinate, is to be determined.[45] Here the Jarzynski identity has allowed the development of fast pulling or steered MD techniques, but as discussed, the quality of the chosen reaction coordinate can be an important limitation. Bastug and Kuyucak[46] show one problem that occurs for fast pulling simulations in complex biological systems, namely, that the long relaxation times in these kinds of systems (attributed by the authors largely to long range interactions) prevents the fast pulling trajectories from sampling the relevant (low energy) paths that contribute most significantly to the resulting free energies. The example discussed in ref 46 nicely illustrates how differently fast growth performs on model systems (a smooth nanotube as a membrane channel) compared to real biological membrane channels, which (with enough time provided) adapt conformationally to the molecule/ion that is being pulled through.

III. Methods

The polymer matrix consists of 50 chains of bisphenol A-polycarbonate (BPA-PC), each five monomers long; the force field is described in ref 47. We started from a melt that was equilibrated at 570 K for 150 ns, in which the chains diffused over 9 radii of gyration.[48] This melt was quenched and equilibrated at 480 K, leading to a cubic unit-cell of 4.8 nm. The equilibration was performed with twin-range 0.9/1.4 nm reaction-field electrostatics for reasons not relevant for this work. For the fast growth we

Table 1. The Length Per Fast-Growth Run for the LJ and Charge (q) Interactions, Number of Fast-Growth Runs (N), and Results (in kJ mol^{-1}) of the Fast-Growth Simulations, e.e.d. is a Bootstrap Error Estimate Assuming the Distribution is Correct, e.e.g. is an Error Estimate Assuming the Distribution is Gaussian, See the Text for Details

compound	time LJ/q (ns)	N	$\langle W \rangle$	$\sigma(W)$	$-{}^1/_2\beta\sigma^2$	$\langle W \rangle - {}^1/_2\beta\sigma^2$	ΔF	e.e.d.	e.e.g.
propane	0.2/−	400	6.9	8.0	−8.0	−1.1	0.1	0.5	1.1
propane	1.0/−	160	2.1	5.1	−3.2	−1.1	−0.6	0.4	0.6
propane	2.0/−	80	1.2	3.9	−1.9	−0.7	−0.7	0.5	0.6
chloroform	1.0/0.2	80	−3.8	6.2	−4.8	−8.6	−8.8	1.2	1.2
DMSO	1.0/0.2	80	−14.9	5.4	−3.7	−18.6	−18.2	0.7	1.0

decided to use particle mesh Ewald (PME)[49] with a direct-space cutoff of 0.9 nm and a grid spacing of 0.12 nm. PME is preferable over the twin-range 0.9/1.4 nm reaction-field because its energy conservation properties are better. This excludes the influence of cutoff errors on the measured work and also allows for a much lower friction in the Langevin thermostat. The additives are propane,[50] chloroform,[51] and DMSO.[52] The bond lengths in the BPA-PC as well as the internal geometries of the additives were kept fixed with the LINCS algorithm.[53] For the Lennard−Jones interaction, a cutoff of 0.9 nm was used with an analytical tail correction. To obtain a canonical distribution when the additive is nearly decoupled from the polymer matrix, a local thermostat is required. We chose to use the Langevin thermostat with a friction coefficient of 1 ps^{-1}. The temperature was set to 480 K; thus $k_BT = 4.0$ kJ mol^{-1}. A Berendsen barostat[54] was used with a reference pressure of 1 bar and a coupling time of 5 ps. All simulations were performed with the GROMACS simulation package.[55]

For the pure polymer matrix we performed a simulation of 10 ns. We checked that all the different bonded, Van der Waals, and electrostatic energy terms as well as the density were not affected by the difference in the simulation setup. From the last 8 ns of this simulation, starting configurations were taken at regular intervals for fast-growth TI simulations.

The fast-growth simulations were performed in two steps. First the Lennard Jones (LJ) interactions of the additive are coupled to the matrix using a soft-core potential:

$$V_{sc}(r) = \lambda V_{LJ}(r_{sc}(r; \lambda)) \tag{4}$$

$$r_{sc}(r; \lambda) = [\alpha\sigma_{LJ}{}^6(1 - \lambda) + r^6]^{1/6} \tag{5}$$

where the coupling parameter λ goes from 0 (uncoupled) to 1 (fully coupled), V_{LJ} is the normal, hard-core, LJ potential, σ_{LJ} is the LJ distance, and the soft core parameter α was set to 0.7. For the chloroform and DMSO additives, a second step is required to couple the Coulomb interactions. In this step the partial charges of the additive molecule (with full LJ interactions) are linearly changed from zero to the normal partial charge of the force field. Although λ changes continuously with time, the integration of the equations of motion is performed in a discretized fashion. So what actually happens is that the system evolves canonically over one MD time step Δt, after which the coupling parameter is changed by $\Delta\lambda = \Delta t/t_s$, where t_s is the total simulation time. The work performed by the change of Hamiltonian at each step is defined by

$$\Delta W = \Delta\lambda \frac{\partial H(\mathbf{x}, \mathbf{p}; \lambda)}{\partial\lambda} \tag{6}$$

The work W for the coupling is the sum of the work over all the MD time steps for both coupling steps. This work then needs to be averaged exponentially over many fast-growth simulations to obtain the free energy difference ΔF (see eq 3). Because we are simulating a constant-pressure ensemble, there is also a $p\Delta V$ term in the excess chemical potential. However at low pressures this term can be neglected, since it is 3 orders of magnitude smaller than the work.

To ensure canonical starting conditions, we generated uniformly distributed random positions and orientations of the additive molecules in the polymer matrix, as well as Gaussian distributed velocity components at 480 K. Since initially the additive is fully decoupled from the matrix, the starting conditions of the matrix do not couple to the starting conditions of the additive. The only remaining issue is whether the 8 ns trajectory of BPA-PC is long enough to be representative for the system. Although the correlation times in the system are much longer than 8 ns, the large number of BPA-PC monomers (250) should hopefully provide enough statistics. The only remaining parameter is the length of the fast-growth simulations. The required length will depend on how the favorable location for additives are distributed in the matrix, as well as on how easily the additive can create space for itself. These two issues make it difficult to estimate the optimal simulation length beforehand. For propane we tried 0.2, 1.0, and 2.0 ns for growing in the LJ interactions; for chloroform and DMSO we used 1.0 ns only. Growing in the charges in chloroform and DMSO is relatively easy, because the charge interaction with the matrix is weak; for DMSO, the work contribution for growing in the charges is between 1 and 2 k_BT. For good sampling, one should sample the rotational correlation time several times, which is around 5 ps. We chose a time of 0.2 ns to grow in the charges. This could be much shorter, but it is still almost negligible compared to the 1.0 ns for the LJ part. The numbers of simulations are shown in Table 1. To check the influence of the electrostatics treatment we repeated the DMSO case with a twin-range 0.9/1.4 reaction-field electrostatics setup and found no significant difference in free energy.

IV. Results

The resulting free energies are given in Table 1. The distributions of the work are shown in Figures 2−5. The figures also show the exponentially reweighted distributions (empty bars), from which one can see that noise in the low-energy tail of the distributions is blown up. Since the distributions are close to Gaussian, we also plotted the Gaussian distribution with the same mean and spread in the figures:

$$G(W; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(W - \mu)^2}{2\sigma^2}\right) \tag{7}$$

The Gaussian distribution can be exponentially reweighted:

$$\frac{e^{-\beta W} G(W; \mu, \sigma)}{\int_{W=-\infty}^{\infty} e^{-\beta W} G(W; \mu, \sigma)\, dW} = G(W; \mu - \beta\sigma^2, \sigma) \tag{8}$$

The result is again Gaussian but shifted by $-\beta\sigma^2$. The shifted Gaussians are also plotted in the figures. One can observe that, as expected, the positive tail of the reweighted curve is sampled accurately, whereas part of the negative tail is missing. Because the weighting for the free energy in the re-weighted distribution is linear, one can easily see how much one is missing (assuming that the underlying distribution is really Gaussian). Except for the 0.2 ns propane simulations (Figure 2), it seems that the missing tail is at least partially compensated by some high peaks. We cannot be sure that the distribution is Gaussian, for that at least an order of magnitude more points are required. However if the distributions are close to Gaussian, the scattered high peaks in the reweighted distribution of, for instance, Figure 4 are exactly what one would expect, because with the given number of simulations there would be 0, 1, or 2 points in each bin of the tail. When the distributions are Gaussian, the effect of the

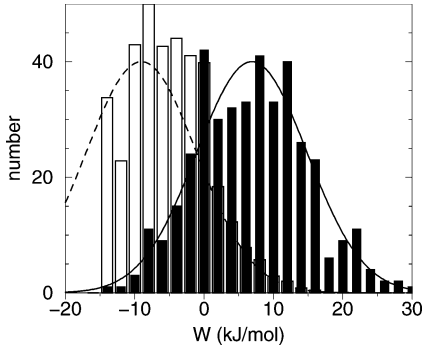

Figure 2. The work distribution for the 0.2 ns simulations of propane (solid bars). The empty bars show the exponentially reweighted distribution. The solid line is a Gaussian curve with average $\langle W \rangle$ and standard deviation $\sigma(W)$. The dashed line is the same Gaussian curve shifted by $-\beta\sigma^2$.

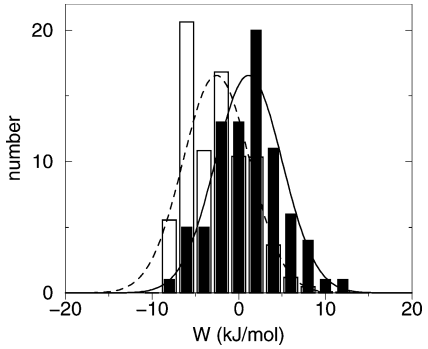

Figure 3. As Figure 2 but for the 2 ns simulations of propane.

missing tail can be quantified by determining the free energy for a Gaussian distribution:

$$\Delta F_G = -k_B T \log \left[\int_{W=-\infty}^{\infty} e^{-\beta W} G(W; \mu, \sigma)\, dW\right] = \mu - \frac{1}{2}\beta\sigma^2 \quad (9)$$

This result has been demonstrated before by Hummer.[38] The irreversible work is $-1/2\beta\sigma^2$, i.e., proportional to the variance of the distribution. In Table 1, ΔF can be compared with ΔF_G. The difference is less than 0.5 kJ mol^{-1}, except for the 0.2 ns propane simulations where the difference is 1.2 kJ mol.

V. Error Analysis

For understanding the different sources that contribute to the error in the free-energy calculated with the fast-growth method, it is useful to again look at the two extreme cases. With infinitely fast growth (the perturbation method), the error depends only on the overlap between the ensembles A and B and the number of perturbations that are performed. For the slow limit (slow growth TI), the error is determined by how accurately the ensemble is sampled at each intermediate state between A and B. This depends on the free-energy barriers between different conformations and the sampling time at each intermediate state. A mixture of these extremes will determine the error for fast-growth simulations in general. It is not possible to determine the relative contribution of the different sources.

In fast growth, on the one hand, one wants the individual simulations to be as long as possible to minimize the irreversible work along each path. However on the other hand, one wants many simulations to obtain good statistics of the work distribution. Because the total computational effort is the number of simulations times the simulation length, the question is what the optimal balance is. If there are no significant barriers in the free-energy landscape, the limit of slow growth will be the most efficient, although not significantly more efficient than fast growth when the irreversible work is around $k_B T$, as will be shown later. The most suitable systems for fast growth are those where there are high barriers between different conformations at fixed λ-states between A and B but no high barriers along paths from conformations in state A to conformations in state B, as illustrated in Figure 1. This is exactly the case for the free-energy of additives in a polymer matrix, as additives can be easily "grown in", but in state B additives will not move from one place in the matrix to another in times reachable by simulation. For a given total simulation time there will be an optimal number of simulations that is larger than one, since one simulation gives bad statistics over the end states.

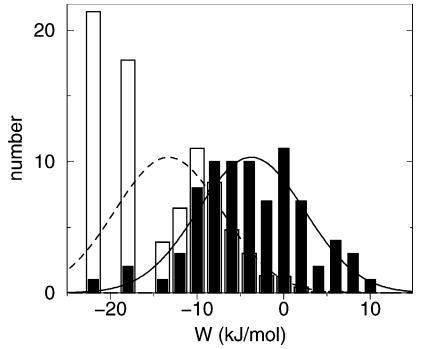

Figure 4. As Figure 2 but for the 1 ns + 0.2 ns simulations of chloroform.

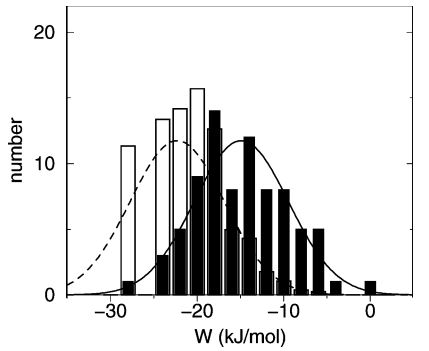

Figure 5. As Figure 2 but for the 1 ns + 0.2 ns simulations of DMSO.

Most ways of determining error estimates for linearly averaged quantities do not work, since the nonlinear averaging produces systematically too high values for smaller numbers of samples. The real problem is that the shape of the work distribution is unknown. There are two ways one can approach this issue. One way is assuming that the observed work distribution of N values is (close to) the correct distribution. Then one can use the bootstrap method to estimate the error. In this method we randomly pull N samples out of the distribution and determine the free energy from those. This procedure is repeated 10^5 times, and the error estimate is then given by the standard error of the resulting free energies, the results are given in Table 1. Generally this error estimate will be too low, because it does not detect a lack of sampling of the important negative tail. The only way to estimate this effect is by making an assumption about the shape of the negative tail. In this case the obvious choice is a Gaussian distribution. We generated N random Gaussian numbers with the variance σ^2 given in Table 1. We repeated this process a million times and determined the standard error as the standard deviation of the resulting free energies. This procedure ignores the uncertainty in σ, but the values of N used here are large enough to make

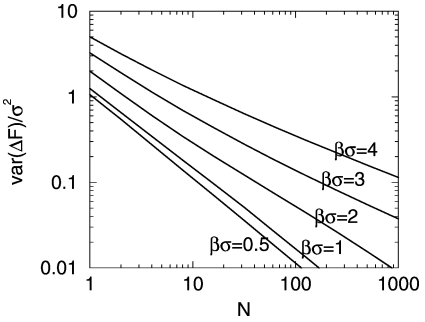

Figure 6. The variance in ΔF normalized by σ^2 for a Gaussian distribution as a function of the number of simulations N for different values of $\beta\sigma$.

Table 2. The Root-Mean-Square Error in ΔF for Propane in Units of k_BT as a Function of the Fast-Growth Run Time and the Total Run Time, Which Is the Run Time Times the Number of Simulations, the Estimates are Made Assuming a Gaussian Distribution with the Indicated Value of σ

		total run time			
fast-growth run time (ns)	$\beta^2\sigma^2$	10 ns	20 ns	40 ns	80 ns
0.2	4.00	0.60	0.45	0.36	0.27
1.0	1.61	0.54	0.40	0.30	0.22
2.0	0.93	0.51	0.37	0.27	0.20

the contribution of this uncertainty negligible. The error estimates are between 0.6 and 1.2 kJ mol^{-1} (see Table 1). For most cases the Gaussian error estimate is only slightly larger than the bootstrap error estimate, which indicates that the sampling is good. Only for the 200 ps propane simulations the difference is more than a factor of 2, because the bootstrap estimate does not see the missing tail. Here the Gaussian estimate gives a reasonable value, which is close to the sum of the statistical error in the variance $\sigma(W)/\sqrt{N-1} = 0.4$ kJ/mol and the systematic shift of ΔF compared with the 1 and 2 ns simulations.

We also determined the error as a function of the total run time, which is the number of simulations times the fast-growth simulation time, assuming that the distributions are Gaussian with σ as given in Table 1. In Table 2 one can see that, of the three simulation times used, 2 ns is always the best choice, but the differences with 1 ns are less than 10%. In the following we will show that simulation times longer than 2 ns (which means $\beta\sigma < 1$) will not give a significant improvement.

A more general error analysis on Gaussian systems has been performed by Gore et al.[56] To provide a complete picture on the tractable parameter range, we also determined the variance of ΔF for σ ranging from 0.5 k_BT to 4 k_BT and N ranging from 1 to 1024, again using a million "experiments" per parameter pair. The results are shown in Figure 6, normalized by σ^2. When $\beta\sigma$ is smaller than 1, the variance decreases as $1/N$, just as for a linearly averaged quantity. As $\beta\sigma$ increases, the error decreases slower with N, especially for relatively large N. Thus it becomes increasingly difficult to reach a reasonable accuracy. This is due to the nonlinear, exponential averaging. Deriving an analytical approximation of the variance which takes the exponential weighting into account is very difficult, if not impossible. Yet there exists an analytical approximation for the variance of ΔF given by Hummer:[38]

$$\mathrm{var}(\Delta F) \approx \frac{\sigma^2}{N} + \frac{\beta^2\sigma^4}{2(N-1)} \qquad (10)$$

where the work is linearly weighted, which gives a useful analytical relationship between the error in the free energy and N and σ. Because of the linear weighting, this analytical approximation ($1/N$ behavior) is limited to cases where σ is smaller than k_BT. We find that for the simulations presented here, eq 10 underestimates the error by a factor of 1.2−1.6.

When σ becomes larger than k_BT, the main source of errors is no longer the inaccurate sampling of μ or σ but insufficient sampling of the negative tail of the work distribution. This is due to the shift of the sampled and the reweighted distribution, which is equal to $-\beta\sigma^2$ (eq 8). Kofke has shown that the missing values in the tail are the most important source of error in many cases.[57] In the Appendix we derive that for a Gaussian work distribution and a given error the required number of simulations increases exponentially with $^1/_2\beta^2\sigma^2$. This can also be observed in Figure 6. Thus, even with short simulation times or large computational resources, the fast-growth method is limited to small values of $\beta\sigma$. This holds not only for Gaussian distributions but for any work distribution that has a small but nonzero negative tail.

With fast growth, one has the choice between doing a few long simulations (small N, small σ) or many short simulations (large N, large σ). To determine the optimal fast-growth run time or value of σ, one needs to know the relationship between these two quantities. In the situation where in the end state there are several conformations separated by high free-energy barriers, one can distinguish two contributions to σ: one from the difference in free energies of conformations in the end state and one from irreversible work along each path. When the latter contribution dominates, σ can be reduced by increasing the simulation time. For a quantitative relation, the simplest assumption is that σ^2 is inversely proportional to the simulation time, i.e., doubling the simulation time will halve σ^2. Table 1 shows that this nearly true for propane when going from 1 to 2 ns. In this case one can see in Figure 6 that values of σ smaller than k_BT will all give nearly the same efficiency, since the variance is then close to σ^2/N. For larger $\beta\sigma$, we have shown that the number of simulations required increases exponentially with $^1/_2\beta^2\sigma^2$ and therefore fast growth rapidly becomes computationally infeasible. This analysis assumes that N is large enough to give a good sampling of the conformations at the end state.

The choice between slow growth TI and fast growth depends on the relaxation times along growth paths and at the end state (or intermediate states). When relaxation along a path is slow and there is only one relevant end conformation, slow growth is the method of choice. A good example is solvating or charging a small molecule in water. If, however, there are multiple conformations in the end state that are separated by high free-energy barriers, fast growth is computationally more efficient. This is the case for free energies of additives in a polymer matrix. For optimal efficiency, the fast-growth simulation length should be chosen such that σ is close to k_BT. At this value, the effect of the exponential reweighting on the efficiency is negligible, while it provides the best sampling of the end state. With larger values of σ, one would have more simulations but the strong exponential reweighting will lead to fewer "relevant" conformations in the end state, i.e., fewer conformations with a non-negligible weight. The number of relevant conformations in the end state is especially important when there are many different conformations in the end state.

VI. Structural Analysis

Finally we take a look at the spatial distribution of the additives in the matrix. Only the spatial distribution seems to

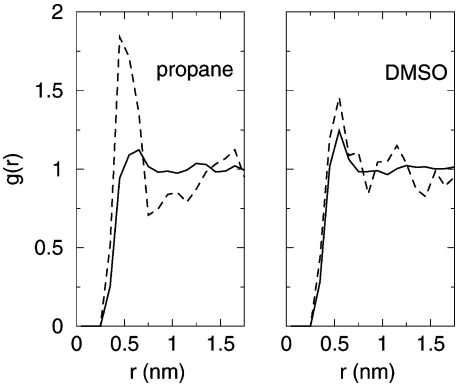

Figure 7. The radial distribution function of the middle carbon of propane (left panel) and DMSO sulfur (right panel) in the fully coupled end state with all phenyl carbons of the BPA-PC matrix (solid lines) and the three phenyl carbons at each end of BPA-PC (dashed lines). The end states have been weighted with $\exp(-\beta W)$.

be important, because we did not observe a correlation of the work with the time of the starting configuration of the BPA-PC matrix. Visual inspection of the spatial distributions revealed that the fully coupled additives do not cluster significantly. From this one can conclude that the additives do not prefer rarely occurring locations (e.g., a few large cavities inside the matrix). One can, however, observe that the additives often end up close to the end groups of the BPA-PC chains. To quantify this we have made radial distribution functions of the fully coupled solutes with all the carbons in the phenyl groups and with only the three phenyl carbons at the ends of the BPA-PC chains. Note that the conformations need to be weighted with the Boltzmann factor of the work. For the propane simulations, both 1 and 2 ns, we used the middle carbon of propane as the reference location and for DMSO we used the sulfur atom. For chloroform, statistically meaningful results could not be obtained, because one conformation (the lowest energy one) contributes a quarter to the average. The distributions (see Figure 7) show that there is a strong preference of propane residing close to the end groups. For DMSO, the preference is much weaker. This is not due to electrostatic effects, since repeating the analysis for DMSO without partial charges, which is obtained as an intermediate state, gives the same behavior. Furthermore we did not find a preference of DMSO for the polar carbonate groups. We speculate that propane is small enough to fit into existing cavities which are present around chain ends, while DMSO is too large. Note that the length of five monomers for BPA-PC used in this study is relatively short; for longer polymers the effect of chain ends will be smaller.

VII. Conclusions

We have illustrated that the fast-growth TI method applied to large-size additive molecules in dense polymer microstructures is capable of providing very accurate estimates of the additives' excess chemical potential. The common application of fast growth to conformation changes in macromolecules suffers from the problems that it is difficult to exactly define the two end states and that there can be very many paths between two states that all need to be sampled to obtain the correct exponential weights for the work. For the free energy of additives in a polymer matrix, the situation is exactly opposite in several aspects. It is trivial to generate a canonical starting ensemble, because the additive is initially completely decoupled from the matrix, however the end state where the additive is fully coupled is unknown. Because it is relatively easy to generate a good ensemble of paths from the beginning to the end state, one obtains a proper ensemble for the end state. With a fast-growth time of 1 or 2 ns, the irreversible work is small enough to obtain good statistics. Furthermore the work distributions are close to Gaussian, which allows for a thorough error analysis.

Fast growth is the method of choice when there are high free-energy barriers between different conformations in the end state. As with any method that uses exponential reweighting, the results are sensitive to the (non-)sampling of the negative tail. We have shown that for (near-)Gaussian distributions, a standard deviation of the work close to k_BT gives the highest computational efficiency; it provides the best tradeoff between the negative effects of exponential reweighting and the number of conformations sampled at the end state. For additives, molecules with three or four heavy atoms such a spread can be reached with fast-growth simulations of around 2 ns. For larger additives, longer simulations might be required. Finally the spatial distributions of fully coupled solutes obtained from a series of fast-growth trajectories provides a good picture of where an additive prefers to reside in the matrix. Because in the presented work, chain lengths are relatively short, we cannot compare calculated excess chemical potentials with experimental data. In a future publication, we will make such a comparision based on higher molecular weight BPA-PC systems for which experimental solubility data have been reported.

Appendix

In this appendix we derive an analytical estimate for how much of the negative tail of a Gaussian work distribution one can afford to miss when one wants to obtain a certain accuracy for the free energy. The free energy with the negative tail missing up to a work of W_m is given by the integral over eq 8, with W_m as the lower limit of the integrals:

$$\Delta F_m = -k_B T \log\left[\frac{\int_{W=W_m}^{\infty} e^{-\beta W} G(W;\mu,\sigma)\,dW}{\int_{W=W_m}^{\infty} G(W;\mu,\sigma)\,dW}\right] \quad (11)$$

When $W_m < -\sigma$ we can approximate the denominator in eq 11 by 1 and we will call the result ΔF_m^*:

$$\Delta F_m^* = -k_B T \log\left[\int_{W=W_m}^{\infty} e^{-\beta W} G(W;\mu,\sigma)\,dW\right] \quad (12)$$

For this approximation we can determine W_m when the shift in free energy is $\log(2)k_BT$, as for this value there is an analytic solution:

$$\Delta F_m^* - \Delta F_G = \log(2) k_B T \Rightarrow W_m^* = \mu - \beta\sigma^2 \quad (13)$$

Integrating the Gaussian distribution (eq 7) up to this value gives the fraction one can miss:

$$\int_{-\infty}^{\mu-\beta\sigma^2} G(W;\mu,\sigma)\,dW \approx \frac{1}{\sqrt{2\pi}\beta\sigma} e^{-1/2\,\beta^2\sigma^2} \quad (14)$$

where the right-hand side is an upper limit which is accurate for $\beta\sigma > 1$. Thus the number of simulations required for an accurate sampling of the negative tail increases exponentially with $^1/_2\beta^2\sigma^2$.

Acknowledgment. The authors thank Ben Reynwar for carefully reading and discussing the manuscript.

References and Notes

(1) Müller-Plathe, F. *Macromolecules* **1991**, *24*, 6475.
(2) Sok, R. M.; Berendsen, H. J. C.; Van Gunsteren, W. F. *J. Chem. Phys.* **1992**, *96*, 4699.
(3) Gusev, A. A.; Suter, U. W. *J. Chem. Phys.* **1993**, *99*, 2228.
(4) Tamai, Y.; Tanaka, H.; Nakanishi, K. *Macromolecules* **1995**, *28* (7), 2544 .
(5) Van der Vegt, N. F. A.; Briels, W. J.; Wessling, M.; Strathmann, H. *J. Chem. Phys.* **1996**, *105*, 8849.
(6) Cuthbert, T. R.; Wagner, N. J.; Paulaitis, M. E. *Macromolecules* **1997**, *30*, 3058.
(7) Hofmann, D.; Fritz, L.; Ulbrich, J.; Schepers, C.; Bohning, M. *Macromol. Theory Simul.* **2000**, *9*, 293.
(8) Van der Vegt, N. F. A. *J. Membr. Sci.* **2002**, *205*, 125.
(9) Neyertz, S.; Douanne, A.; Brown, D. *J. Membr. Sci.* **2006**, *280*, 517.
(10) Spyriouni, T.; Economou, I. G.; Theodorou, D. N. *Macromolecules* **1997**, *30*, 4744.
(11) Raptis, V.; Economou, I. G.; Theodorou, D. N.; Petrou, J.; Petropoulos, J. H. *Macromolecules* **2004**, *37*, 1102.
(12) Makrodimitri, Z. A.; Dohrn, R.; Economou, I. G. *Macromolecules* **2007**, *40*, 1720.
(13) Stapleton, M. R.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1990**, *92*, 1285.
(14) Knopp, B.; Suter, U. W.; Gusev, A. A. *Macromolecules* **1997**, *30*, 6107.
(15) Knopp, B.; Suter, U. W. *Macromolecules* **1997**, *30*, 6114.
(16) van der Vegt, N. F. A.; Briels, W. J. *J. Chem. Phys.* **1998**, *109* (17), 7578.
(17) Hummer, G.; Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (7), 3658.
(18) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78* (14), 2690.
(19) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(20) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(21) Bennett, C. H. *J. Comput. Phys.* **1976**, *22* (2), 245.
(22) Lu, N. D.; Singh, J. K.; Kofke, D. A. *J. Chem. Phys.* **2003**, *118* (7), 2977.
(23) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122* (14), 144107.
(24) Pitera, J. W.; van Gunsteren, W. F. *J. Phys. Chem. B* **2001**, *105*, 11264.
(25) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2007**, *126* (18), 184110.
(26) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578.
(27) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635.
(28) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.
(29) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275.
(30) Rodinger, T.; Pomes, R. *Curr. Opin. Struct. Biol* **2005**, *15* (2), 164.
(31) Meirovitch, H. *Curr. Opin. Struct. Biol.* **2007**, *17* (2), 181.
(32) Jarzynski, C. *Phys. Rev. E* **1997**, *56* (5), 5018.
(33) Crooks, G. E. *J. Stat. Phys.* **1998**, *90* (5−6), 1481.
(34) Crooks, G. E. *Phys. Rev. E* **1999**, *60* (3), 2721.
(35) Scholl-Paschinger, E.; Dellago, C. *J. Chem. Phys.* **2006**, *125* (5), 054105.
(36) Berendsen, H. J. C. *Simulating the Physical World*; Cambridge University Press: Cambridge, U.K., 2007.
(37) Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I.; Bustamante, C. *Science* **2002**, *296* (5574), 1832.
(38) Hummer, G. *J. Chem. Phys.* **2001**, *114* (17), 7330.
(39) Hendrix, D. A.; Jarzynski, C. *J. Chem. Phys.* **2001**, *114* (14), 5974.
(40) Lua, R. C.; Grosberg, A. Y. *J. Phys. Chem. B* **2005**, *109* (14), 6805.
(41) Oberhofer, H.; Dellago, C.; Geissler, P. L. *J. Phys. Chem. B* **2005**, *109* (14), 6902.
(42) Jarzynski, C. *Phys. Rev. E* **2006**, *73* (4), 046105.
(43) Oostenbrink, C.; van Gunsteren, W. F. *Chem. Phys.* **2006**, *323* (1), 102.
(44) Ytreberg, F. M.; Swendsen, R. H.; Zuckerman, D. M. *J. Chem. Phys.* **2006**, *125* (18), 184114.
(45) Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120* (13), 5946.
(46) Bastug, T.; Kuyucak, S. *Chem. Phys. Lett.* **2007**, *436* (4−6), 383.
(47) Hahn, O.; Mooney, D. A.; Müller-Plathe, D. A.; Kremer, K. J. *Chem. Phys.* **1999**, *111*, 6061.
(48) León, S.; van der Vegt, N.; Delle Site, L.; Kremer, K. *Macromolecules* **2005**, *38*, 8078.
(49) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(50) Schuler, L. D.; Daura, X.; Van Gunsteren, W. F. *J. Comput. Chem* **2001**, *22* (11), 1205.
(51) Tironi, I. G.; van Gunsteren, W. F. *Mol. Phys.* **1994**, *83*, 381.
(52) Geerke, D. P.; Oostenbrink, C.; van der Vegt, N. F. A.; Van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108* (4), 1436.
(53) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.
(54) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(55) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701.
(56) Gore, J.; Ritort, F.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12564.
(57) Kofke, D. A. *Mol. Phys.* **2006**, *104* (22), 3701.
(58) Here we use the phrasing "free energy difference". Because states A and B refer to an empty polymer matrix plus ideal gas penetrant and the polymer matrix plus the dissolved (fully interacting) penetrant, respectively, the free energy difference between states A and B equals the solute excess chemical potential. The *excess* chemical potential is defined relative to an ideal gas at the same number density; i.e. for penetrants without internal degrees of freedom the ideal gas part of the chemical potential equals $k_B T \log(\rho\Lambda^3)$ with ρ the number density of the dissolved penetrant and Λ^3 the momentum partition function.

MA702070N