JCTC Journal of Chemical Theory and Computation

Coarse Graining of Short Polythylene Chains for Studying Polymer Crystallization

Thomas Vettorel[†,‡] and Hendrik Meyer*[,†]

Institut Charles Sadron, 6 rue Boussingault, 67083 Strasbourg, France, and Institut für Physik, WA 331, Johannes-Gutenberg Universität, Staudinger Weg 7, D-55099 Mainz, Germany

Abstract: We derive coarse-grained models of polyethylene in the melt state with the aim to study polymer crystallization. This requires a low level of coarse-graining: We use a mapping of two CH_2 groups onto one bead. The coarse-grained beads are connected with harmonic springs, an optimized angular potential, and an optional torsional potential. Coarse-grained potentials are derived from detailed all-atom simulations, and an optimized form of the force field is then derived which achieves a good accuracy in reproducing the static properties of the chains. We address the question over which temperature range such models can be used, and in particular if the model is capable of reproducing the phase transition to an ordered state; it is found that the qualitative behavior of short polyethylene chains is well described, and the experimental melting temperature of $C_{44}H_{90}$ is approached when using the most accurate optimized model.

1. Introduction

Despite the progress of computer power, the simulation of complex materials with all atomic and electronic details remains unfeasible for many relevant applications. So, much effort is spent to construct simplified models retaining properties of the original chemistry.[1−3] Depending on the properties one is interested in, more or less details can be averaged out. Classical all-atom simulations can already be considered as a coarse-grained model with respect to quantum chemistry calculations by abandoning the electrons. The next steps are united-atom models which eliminate the pending hydrogens. A natural extension to coarser scales is to lump groups of atoms into one sphere. One now needs effective interaction parameters between these spheres which can be obtained via distribution functions from simulations on a more detailed level. The Boltzmann relation $P(x) = \exp(-F(x)/k_BT)$ relates the distribution with the associated free energy, the so-called potential of mean force $F(x) \sim -k_BT \ln P(x)$. One applies this to different distributions which may occur, as bond lengths, angles, and torsions along chains and radial distribution functions. One wants to avoid cross-correlations between these degrees of freedom for reasons of efficiency. So one should choose mapping centers such that cross-correlations are as small as possible.[4]

There are several further difficulties by applying this procedure: When $P(x)$ is the radial distribution function, $F(x)$ is used as a pair potential. However, $P(x)$ includes the effect of many-body interactions. Take for example a dense Lennard-Jones or even a hard sphere fluid: The potential has a single minimum or even no minimum; however, the radial distribution function exhibits long-range oscillations. This means that complex behavior is generated by a simple-looking potential. It is thus justified to make some simplifying assumptions about the interaction potential, e.g. about the cutoff of the interaction, and an optimization starting from the potential of mean force may be necessary. Different procedures for such an optimization process have been proposed to work e.g. with functional forms[4,5] or complete distribution functions.[6] Self-consistent optimization has been shown to work with rather long cutoff distances.[6−8] As we are working in dense melts, and to optimize the computational performance, we here choose rather short cutoff distances.

A related problem is that the distributions have been determined at a certain state point and that it is a priori not

* Corresponding author e-mail: hmeyer@ics.u-strasbg.fr.
† Institut Charles Sadron.
‡ Johannes-Gutenberg Universität.

Coarse Graining of Short Polythylene Chains

J. Chem. Theory Comput., Vol. 2, No. 3, 2006 **617**

clear how far the temperature or density can be modified to still yield reasonable results. This is one question we address in the present paper.

Another difficulty is the question of transferability when different chemical species are mixed. This is already well-known from the construction of atomistic force fields. Recent attempts have been made with simultaneous optimization of hydrocarbon substances[9] and an example to handle tacticity in polystyrene on a united monomer level.[10] When going to higher degrees of coarse graining, chemical specificity is lost very rapidly.

Finally, the question arises of which level of coarse-graining one should choose. Actually, when many particles are lumped together, their effective interaction becomes very soft.[11,12] This is very popular in the context of so-called dissipative-particle dynamics[13] and may be very efficient to handle phase separation of copolymer systems. For polymers, the softness may result in the problem that bonds of neighboring chains can cross each other, thus violating the fundamental property of topological constraints. Some effort has been spent by Padding and Briels[14] to circumvent this problem for the simulation of polymer rheology.

The present work is motivated by research on the polymer crystallization process. In this context, the role of possible precursor or transient phases is debated.[15−18] Polymer crystals exhibit not only positional order but also orientational and conformational order. This means that a simulation model must be able to represent the melt as well as a crystalline state. This restricts the possibility of coarse graining to a rather low level. We choose in this work the "united monomer" representation, i.e., two backbone carbons per coarse-grained bead.

The paper is organized as follows: Section 2 describes the atomistic simulations we performed to get reference data of distribution functions and describes then a coarse-grained model of poly-vinyl alcohol "CG-PVA" which has already been used to model polymer crystallization.[19,20] Section 3 describes the derivation of coarse-grained models of polyethylene on the "united-monomer" basis where still a lot of choices are possible. Finally, we discuss in section 4 how the optimized models perform when extending the temperature range to get crystallization.

2. Simulation Methods on Different Scales

2.1. All-Atom Model: *n*-Alkanes. We performed all-atom simulations of short polyethylene chains with the YASP simulation program[24] using a model derived from the OPLS force-field. The OPLS-AA force field[21] was developed on the basis of preexisting force fields in order to simulate organic compounds accurately. Parameters describing bond stretching and angle bending interactions were taken from the AMBER and CHARMM/22 force fields,[22,23] ab initio molecular orbital calculations were used for torsions, and Monte Carlo simulations were used for nonbonded interactions in adjusting the thermodynamic and structural properties.

The interactions described by this force field as given in ref 21 are the following: connectivity interaction (harmonic potential $U_{\text{bond}}(b)$ constraining the distance b between co-
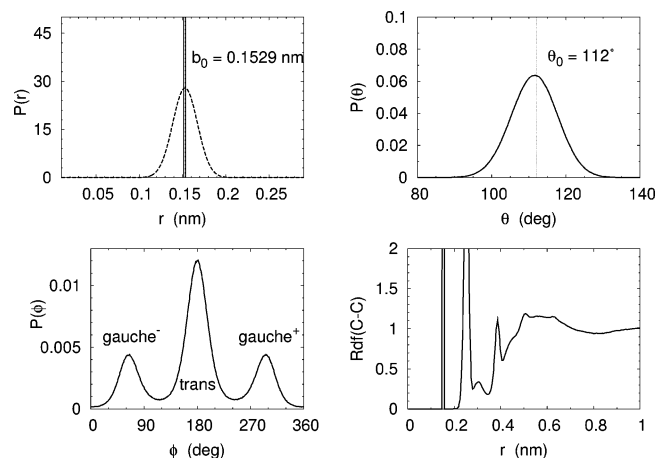

Figure 1. Probability distributions for the following quantities: bond length, bond angle, torsion angle, and radial distribution function (RDF). Only carbon atoms are considered here; these quantities are the result of the interactions of the force field. For the bond length distribution, the OPLS-AA prescription (dashed-line Gaussian curve) is shown in comparison with the delta function enforced by the constraint algorithm. The structure of the torsion potential exhibits the three favorable states trans, gauche⁺, and gauche⁻ which determine the flexibility of the chains. The first four peaks one can observe in the RDF are related to the three bonded interactions.

valently bound atoms); angular interaction (harmonic potential $U_{\text{ang}}(\theta)$ that determines the angle θ between two consecutive bonds); torsional interaction (3-term Fourier series)

$$U_{\text{tors}}(\phi) = \sum_{n=1}^{3}\left(\frac{1}{2}k_n^{\text{tors}}(1 - \cos(n\phi - \phi_n^0))\right) \qquad (1)$$

for the angle ϕ formed by two consecutive dihedrals (similar potential terms exist for torsion angle involving both carbon and hydrogen atoms); and nonbonded interactions (Lennard-Jones potential $U_{\text{mol}}(r)$ for atoms of different chains or of the same chain provided they are far enough from each other along the chain). The electrostatic interaction due to the partial charges q_a at every atom is also taken into account via the Coulomb interaction, which is cut off for distances larger than 0.9 nm.

Figure 1 shows the distributions of the quantities that are directly or mostly influenced by the above mentioned interactions, considering the carbon atoms of the chains. The bond lengths, bond angles, and torsion angles are mainly determined by their corresponding potentials; however, some influence by the other degrees of freedom exists as we shall see later on. The radial distribution function (RDF) is determined by all the different interactions, more or less importantly depending on the distance. The quantities of Figure 1 are the ones we shall consider while discussing the coarse-grained description of the *n*-alkane chains.

The YASP simulation package implements a constraint algorithm (SHAKE) that has been used instead of the harmonic potential parametrized in the OPLS-AA force field, to access larger time steps by neglecting the rapid (and

unimportant) vibrations around the average value b_0 of the bond length. The YASP program makes use of the Berendsen thermostat (velocity rescaling; the coupling time used is 0.2 ps) and barostat (coupling factor 2 ps × 10^{-6} kPa) to keep temperature and pressure constant. Runs have been performed at constant atmospheric pressure and various temperatures. We also implemented a Langevin thermostat which essentially yielded the same results.

The parameters for the torsional potential of eq 1 have been adjusted (optimization of the k_n^{tors} coefficients) in order to reproduce the probabilities of trans and gauche states as reported in other simulation studies where the parameters of the force field had been optimized to recover accurate static properties of short polyethylene chains.[25] The optimized values of the parameters are given as follows (no changes were made to the terms involving both carbon and hydrogen sites):

$$k_1^{tors} = 5; \quad k_2^{tors} = -1; \quad k_3^{tors} = 0.6 \qquad (2)$$

This optimization appeared necessary in view of the very strong influence of the torsional potential on the chains' conformation. Since torsions are the most flexible degree of freedom of the chains, they are mainly responsible for significant movements of the atoms on the local scale, and they also statistically determine the characteristics on the scale of the chain's size.

Atomistic simulations are calibrated in absolute units; the appropriate units here are nanometers (nm), picoseconds (ps), and kiloJoule (kJ) per mole. The time step used is 0.002 ps. We studied *n*-alkane molecules in the melt at atmospheric pressure and $T = 500$ K for chains containing 8, 10, 13, 14, 15, 16, 30, and 44 carbon atoms, in simulations runs lasting up to 20 ns. Table 1 summarizes a few quantities characterizing the simulation of systems with chains of different molecular weight at $T = 500$ K.

The model described above has been simulated at various temperatures, both above and below the experimental melting temperatures of the different *n*-alkanes. A continuous cooling protocol has been applied in order to trigger the transition toward the crystalline state, that comprises several different phases well characterized experimentally.[26] During cooling (at rates between -0.05 and -0.5 K ps^{-1}), no ordering of the chains was observed. The dynamics of the systems rather slowed as a glassy state was approached. Isothermal relaxation below the experimental melting point also could not reproduce the crystalline structure. One possible reason for this failure is a lack of precision in the model considered here, but it can also be related to the fact that nucleation barriers are rather high and thus makes nucleation in the small simulation volume a rare event.

There are many force fields one could use in order to simulate alkanes; the optimization of the parameters of such data sets is a difficult task. It is therefore not straightforward to find the most appropriate force field for a particular study. Moreover the optimization used here to *tune* the torsional interaction is questionable. However, the results obtained for static quantities show that the model still provides a reasonable description of *n*-alkane melts in the liquid state.

Furthermore, using this model to simulate the melting of perfectly crystalline arrangements of the chains proved that such structures are stable at low temperatures, and a "rotator" phase (transitory phase showing a hexagonal symmetry; see ref 26 for experimental details and ref 27 for early simulation studies) was observed while the system transformed from the solid to the liquid state. This fact demonstrates the ability of the present model to account for the behavior of the low-temperature phases of short polyethylene chains.

On the other hand, it is possible to observe a phase transition using continuous cooling from the melt for artificially stiff chains. Modifying the torsional interaction to make the trans state much more favorable turned out to enforce the crystallization of a $C_{16}H_{34}$ system at a temperature $T_{cryst} \approx 350$ K, which is large compared to the experimental value (around 300 K), whereas one would rather expect to find a crystallization temperature lower than the experimental value when applying such high cooling rates. At $T = 500$ K, this unrealistically stiff model has a probability of trans states more than 20% higher than our regular *n*-alkane model, which results in a persistence length and a squared radius of gyration, respectively, 35% and 15% larger.

There is another reason our all-atom simulations cannot reproduce a crystalline configuration, at least for a reasonably stiff model: One is restricted to studying very small systems over short time periods, otherwise the computational time needed becomes prohibitive. Therefore, the probability to form a crystal is drastically lowered: As mentioned above, the smaller the box, the lower the probability of a nucleation event, and the very fast cooling rate hinders the rearrangement of the chains in ordered conformations. This suggests that there might be other possible simulation schemes that would be more appropriate to approach the liquid−crystal transition; more precisely, as crystallization implies a modification of the chain's conformation it is more important to consider what happens on the length scale of the monomers rather than on that of the individual atoms. This is the idea the coarse-grained models rely on. We shall see how the data obtained from all-atom simulations can be used to develop a coarse-grained model that retains prominent features of the molecules while making their simulation more efficient and thus complex processes such as crystallization accessible.

2.2. Coarse-Grained Model: CG-PVA. The CG-PVA model is a coarse-grained model that allows a qualitatively accurate description of a crystallizable polymer, in that the typical semicrystalline structure is obtained after cooling simulations.[19,20] Figure 2 shows a schematic representation of both the coarse-grained model and the underlying PVA molecule: It can be seen that the coarser model lumps one monomer (i.e. 7 atoms here) into one bead, thus making the simulation much faster. The figure also represents the different interactions between such coarse-grained beads.

Besides the gain in computational efficiency this scheme is particularly interesting because the coarse-grained chain still has a close link to the underlying atomistic description via the definition of the effective potentials governing the behavior of the coarse-grained monomers. These interactions are determined using the all-atom simulation data, and thus
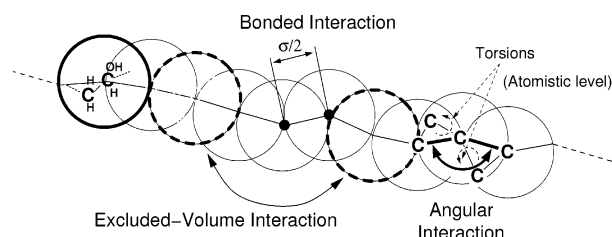
Coarse Graining of Short Polythylene Chains

J. Chem. Theory Comput., Vol. 2, No. 3, 2006 **619**

Figure 2. Schematic representation of the CG-PVA model. Atoms of the underlying molecule are shown to emphasize the mapping of one monomer onto one coarse-grained particle. The effective interactions between the latter are also depicted.

the coarse-grained model is really designed so as to reproduce the properties of the more detailed all-atom model.[4]

The interactions are the following: connectivity (harmonic potential $U_{bond}(b)$), angular interaction (tabulated potential $U_{ang}(\theta)$, see below), and nonbonded interaction (soft Lennard-Jones 6−9 potential $U_{mol}(r)$, without attractive part). There is no electrostatic interaction. Electrostatic effects accounted for by an explicit potential in the all-atom model are here absorbed in the effective nonbonded interaction.

The coarse-grained simulations were carried out using a simulation program similar to the YASP package but optimized for linear polymer chains and working with LJ units (see below). It is based on a velocity-Verlet integrator (time step 0.005 τ) and uses the Langevin thermostat instead of the Berendsen algorithm, with friction 0.5 τ^{-1}. The pressure is kept constant via position rescaling (Berendsen, coupling constant 10^{-5} $k_B T \sigma^{-3} \tau$).

Figure 3 presents the three effective potentials. The main difference when compared to the all-atom situation is the absence of a torsional potential and the form of the angular potential. The angular degree of freedom in the coarse-grained model actually plays the role of torsions on the atomistic scale: More precisely, the coarse-grained angle θ is exactly determined by two consecutive torsional states of the atomistic backbone. This is reflected in the structure of the angular potential which displays three minima corresponding to trans−trans, trans−gauche, and gauche−gauche states of the underlying chain. Since in the all-atom case the *torsional* degree of freedom is most important for changes in the chains' conformations, the *angular* degree of freedom becomes the most important ingredient of the coarse-grained model. The torsion angle *on the coarse-grained level* is far less relevant and shows a weak structure only. The narrow minimum at 180 ° in $U_{ang}(\theta)$ is responsible for the possibility that parts of the chains stretch (forming stems that constitute the crystalline lamellae), and the other minima allow that the folds between two stems are locally stable. In the melt state at high temperature, the barriers can be crossed easily. The procedure used to derive the coarse-grained effective potentials was similar to what is presented in the next section 3 for the case of polyethylene.

The units used in the coarse-grained simulation are the following (the same units will be used for the description of the coarse-grained models of PE): length scale: $\sigma = 0.52$ nm, average bond length between the coarse-grained monomers $b_0 = \sigma/2$ (please note that there is no relation between
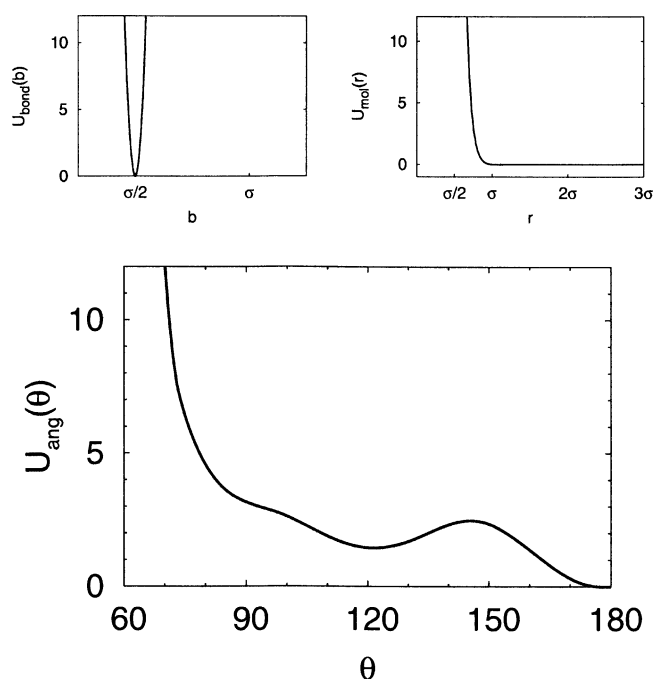

Figure 3. Effective interaction potentials between monomers in the CG-PVA model. The harmonic binding potential and soft Lennard-Jones nonbonded potential have simple analytical expressions, while the angular potential, also determined from all-atom simulations of PVA, has a more complex structure and has to be tabulated. The minima correspond to trans−trans, trans−gauche, and gauche−gauche states of the underlying atomistic backbone and lead to the formation of both stretched and folded parts of the chains at low temperature.

σ and the characteristic length σ_0 entering the Lennard-Jones potential definition, see eqs 4 and 6; mass scale: mass of one monomer taken to be 1 (all particles identical); and energy scale: $k_B T = 1$ for a reference temperature, here $T = 550$ K (temperature at which the data from atomistic simulations of PVA were taken to perform the mapping). The time scale cannot be determined in a straightforward manner as for the length scale, which is directly related to the way the mapping is performed. One possible way of setting this scale consists of a mapping involving the diffusion coefficient.

The properties of the CG-PVA model which allow for describing crystallizable polymer melts and to recover the correct qualitative structure are discussed in refs 19 and 28.

3. Coarse-Graining of the All-Atom Model

3.1. Derivation of Different Coarse-Grained Models. The coarse-graining procedure consists of defining effective interactions governing the behavior of particles on a larger length scale than the original model (thus also enforcing a longer time scale). The coarse-grained interactions are to be adjusted such that the static properties of the two models coincide. Obviously, since the coarse-graining process is devoted to "losing" some (irrelevant) information contained in the original model in order to make the simulation more efficient, the definition of the coarse-grained model cannot be unique, and choices have to be made so as to create an
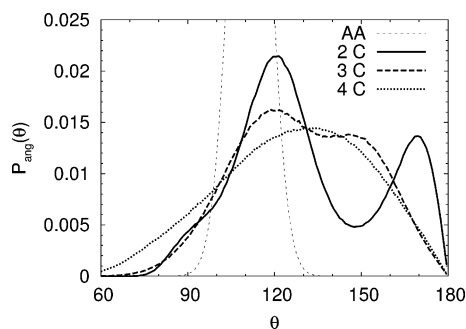
Figure 4. Comparison of the angular distributions obtained for different choices of coarse-graining level at $T = 550$ K. The original (harmonic) bond angle distribution is also indicated (AA). The distribution's structure weakens for mapping centers located 3 carbons away and completely disappears for centers separated by 4 bonds.

Figure 5. Distribution of bond lengths between two consecutive mapping centers measured from the all-atom simulation data. The inset shows the good agreement with the Gaussian distribution obtained using the harmonic potential of eq 4. The target distribution is slightly shifted toward smaller bond lengths.

appropriate model. In the case we are interested in, polyethylene (PE) melts and their behavior at low temperature experience acquired using the CG-PVA model tells us that a suitable coarse-grained model to access the semicrystalline structure through cooling of a melt has to retain enough features from the original chain. Moreover, the similarity of the PVA and PE chains prompts us to define a coarse-grained PE model in a way similar to the case of PVA. Since our studies of CG-PVA have shown that the key parameter in the model is the angular potential, it is natural to make the same choice here.[29] Thus we place the mapping centers on every other carbon atom of the atomistic backbone, which will yield a similar angular potential for the CG-PE model.

Figure 4 shows the influence of the degree of coarse graining on the structure of the angular potential: We computed the bond angle distributions for several choices of mapping centers. It can be seen that, when defining coarser models in which the interaction sites lie 2 or 3 carbons away from each other, the structure of the distribution (and thus of the effective angular potential to be derived) vanishes, leading to a bead-spring model with a generic, average rigidity as considered in the wormlike chain model.

Once the position of the mapping centers is chosen, one still has to define the type and characteristics of the effective interactions between such centers. In the case of simple linear polymer chains as for PVA and PE, the most natural choice is to define the very same type of interactions as for the standard all-atom models. Then the parameters of the bond, angular, torsional, and nonbonded interactions remain to be determined. To this end, using the assumption that the quantities constrained by the different effective interactions are weakly correlated, one can invert the probability distribution measured from the detailed simulation data via the Boltzmann formula in order to get the corresponding effective potential:

$$U(x) = -k_BT \ln P(x) \tag{3}$$

Of course, this method can only be approximate, since the quantity x is necessarily correlated to the other degrees of freedom of the system, and therefore $U(x)$ is not, strictly speaking, a potential energy. Furthermore, $U(x)$ as defined in eq 3 is only determined up to an additive constant, and a
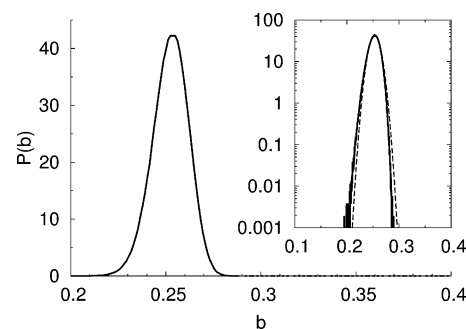
particular reference can be chosen for the energy. The inversion of the probability distribution is straightforward and accurate for the stiff degrees of freedom like the bond length. It becomes less obvious as the energy level of the effective potential decreases and the target quantity becomes more influenced by the medium. For the definition of the nonbonded interaction from the measured RDF for instance, more choices will have to be made for the form of the effective interaction.

With this methodology, we can now create several different coarse-grained models and adjust their parameters so as to reproduce first the target distributions $P(x)$ and then the other properties. Besides this, an obvious and very simple choice for a coarse-grained model of PE consists of taking the parameters from the CG-PVA model and modifying only the most important interaction from the all-atom simulation data, i.e., the angular potential. One can justify such an approach by assuming that, the structures of PVA and PE being overall very similar, the other parameters of CG-PVA should constitute an acceptable approximation for alkanes. We call this model "CG-PE$_0$"; the modified angular potential is derived in section 3.1.2, as part of the "regular" coarse-graining procedure that we shall present now.

In the following, we focus on simulations of 100 chains of $N = 22$ monomers which match our results for $C_{44}H_{90}$ obtained with the all-atom model.

3.1.1. Connectivity. The stiffest interaction in the system corresponds to the potential that ensures the connectivity of the consecutive particles inside a chain. As for covalent bonds between atoms in all-atom simulations (which in the case of our alkane model are treated as rigid constraints, see section 2.1), the fluctuations of coarse-grained beads around the average bond length is very rapid and relatively unaffected by the other interactions. Therefore the simplest description of the bonded interaction consists of assuming a harmonic form of the interaction

$$U_{bond}(b) = \frac{1}{2}k_{bond}(b - b_0)^2 \tag{4}$$

and determining the appropriate coupling constant k_{bond} via a fit of the data obtained from the all-atom simulations. Figure 5 presents the target distribution $P_{bond}(b)$ and the
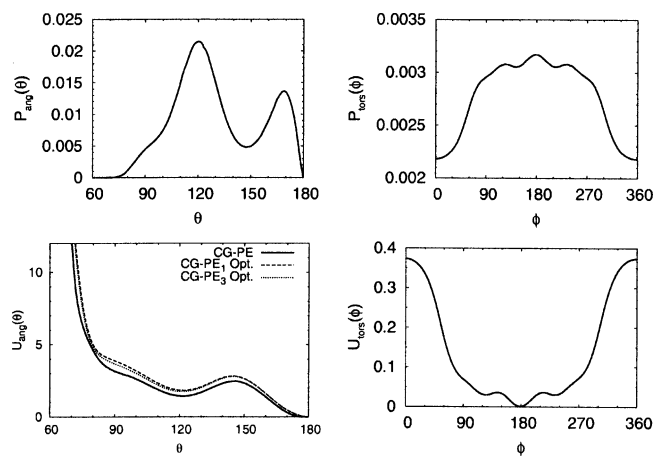
Figure 6. Bond angle and torsional angle distributions measured from the all-atom simulation and the corresponding effective potential derived from them using eq 3.

Gaussian fit reproducing well the all-atom data, except for a slight shift of the distribution toward the lower bond lengths. This effect is unimportant for our coarse-grained models. The fitting parameter takes the value $k_{bond} = 11476$ $k_B T \sigma^{-2}$; the average bond length $b_0 = 0.2535$ nm $\approx \sigma/2$ is the reference length in the coarse-grained simulations, see section 2.2.

3.1.2. Bond Angles. The bond angle distribution measured between every other carbon atom in the atomistic simulations exhibits a more complex structure than the bond length distribution or angular distribution on the atomistic level, as a result of the influence of the underlying torsional states as already mentioned (cf. Figure 6). Here it is thus not possible to find a simple analytic expression for the effective potential $U_{ang}(\theta)$, and one has to use the Boltzmann inversion

$$U_{ang}(\theta) = -k_B T \ln\frac{P_{ang}(\theta)}{\sin\theta} + U_{ang}^0 \qquad (5)$$

The resulting potential is also shown in Figure 6; the arbitrary constant U_{ang}^0 has been set so that the reference angular energy is 0 for trans−trans states. The structure of the coarse-grained angular potential obtained for *n*-alkanes is much less pronounced than the corresponding CG-PVA potential. Only two of the three expected favorable states are observable, the gauche−gauche state being almost totally absorbed into the neighboring trans−gauche state. This comes from the fact that the angular distribution is sharper in the case of PVA, with very weak probabilities of transition between two states. This is due to intrachain hydrogen bonds that stabilize the conformations of the PVA chains. The absence of hydrogen bonds in the PE melt causes the chains to become more flexible. Nevertheless, we shall see that the specific shape of the angular potential will strongly favor stretched states and thus drive the crystallization at low temperature.

3.1.3. Torsions. The distribution of torsion angles measured from the conformation of four consecutive mapping centers is shown in Figure 6 and can be inverted in the same fashion as has been done for the angular degree of freedom (except there is no sine from the jacobian in the Boltzmann inversion) to yield an effective torsional potential $U_{tors}(\phi)$.
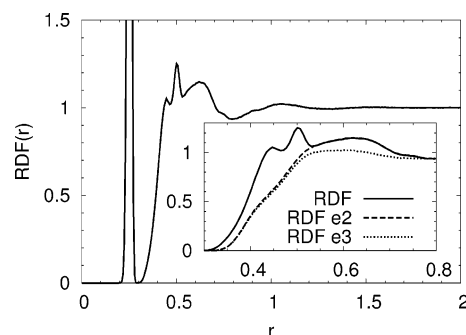

Figure 7. RDF measured from the position of every other carbon atom along the atomistic backbone. In the inset, the same distribution is compared to variants calculated with the 2 closest neighbors excluded (RDF e2) and the 3 closest neighbors excluded (RDF e3). The latter functions are used in the Boltzmann inversion formula when deriving the non-bonded effective potential in the case of a completely repulsive potential and a potential with an attractive part.

Although much less stiff than the angular potential (smaller energy scale), it still clearly favors the trans configuration of the coarse-grained chains. The trans state of the coarse-grained backbone should not be confused with the trans state of the original atomistic model, which determines the angular interaction. The definition of a torsional interaction on the coarse-grained level can be viewed as an attempt to fine-tune the description of the chains' flexibility on a larger length scale.

The torsional potential derived for CG-PE appears weak but still clearly favors the stretched conformations, thus it seemed worth including it in the coarse-grained model. To be sure about the influence, we constructed two models, one with and another without torsions.

3.1.4. Nonbonded Interactions. The effective interactions determined above are rather straightforward to parametrize, using either an analytic expression or direct Boltzmann inversion to map the data from all-atom simulations. Whether using these mappings in coarse-grained simulations will yield distributions that match the target functions is of course not obvious, but the method is clear. In the case of the nonbonded interaction $U_{mol}(r)$, it is not possible to neglect the influence of the other degrees of freedom anymore, and more choices have to be made. This is apparent when looking at the RDF presented in Figure 7: The first peaks are directly related to the positions of the nearest neighbors of one "monomer" (i.e. an ethylene group in the case of the PE chains considered here), and thus RDFs have to be calculated excluding these contributions the effect of which has already been taken into account when deriving the previous effective interactions. Figure 7 also shows RDFs where respectively the first two and three neighbors of one particle have not been considered in the calculation, in comparison with the full RDF.

One possible solution to determine $U_{mol}(r)$ consists of guessing a functional form as we did in the case of the binding interaction in section 3.1.1; CG-PVA uses a softer Lennard-Jones potential of the form

$$U_{mol}(r) = \epsilon_0\left(\left(\frac{\sigma_0}{r}\right)^9 - \left(\frac{\sigma_0}{r}\right)^6\right) \qquad (6)$$

which is truncated and shifted at the position of the minimum, yielding a fully repulsive interaction. It is possible to choose such a functional form and then try to adjust the values of the coefficients ϵ_0 and σ_0 in order to reproduce schematically the trend exhibited by the RDF. However, this method cannot be very precise, and it is possible to attempt to match the RDF more closely either with an optimization of (ϵ_0, σ_0) as presented in ref 4 or using the Boltzmann inversion on part of the target RDF.

In any case, the first choice to be made is related to the torsional interaction: Depending on whether an effective interaction is included in the coarse-grained model or not (in the case of CG-PVA, it was found unnecessary), the mapping for the nonbonded interaction is different. The reason for this is that we take the viewpoint that all interactions which are explicitly taken into account by bonded interaction potentials should not be considered in the nonbonded potential. Thus, if torsions are included, the 1−4 interaction (i.e. particles separated by three consecutive bonds) should be excluded from the nonbonded interaction and from the target distribution. If no torsions are considered, the 1−4 interaction should be also included in the RDF target distribution.

Since these RDFs take into account the correlations between any two particles in the system (except the close neighbors along the chain), they contain several peaks indicating the average positions of the neighbor shells. A direct Boltzmann inversion of such a function would yield a long-range potential which takes many-body interactions into account. To avoid this, a cutoff distance has to be chosen beyond which the effective potential is zero (see also ref 6). Here again, many possibilities arise. Since an essential aim of the coarse-graining procedure is to save computer time, we wish to have a potential with the shortest possible interaction range. Thus, the most natural possibilities consist in cutting off the potential either at the first *minimum* (yielding a fully repulsive interaction, as in the case of the CG-PVA model) or at the first *maximum* (so as to include an attractive part). The latter options appear equally sensible, since in the molten state the high density of the system screens the nonbonded interactions beyond the nearest-neighbor distance. Thus, keeping the density of the system constant either in constant-volume or constant-pressure simulations should yield equivalent results, at least for simulations at the given state point. However, note that the thermodynamic properties involving derivatives as the equation of state are strongly influenced by the attractive tail of the potential. Thus, to ensure a better transferability in the temperature or pressure range, a potential with an attractive part is probably the better choice.

The simulation procedure applied consisted in a first equilibration run at constant temperature and volume, using for the density the value extracted from the all-atom simulations. During this NVT simulation, the average pressure is measured. This pressure is applied in the following constant-pressure simulations for a second equilibration step and the cooling-heating cycles. The value of the pressure is arbitrary in our simulations and differs depending on the model. The values of the pressure (in $k_B T \sigma^{-3}$) for the different

Table 1: Static Properties of Different *n*-Alkanes Obtained by All-Atom Simulations[a]

system	chains	duration (ns)	l_p/b_0	l_p (nm)	R_g^2 (nm²)	R_e^2/R_g^2	D (nm² ps⁻¹)
C_8H_{18}	256	2	2.17	0.332	0.0707	7.78	0.01273
$C_{16}H_{34}$	84	20	2.33	0.357	0.205	7.93	0.00353
$C_{44}H_{90}$	100	4	2.35	0.359	0.831	6.98	0.0080

[a] For the three different chain lengths presented here, simulation details (number of chains and length of the simulation runs) are indicated, together with the following properties: Persistence length l_p (measured as the characteristic length of the vector orientation decorrelation along one chain, for vectors joining every other carbon atom), squared radius of gyration R_g^2 and end-to-end distance R_e^2, and diffusion coefficient.

Table 2: Summary of the Force-Field Characteristics for the Different CG-PE Models

model	nonbonded potential	torsional potential	applied pressure (at $T=1$)
$CG-PE_1$	repulsive	explicit	3.73
$CG-PE_2$	repulsive	implicit	4.36
$CG-PE_3$	attractive	explicit	2.74
$CG-PE_4$	attractive	implicit	1.80

coarse-grained models are indicated in Table 2, together with a summary of the force-field characteristics. It can be seen that a lower pressure is sufficient to achieve the right density in the case of the models with an attractive nonbonded interaction.

In case surfaces are involved in the study, the model must be optimized such that the applied pressure is 0. References 6 and 10 solved this by adding a linear term to the nonbonded potential. However, one should probably include the pressure (as well as other thermodynamic quantities as mentioned above) in the optimization procedure. Here, we are only interested in simulations of the melt and can treat the pressure as a force-field parameter.

3.2. Optimization of the Coarse-Grained Models. The prescription given above for the different coarse-grained models has been tested in simulation; the first requirement of the coarse-graining approach is that the target distributions extracted from the all-atom simulation data are reasonably well reproduced with the simplified scheme. This should ensure that the static properties are very close in the two simulation models.

The bond length distribution, as expected, exactly matches (within the line width) the Gaussian function obtained by inversion of the prescribed potential, whatever model is used. This is not surprising since the latter potential is very stiff, and therefore the bond interaction is not influenced by the other degrees of freedom.

Contrary to that, the angular distribution shows significant deviations from the target, for all models (cf. Figure 8). This is very different from what has been observed for the CG-PVA model for which the measured angular distribution reproduces almost exactly the prescribed one (with only minor differences which do not significantly affect the population of the different, well-separated states). The reason for that is the much smoother form of the angular distribution for PE, as already mentioned above. In particular, it can be
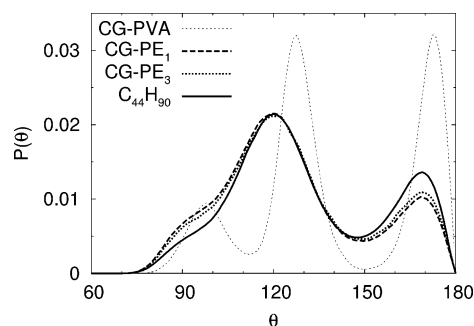
Figure 8. Comparison of the angular distributions obtained with the various models at $T = 1$. The probability of trans−trans states appears reduced compared to the target distribution from all-atom models.
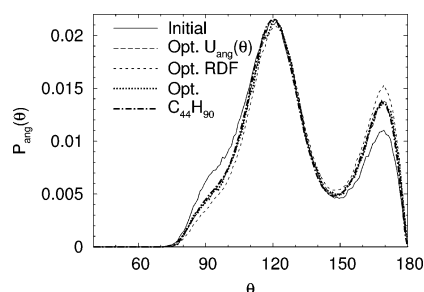

Figure 9. Evolution of the angular distribution $P_{ang}(\theta)$ during optimization of $U_{ang}(\theta)$ for the CG-PE$_3$ model. The *initial* distribution is obtained directly from the potential inverted from the atomistic $P_{ang}^{target}(\theta)$, i.e., without any optimization. After optimizing $U_{ang}(\theta)$ the distribution matches the target (Opt. $U_{ang}(\theta)$) but deviates again after the optimization of the RDF (Opt. RDF). The agreement is fine again after a second optimization of both $U_{ang}(\theta)$ and $U_{mol}(r)$ (Opt.).

noticed that the population of the trans−trans state is significantly lowered for all coarse-grained models, whereas the shoulder at 90° is higher than in the target distribution. This effect seems slightly more pronounced in the case of models having no explicit torsional potential, surprisingly (one would expect the torsional potential to make the chains stiffer, but this probably only occurs on larger length scales). This lowering of the probability of trans−trans states is probably due to the density of the melt which tends to compress the chains and makes folded angular states more favorable than the stretched states.

The torsional distribution that is enforced via an explicit torsional potential appears to be well reproduced for the models CG-PE$_1$ and CG-PE$_3$, whereas the other models have a less pronounced distribution of torsional angles, showing nevertheless the same trend (more favorable trans state), see also Figure 10.

The RDF measured on the data obtained with the various coarse-grained models proved to depart significantly from the original function. These discrepancies are directly related to the mismatch described above for the angular interaction: The peaks associated with the angular states are too low, and a shoulder appears between the first and second peak in the global RDF, corresponding to the too large proportion of folded states already observed. Apart from these differences concerning the short-range particle−particle
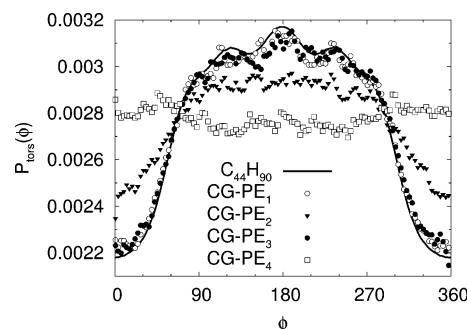

Figure 10. Torsional distribution obtained for the different coarse-grained models. The results for CG-PE$_1$ and CG-PE$_3$ which both include an explicit torsional potential are very close to the target distribution, whereas the other two models exhibit significant deviations; for CG-PE$_4$, the distribution is almost completely flat, indicating that no particular torsion angle is favored. CG-PE$_2$ has some structure identical in trend to the stiffer models. (Gliding averages have been applied to the data.)

correlations, one can notice that the subsequent peaks of the RDF, associated to the neighbor shells, are shifted. These long-range differences are more difficult to relate to the shape of the effective potentials, which all necessarily have short ranges. The mismatch between the coarse-grained and the all-atom RDF varies with the type of coarse-grained model but not qualitatively.

The previous observations show that the coarse-grained models "simply" derived, i.e., with effective potentials directly inverted from the target distributions, are not sufficient to reproduce the correct behavior of a dense polymer melt, even at high temperatures where one can hope that the different degrees of freedom are relatively independent from each other. This only works well in the case of very sharp distributions. Thus, an additional step is needed in the coarse-graining procedure: The effective potentials associated with distributions that are not well reproduced in the coarse-grained simulations have to be optimized.

The optimization of the effective potentials is an iterative procedure during which one coarse-grained potential is modified until the corresponding probability distribution matches the target distribution.[2,3] This has to be performed for all the degrees of freedom that are not well described compared to the reference all-atom simulation. In doing so, it is sensible to consider the interaction in decreasing order of intensity, i.e., the strongest potential should be optimized first, and the method should be applied in turn to the other interactions, ideally refining the model at every step (provided the different optimizations do not have contradictory effects).[4] The order in which the different potentials should be optimized is therefore, in our case, the following:

$$U_{bond}(b) \rightarrow U_{ang}(b) \rightarrow U_{mol}(b) \rightarrow U_{tors}(b) \qquad (7)$$

Optimization of the binding interaction was not needed for any of the models. The procedure would have been slightly different than for the other potentials, since for this degree of freedom we used an analytic form for the potential (whereas it is tabulated for the other interactions). In that

case, an optimization of the parameters (here, b_0 and k_{bond}) using e.g. a simplex method is needed.

The angular potential is the potential that has to be optimized first, and this is a particularly important step since we know its prominent influence on the crystallization process. The method consists of modifying $U_{ang}(\theta)$ iteratively[6] via

$$U_{ang}^{n+1}(\theta) = U_{ang}^n(\theta) + k_B T \ln \frac{P_{ang}^n(\theta)}{P_{ang}^{target}(\theta)} \qquad (8)$$

with $U_{ang}^n(\theta)$ being the potential at step n, $P_{ang}^n(\theta)$ the probability distribution of angles when applying this potential in the simulation, and $P_{ang}^{target}(\theta)$ the target distribution from the all-atom simulation. The optimization procedure was observed to converge extremely fast, with most of the changes already accomplished after the first iteration, subsequent iterations only slightly refining the agreement between $P_{ang}^n(\theta)$ and $P_{ang}^{target}(\theta)$. Ten iterations were performed, yielding a satisfactory agreement; the resulting distributions at various steps during the optimization procedure are shown in Figure 9.

Since the measured torsional distribution was very close to the prescription already, it was not necessary to apply the optimization algorithm to this interaction; this is shown in Figure 10.

The same procedure as described in the case of the angular potential was applied to optimize the effective nonbonded potential, and it also converged very fast—again one iteration was sufficient to yield a very small discrepancy between $RDF^n(r)$ and $RDF^{target}(r)$ (the total RDF were used as a means of comparison between the nth and $n + $ first distributions). After 10 iterations, the RDF did not change appreciably, and the mismatch with the target RDF was small (but still nonzero, especially in the case of CG-PE$_1$ for which the nonbonded interaction has a very short range and therefore cannot strongly influence the large-r region behavior of the RDF).

After optimizing $U_{mol}(r)$, it appeared that the angular distribution has been modified as well and no longer matched the target $P_{ang}^{target}(\theta)$. This is not surprising, since the structure of the melt is necessarily strongly influenced by the angular distribution. Therefore convergence of both the angular distribution and the RDF at the same time cannot be guaranteed. However, repeating the iterative process for the angular distribution first, then for the RDF yielded satisfactory results for both distributions, in the case of CG-PE$_1$ and even better for CG-PE$_3$ (see Figure 11)—which could reproduce very accurately the RDF on a larger length scale due to the larger cutoff in $U_{mol}(r)$; the optimized nonbonded potential thus has a stronger influence on the large-scale structure and facilitates convergence to the target distribution. Again, 10 iterations were performed for each degree of freedom, while most changes occurred after the first step.

The pressures that were used to maintain the density to its value obtained from the all-atom simulation were, respectively. 2.39 and 1.21 $k_B T \sigma^{-3}$ for CG-PE$_1$ Opt. and CG-PE$_3$ Opt., which is lower than the values of the unoptimized models indicated in Table 2. This means that it is easier to
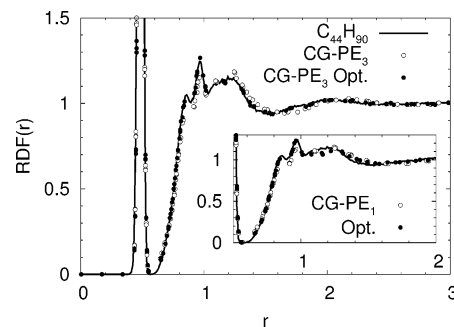

Figure 11. Comparison of the (total) RDF for the original C$_{44}$H$_{90}$ model, CG-PE$_3$, and its optimized form (after the second optimization procedure). The inset shows the same quantities in the case of CG-PE$_1$. It can be seen that the RDF is very well reproduced using optimized effective potential, even better in the case of CG-PE$_3$ which has an attractive part and therefore a longer range.
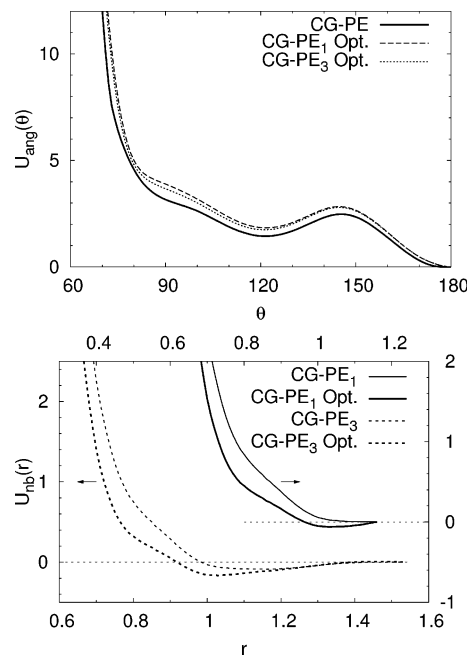

Figure 12. Optimized angular and nonbonded potentials for the CG-PE$_1$ and CG-PE$_3$ models. The original potential $U_{ang}(\theta)$ is unique, whereas $U_{mol}(r)$ depends on the CG-PE model (smaller or larger cutoff). One can notice that the optimization process for CG-PE$_1$ which is originally completely repulsive, enforces the apparition of a slightly attractive part.

constrain the target density in the case of optimized models. Figure 12 presents the final form of the angular and nonbonded potentials, compared to the original CG-PE model, i.e., before and after optimization. As a conclusion, optimization of the coarse-grained potentials of the CG-PE models appeared quite successful in improving the ability to reproduce the static structure of the original all-atom model, as can be seen from the values of the squared radius of gyration R_g^2 which are 2.59 for CG-PE$_1$ and 2.70 for CG-PE$_3$ and become 2.90 and 2.94, respectively, for CG-PE$_1$ Opt. and CG-PE$_3$ Opt.; this is comparable to the value 3.03 obtained for R_g^2 with the corresponding detailed simulations of C$_{44}$H$_{90}$.
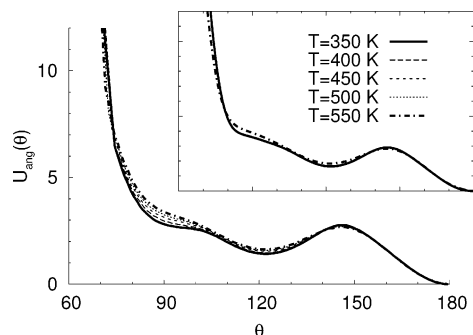
Figure 13. Comparison of the effective angular potentials $U_{ang}(\theta)$ (no optimization) inverted from the atomistic simulations data corresponding to equilibrium at several temperatures. The curves differ particularly for the gauche−gauche state at around 90° for the extreme temperatures, whereas for the optimized angular potentials (inset) they are found to be closer to each other (potentials for the models CG-PE$_1$ and CG-PE$_1^{T=350}$).

3.3. Effect of the Temperature on the Mapping. In this study of coarse-grained models for *n*-alkane melts, we make the assumption that the effective potentials describing the interactions between the beads are valid at *any* temperature, even though they are determined from the atomistic simulations *at one temperature only*. This is certainly justified in case we only consider the melt at this particular temperature or for small variations around it but is obviously wrong when we use the simple coarse-grained model over a large range of temperatures, as for a cooling toward the crystalline state of the polymer system, as presented in section 4. Therefore we will consider hereafter the effect of temperature changes on the parameters in the models and check whether it can be valid to make use of such a simplistic approach.

To check the validity of our simplified approach, we can have a look at the different effective potentials and monitor their behavior as the temperature is modified. The binding potential $U_{bond}(b)$ is a harmonic potential [see eq 4] whose parameters were determined by a fit: One can calculate the parameters b_0 and k_{bond} at other temperatures as well. Doing so between $T = 350$ K and $T = 550$ K, it was found that the parameters are temperature independent, i.e., that the change in the bond length distribution observed for two different temperatures T_1 and T_2 is completely determined by the ratio T_1/T_2. This comes from the independence of the bond length from the other parameters in the system and is not surprising when considering that, in this case, the harmonic interaction on the coarse-grained level is directly related to the harmonic interaction constraining the bond angles on the atomistic level. Thus the effective binding potential is a true potential energy for our models.

The angular potential is determined via the Boltzmann inversion of an angular distribution. Figure 13 shows the comparison of the effective potentials that are derived directly from the atomistic simulations, without optimization, for various temperatures. It is observed that in this case the potentials are different; there is a temperature dependence of the angular potential. To obtain another comparison of the angular potentials and the effect of temperature, we proceeded to the derivation of coarse-grained effective

potentials using atomistic simulation data at a lower temperature, i.e., $T = 350$ K (corresponding to $T = 0.64$ for coarse-grained simulations). The method used was exactly the same as presented in sections 3.1 and 3.2: After the first derivation, the effective potentials were optimized to yield a model CG-PE$_1^{T=350}$. This model again significantly improves the reproduction of the structure of the atomistic model, with for instance a squared radius of gyration $R_g^2(T = 0.64) \simeq 3.13$ instead of 3.62 before optimization; this is to be compared to the value measured in the atomistic simulation, $R_g^2(T = 350K) \simeq 3.19$. The model CG-PE$_1^{T=350}$ is the equivalent of the CG-PE$_1$ model, i.e., it has an explicit torsional potential and no attractive part. The model CG-PE$_3^{T=350}$ with an attractive part in the nonbonded potential could not be optimized nor used, since at the temperature $T = 0.64$ it already started to crystallize during the equilibration phase of the simulation.

The inset in Figure 13 compares the *optimized* effective angular potentials of CG-PE$_1$ (mapping done at $T = 550$ K) and CG-PE$_1^{T=350}$: The two functions only slightly differ, compared to the differences observed for the nonoptimized potentials. The two potentials most notably differ (by approximately 10%) at around $\theta = 90°$ and $\theta = 120°$. This shows that the optimized potentials are much closer to being potential energies than the simply derived potential which are in fact free energies and take into account a certain amount of entropy corresponding to the influence of the other degrees of freedom in the system.

The torsional potential was also computed from the atomistic simulations at different temperatures; here the changes were more drastic, since the probability distribution of the angle ϕ is qualitatively different depending on the temperature: The local minima on each side of the maximum at $\phi = 180°$ (see for instance Figure 10) are more pronounced at temperatures lower than 400 K (not shown). The coarse-grained models do not reproduce this tendency, and this discrepancy could only be treated by introducing the appropriate temperature dependence in the definition of the effective potentials. However, one can argue that this is a weak effect and that the torsional degree of freedom does not play a key role for our models. Actually, cooling simulations of CG-PE$_1$ and CG-PE$_2$ show that the two models differing by the torsional potential have essentially the same behavior and similar crystallization temperatures.

The observations for the nonbonded interaction are the same as for the angular potential: The inverted potentials are indeed rather different depending on the temperature chosen for the mapping, but when optimized the mismatch tends to disappear, as observed when comparing CG-PE$_1$ to CG-PE$_1^{T=350}$. So once again the effective potentials determined for the optimized models are a good approximation to potential energies.

Thus, it is possible as a first approximation to use coarse-grained potentials mapped at one single temperature from a reference (here an all-atom simulation) but only when using *optimized* models. For the latter, the correlations that might exist between the different degrees of freedom are taken into account during the optimization procedure, and the effective potentials therefore correspond more closely to the ideal

potential energies. For our CG-PE models, it could be checked that the effective potentials are not too different at different temperatures (above T_{cryst} of course), which means that they are a good approximation to the "real" potentials. However, this is still not a sufficient condition for the coarse-grained models to reproduce the phase transition to the crystalline state at low temperature; this is what is considered in the next section.

4. Crystallization Simulations

Crystallization of the different models of PE could be observed in nonequilibrium simulations consisting in a continuous cooling, i.e., a steady decrease of temperature imposed by the thermostating algorithm. Continuous cooling has been performed at a constant rate of $-5 \times 10^{-6} \, \tau^{-1}$ and was followed by continuous heating (rate $2 \times 10^{-5} \, \tau^{-1}$).

4.1. Time Scales. The use of identical cooling and heating rates for the nonequilibrium simulations of the different models is justified by the similarity of the time scales that can be derived for CG-PE: The characteristic time $\tau_R = R_g^2/D$ takes the value $372 \, \tau$, $\pm 54 \, \tau$ depending on the model. This allows us to estimate the time unit τ in the coarse-grained simulations: Comparing $\tau_R = 870$ ps for $C_{44}H_{90}$ to $\tau_R = 372 \, \tau$, one obtains $\tau \approx 2.3$ ps. It is possible with this figure to compare more quantitatively the "efficiency" of the two simulation schemes: The all-atom simulations' integration time-step is 0.002 ps, whereas it is $0.005 \, \tau \approx 0.01$ ps for the coarse-grained method.

This estimation implies that during cooling at the rate used here, the temperature changes by about 10^{-3} during the chain's relaxation time τ_R, meaning that the system is not so far from the equilibrium, at least at the beginning of the cooling run. At $T = 0.6$, the relaxation time is increased by almost a factor 20, i.e., the chain relaxes during cooling of 0.02. This is still very rapid and means that the slowing down should not be an important driving force of the crystallization.

Upon optimization of the potentials, D is observed to increase slightly. The behavior of the diffusion coefficient with temperature is examined in Figure 14; it is seen that data for the coarse-grained models exhibit the same trend typical of an Arrhenius law, $D \sim \exp(-U/k_B T)$. Using the length and time scales σ and τ, it is possible to compare the coarse-grained and all-atom data and this shows that both simulations yield a similar behavior, at least for high temperatures. This also corroborates the statement that the systems are not far from equilibrium during cooling. The decrease of D with $1/T$ is faster for the all-atom simulations than for the coarse-grained models; this points out the limitations of the simple coarse-graining procedure that does not allow for reproducing the dynamics of the system on a large temperature scale (note that the deviation has the same trend as the deviation of the thermal expansion coefficient shown in Figure 17).

4.2. Crystal Structure. Figure 15 shows the results for cooling of the models with nonoptimized and optimized potentials; they crystallize and melt at the temperatures indicated in Table 3 (melting and crystallization temperatures are determined as the inflection point in the behavior of the volume as a function of temperature). The experimental value
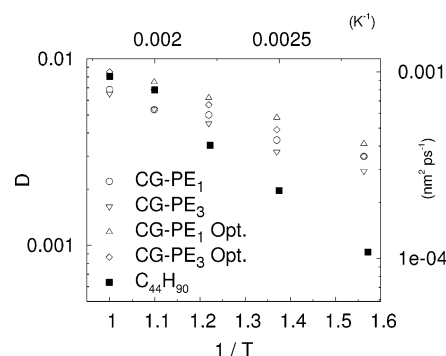
Figure 14. Evolution of the diffusion coefficient as a function of the inverse temperature. The "coarse-grained" units are used; the values for $C_{44}H_{90}$ have been rescaled by a factor σ^2/τ with $\tau = 2.3$ ps as derived from the mapping of the relaxation times τ_R, i.e., $\tau \approx 2.3$ ps. The points corresponding to the coarse-grained models show very similar behavior, whereas the data for $C_{44}H_{90}$ depart from it at temperatures lower than 500 K. This is to be connected to the mismatch in the expansion coefficient shown by Figure 17. (The top and right-hand axes show for reference the corresponding values in units appropriate to all-atom simulations, i.e., K^{-1} and nm^2 ps^{-1}.)
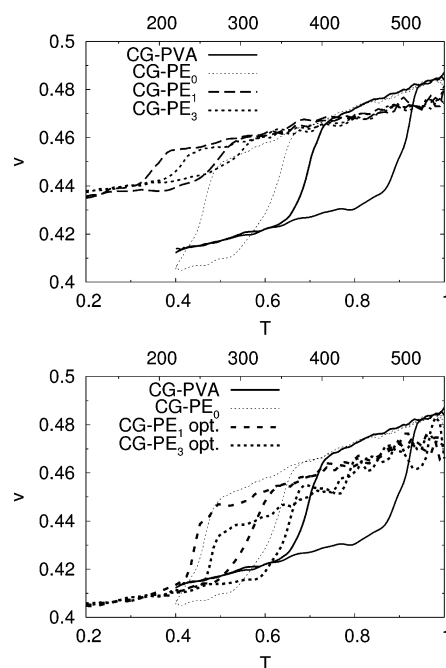
Figure 15. Phase diagrams for the different coarse-grained models and for CG-PVA: Evolution of the volume per monomer as a function of temperature during continuous cooling and subsequent reheating. The upper figure shows the results for the nonoptimized models, and the phase diagrams for the models with optimized effective potentials are presented below. It can be seen that the model which matches the high-temperature structure of the atomistic model the best (CG-PE$_3$ Opt.) also yields the melting point that is closest to the experimental data for $C_{44}H_{90}$ ($T_{melt} \approx 360$ K). (The simulation data have been smoothed.)

for the melting temperature of $C_{44}H_{90}$ is $T_{melt} \approx 360$ K, which corresponds to $T = 0.65$ in the units relevant to our models. It has to be pointed out that the crystalline structures formed

Coarse Graining of Short Polythylene Chains

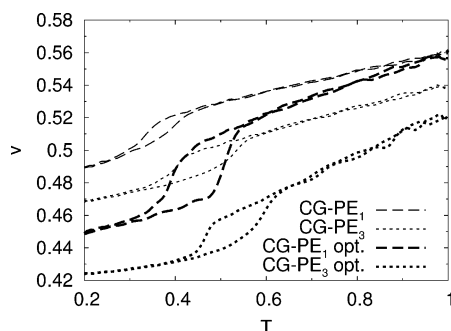*J. Chem. Theory Comput., Vol. 2, No. 3, 2006* **627**

Figure 16. Phase diagrams obtained for chains of 50 monomers simulated with the different coarse-grained models for PE. The hierarchy of the different coarse-grained models is unchanged: The optimized $CG-PE_3$ has the higher melting point, and $CG-PE_1$ the lowest. It is also worth noticing that, for this longer chain length, the densities measured for the different models do not match anymore, as was the case for the system used for the mapping of the potentials ($N = 22$, density 2.1 at $T = 1$). The phase diagram obtained for $N = 100$ is extremely similar to what is observed here.
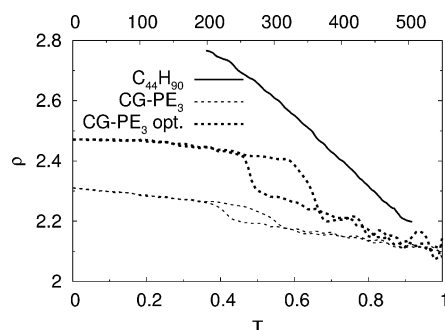

Figure 17. Density of the melt vs temperature measured during continuous cooling for the original atomistic model and CG-PE. The expansion coefficients for the latter are much smaller than for $C_{44}H_{90}$.

Table 3: Crystallization and Melting Temperatures Measured during Continuous Cooling and Heating of the CG-PE Models

model	$N = 22$		$N = 50$		$N = 100$	
	T_{cryst}	T_{melt}	T_{cryst}	T_{melt}	T_{cryst}	T_{melt}
$CG-PE_1$	0.35	0.49	0.33	0.39	0.33	0.42
$CG-PE_2$	0.40	0.52				
$CG-PE_3$	0.41	0.54	0.40	0.53	0.40	0.52
$CG-PE_1$ Opt.	0.44	0.56	0.39	0.51	0.40	0.49
$CG-PE_3$ Opt.	0.48	0.65	0.47	0.59	0.47	0.62

chains for higher molecular weight molecules, as we shall see later on.

The temperatures of crystallization/melting are correlated to the probability of finding stretched states in the melt *at high temperature* ($T = 1$), showing the prominent influence of the bond angle degree of freedom on the properties of the model and particularly its low-temperature behavior.[29] The nonoptimized models which have a too low probability of trans−trans states compared to the original atomistic model yield lower crystallization and melting points, since chain rigidity favors formation of the crystalline phase and also prevents it from melting at lower temperature. The pronounced hysteresis found not only relates to the large rates for cooling and heating but also depends on the activation barrier to be overcome when melting the crystalline phase.

Since the results obtained from the simulations of the coarse-grained models with $N = 22$ (corresponding to $C_{44}H_{90}$) appeared reasonable, it was interesting to extend the simulation to longer chains, as *n*-alkanes are just short PE molecules. The phase diagrams obtained for $N = 50$ are shown in Figure 16, and Table 3 summarizes the crystallization and melting temperatures also for $N = 100$. Simulations were performed with 72 chains for $N = 50$ and 192 chains for $N = 100$; the cooling and heating rates applied are the same as in the case of $N = 22$.

Note that the typical relaxation times increase as N^2 in this length region. Thus the temperature difference corresponding to the relaxation time becomes significant for the longest chains, and they may be already out of equilibrium. The different models crystallize and melt at about the same temperatures as $N = 22$, whereas one might have expected that the crystallization temperature increases with chain length. The phase diagrams appear very similar for the longer chains $N \geq 50$, indicating a fast saturation of T_{cryst} and T_{melt}. It is to be noted that the densities of the melt at $T = 1$ are not identical anymore depending on the model; the density was a constraint in the derivation of the coarse-grained models, as the value of the pressure has been adjusted so as to reproduce the density of the all-atom simulation.

The density varies linearly with the inverse chain length $1/N$ due to chain-end effects (this has been checked for the CG-PVA model, with chains from $N = 10$ to $N = 1000$). However, the proportionality coefficient necessarily depends on the force field and therefore need not be the same for different models. Another constraint could have been introduced during the coarse-graining procedure to enforce the density variation as a function of the chain length, or temperature.

The morphology of the crystals depends on chain length and on the force field used. For short chains, the melt turns into a crystal of rods as the chains can still fully stretch upon cooling, whereas for longer chains the qualitative behavior of polymer systems is recovered: The chains are constrained, and parts aggregate during crystallization before the whole chain has had the possibility to stretch in order to reach the all-trans state. This gives rise to ordered lamellar regions and amorphous zones surrounding them. This is observed in our simulations of CG-PE as well as for CG-PVA;[20] the crossover between the rod and the polymer regime is located

during the cooling simulations of coarse-grained models cannot reproduce all features of *n*-alkane systems: The crystals have a hexagonal symmetry due to the cylindrical shape of the chains' stretched parts. This is qualitatively different from the orthorhombic (or triclinic, depending on the chain length[26]) phase that is characteristic of *n*-alkanes and for the description of which the explicit treatment of the hydrogen atoms is mandatory. However, the larger scale features of the crystals, i.e., the lamellar-like structure, is well reproduced, and so is the apparition of folds along the

around $N = 20$ at such cooling rates. For the nonoptimized (and therefore not stiff enough) models most chains are folded once in the crystal (for $N = 22$), and for the optimized models all-trans states are obtained, with chains slightly tilted as a result of the small size of the simulation box.

For $N = 50$, the CG-PE$_1$ Opt. model still has fully elongated chains, CG-PE$_3$ Opt. yields one well-ordered lamella surrounded by amorphous chain parts, and the nonoptimized models form more disordered structures. The models having an attractive nonbonded potential prove to be significantly denser at low temperature. The structure of the crystals obtained for the different models can be characterized by the proportion of trans−trans states at low temperature, which is above 90% for $N = 22$ (92% for the nonoptimized models and 98% for the stiffer optimized models) and slightly smaller for $N = 50$ as a result of more disorder in the final configurations.

5. Conclusions

We have constructed coarse-grained models of PE in the same spirit as the CG-PVA model, this means on a "united-monomer" level with two CH$_2$ groups per CG-bead, and including at least an angular potential beyond the bead-spring interactions. It has been shown that such models yield good results in the description of the melt in the sense that *optimized* coarse-grained models successfully reproduce the structure obtained through the more detailed (all-atom) simulations of C$_{44}$H$_{90}$. This simplification scheme yields a reasonable agreement with the structural properties of the original model in equilibrium at high temperature (it was optimized to this end). It also allowed us to reproduce qualitatively the features of the crystals that develop at low temperature, whereas this could not be achieved using the all-atom simulations. Furthermore, the refinement of the coarse-grained models via optimization of the angular and nonbonded potentials made it possible to approach the experimental value of the melting point. Even though the agreement with this value does not necessarily mean that the model describes completely the system under consideration, it is still very interesting that with a rather simple method one is able to obtain an effective model which achieves very good qualitative agreement with experiment.

The results of the present study show that the coarse-graining procedure used to address the crystallization of PVA can be applied to the even more generic PE and yield both the same qualitative results for the crystalline state (which is interesting because of the abundant literature about crystallization of PE) and a simplified model that reproduces accurately many static properties of the melt at high temperature.

One of the main reasons for the success of our coarse-graining approach lies in the chosen "scale" at which the mapping is performed, i.e., the position of the mapping centers. The choice to retain every other carbon atom from the original backbone is indeed crucial, since this leads to a shift of the different interactions (the binding potential is related to the atomistic bond angles, the angular potential to the torsions, and the torsional potential has very little influence): Thus the global structure of the model is preserved, while the unimportant degrees of freedom are averaged out without losing the connection to the underlying all-atom description. The angular potential in its structure is reminiscent of the torsional states, and this connection captures the most important features of the chains we are modeling. Simulations of CG-PVA-like models with a structureless angular distribution such as the ones obtained for coarser models (see Figure 4) only led to the glassy state upon continuous cooling since nothing drives the crystallization anymore in this case. To be consistent when deriving a coarse-grained model the way we described it here in the case of more remote mapping centers, one would need to take into account the bond length distribution which is not Gaussian anymore but presents several peaks. However, this is not likely to lead to the formation of an ordered state at low temperature since nothing would then force the chains to stretch.

The process of coarse-graining consists of discarding part of the information available from a system; by doing so one makes the simulation more efficient but also introduces possible inconsistencies in the new models. Even though the high-temperature structure has been shown to match rather precisely the properties of the original all-atom model, the continuous cooling simulations revealed that the expansion coefficient of the CG-PE models do not correspond to the values measured from the atomistic simulations: Figure 17 compares the particle density for both simulation schemes, indicating the mismatch. The value obtained for C$_{44}$H$_{90}$ (all-atom data yield a result close to 9×10^{-4} K^{-1} between 350 and 450 K) is rather close to the experimental data (6×10^{-4} K^{-1} is reported in the melt[30]), while the coefficient measured for the different CG-PE models is much smaller; this also explains why crystallization is easier to observe using the coarse-grained than the atomistic models. To improve the model, one certainly has to better reproduce the thermodynamics of the model, in particular the compressibility and the thermal expansion coefficient. This was also found to be a prerequisite to successfully reintroduce atomistic details into the simulation.[31]

This work shows that except for extreme cases, the optimization of the coarse-grained potentials is both needed and very efficient (optimization was found unnecessary e.g. for very flat potentials as in the case of torsions for CG-PE or for very sharp potentials such as the angular potential of CG-PVA). Concerning the other choices for the force-field parameters, a compromise between the speed and accuracy of the simulations has to be adopted. The CG-PVA model was derived with the idea of providing a very efficient model for polymer melt simulations, i.e., with short-range potentials, tabulated values with linear interpolation. The torsional potential was discarded because of its weak influence on the structure of the melt. In the case of CG-PE, it was found that taking into account a torsional potential (which has a somewhat more pronounced structure) slightly improves the accuracy of the static properties without influencing notably the performance of the simulation code. This is the reason we focused on the models *with* explicit torsional potential here (CG-PE$_1$ and CG-PE$_3$). The form of the nonbonded interaction is a more complicated issue. The choice of a

Coarse Graining of Short Polythylene Chains

J. Chem. Theory Comput., Vol. 2, No. 3, 2006 **629**

smaller cutoff distance (CG-PE$_1$) obviously makes the simulation faster, but since the model with a larger cutoff (CG-PE$_3$) achieves a higher accuracy in reproducing the structure of the melt it cannot be deemed "less efficient". And even though the (optimized) CG-PE$_3$ takes roughly twice as much CPU time as CG-PE$_1$, it remains much faster than the all-atom scheme.

References

(1) Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Müller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V. *Adv. Polymer Sci.* **2000**, *152*, 41−156 review.

(2) Müller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754 review.

(3) Müller-Plathe, F. *Soft Mater.* **2003**, *1*, 1−31 review.

(4) Reith, D.; Meyer, H.; Müller-Plathe, F. *Macromolecules* **2001**, *34*, 2335.

(5) Meyer, H.; Biermann, O.; Faller, R.; Reith, D.; Müller-Plathe, F. *J. Chem. Phys.* **2000**, *113*, 6264−75.

(6) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624−36.

(7) Faller, R.; Reith, D. *Macromolecules* **2003**, *36*, 5406.

(8) Sun, Q.; Faller, R. *Comput. Chem. Eng.* **2005**, *29*, 2380.

(9) Bourasseau, E.; Haboudou, M.; Boutin, A.; Fuchs, A.; Ungerer, P. *J. Chem. Phys.* **2003**, *118*, 3020−34.

(10) Milano, G.; Müller-Plathe, F. *J. Phys. Chem. B* **2005**, *109*, 18609−19.

(11) Ashbaugh, H. S.; Patel, H. A.; Kumar, S. K.; Garde, S. *J. Chem. Phys.* **2005**, *122*, 104908.

(12) Yelash, L.; Müller, M.; Paul, W.; Binder, K. *J. Chem. Theory Comput.* **2006**, *2*, 588−597.

(13) Guerrault, X.; Rousseau, B.; Farago, J. *J. Chem. Phys.* **2004**, *121*, 6538−46.

(14) Padding, J. T.; Briels, W. J. *J. Chem. Phys.* **2002**, *117*, 925−943.

(15) Strobl, G. *Eur. Phys. J. E* **2000**, *3*, 165.

(16) Lotz, B. *Eur. Phys. J. E* **2000**, *3*, 185.

(17) Muthukumar, M. *Adv. Chem. Phys.* **2004**, *128*, 1−63.

(18) Ungar, G.; Zeng, X.-B. *Chem. Rev.* **2001**, 4157.

(19) Meyer, H.; Müller-Plathe, F. *J. Chem. Phys.* **2001**, *115*, 7807.

(20) Meyer, H.; Müller-Plathe, F. *Macromolecules* **2002**, *35*, 1241.

(21) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(22) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. R.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.

(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(24) Müller-Plathe, F. *Comput. Phys. Comm.* **1993**, *78*, 77.

(25) Yoon, D. Y.; Smith, G. D.; Matsuda, T. *J. Chem. Phys.* **1993**, *98*, 10037.

(26) Dorset, D. L. *Crystallography of the polymethylene chain. An inquiry into the structure of waxes;* Oxford University Press: Oxford, 2004.

(27) Ryckaert, J.-P.; Klein, M. L. *J. Chem. Phys.* **1986**, *85*, 1613.

(28) Meyer, H. Structure Formation and Chain-Folding in Supercooled Polymer Melts. Some Ideas from MD Simulations with a Coarse-Grained Model. In *Polymer Crystallization: Observations, Concepts and Interpretations*; Sommer, J.-U., Reiter, G., Eds.; Springer: Berlin, 2003; Vol. 606.

(29) Vettorel, T. Ph.D. Thesis, University Louis Pasteur, Strasbourg, France, 2005. http://eprints-scd-ulp.u-strasbg.fr: 8080/archive/00000334/.

(30) Van Krevelen, D. W. *Properties of Polymers*; Elsevier: Amsterdam, 1997.

(31) Praprotnik, M.; Delle Site, L.; Kremer, K. *J. Chem. Phys.* **2005**, *123*, 224106.