THE EUROPEAN
PHYSICAL JOURNAL
SPECIAL TOPICS

Regular Article

Single-chain behavior of poly(3-hexylthiophene)

Momchil Ivanov[a], Jonathan Gross[b], and Wolfhard Janke[c]

Institut für Theoretische Physik, Universität Leipzig, Postfach 100 920,
04009 Leipzig, Germany

Abstract. Poly(3-hexylthiophene) (P3HT) has been in the focus of recent studies due to its promising future use in organic photovoltaics, electronics and photonics. Recent publications investigate the melt behavior of P3HT, its interaction with other molecules, mainly various fullerene derivates, and isolated chains interacting with substrates. In this work we lay the focus on the single-chain properties of P3HT in vacuum. We compare structural properties obtained from simulations using two coarse-grained models and an atomistic model of the polymer for various chain lengths and temperatures.

1 Introduction

The shift towards "renewable" energy sources worldwide in recent years has pushed the production and development of devices harvesting the solar energy. Optoelectronic devices based on organic polymers [1] are in the current focus of research as they are a promising low-cost alternative to conventional silicon based devices. Other areas of research such as organic electronics [2–4] and photonics [5] have profited from the increased interest on organic photovoltaics as well.

Decreasing the polymer film thickness increases the influence of interfacial constraints at the contact points between the polymer film and the electric contacts. These effects determine the structure formation of the polymer networks and therewith strongly influence the electronic properties of devices [6,7].

Previous studies have investigated these issues mainly for polymer solutions [8] or using wet-chemical preparation [9]. Recently, it became experimentally feasible to eliminate solvent effects by studying polymer molecules at substrates under ultrahigh-vacuum conditions [10]. In particular, this was achieved by depositing a dilute monolayer of poly(3-hexylthiophene) (P3HT) on an ideally flat Au(001) surface [10,11]. A very high resolution image of the surface was obtained via scanning tunneling microscopy (STM). Single polymer chains were clearly resolved with that technique. Statistics of isolated polymer conformations obtained by the experiment showed

---

[a] e-mail: `momchil@itp.uni-leipzig.de`

[b] e-mail: `gross@itp.uni-leipzig.de`

[c] e-mail: `janke@itp.uni-leipzig.de`

(a) Model A   (b) Model B   (c) Model C
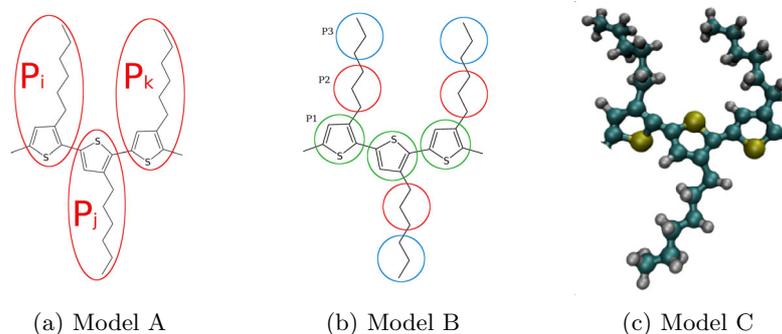
Fig. 1. (a) Two-dimensional representation of a part of a poly(3-hexylthiophene) (P3HT) molecule with overlaying coarse-grained particles P_i, P_j and P_k. The subscripts i, j and k label repeat units. (b) Two-dimensional representation of a part of a P3HT molecule with overlaying coarse-grained particles of types P1, P2 and P3 labeled in color as green, red and blue, respectively. (c) Snapshot of three repeat units of the all-atom model.

strong attraction of the polymer to the surface with the polymer chain laying flat, two-dimensional, on the substrate surface. Two distinct types of conformations were present. One, where the polymer digs into the top surface layer and remains in a rod like conformation, thus modifying the substrate. And another, where the polymer lays on top of the surface in a more or less random conformation. The statistical properties of the second type of conformations has been investigated in computer simulation studies of a coarse-grained (CG) model employing the Markov Chain Monte Carlo (MCMC) approach. A comparison with the experimental data was reported in reference [11].

In the present work we discuss the single-chain behavior of isolated P3HT molecules in vacuum utilizing computer simulations of three different existing models [12–15]. Molecular dynamics (MD) and MCMC were used to simulate an isolated P3HT molecule in the canonical ensemble. We report conformation statistics obtained at different temperatures and compare the results of the three models. This data will be useful for a future detailed study of the effects of a surface constraint on the polymer conformations. Some preliminary data has recently been given in reference [16].

2 Single-chain behavior studies

P3HT is a bottle-brush polymer. The backbone of the polymer is comprised of thiophene rings, a single hexyl side chain is attached to each thiophene ring. Regioregularity describes the distribution of the position where the side chain is connected to the backbone in comparison to the previous thiophene ring. Typical couplings are head-to-tail (HT), head-to-head (HH) or tail-to-tail (TT) (see Fig. 1 in Ref. [17]). If the complete polymer consists of a single coupling type it is considered regioregular (rr-). Polymers consisting of mixed couplings are considered regiorandom (rra-). Adachi et al. [18] studied the effect of the regioregularity on single polythiophene chain conformations. They found that regioregularity affects not only the morphologies of bulk films, but single chains as well. With changing morphologies the properties also change. Hence, it is worthwhile to investigate morphology changes in single chains. In their study, Adachi et al. examine isolated (rr-) and (rra-) P3HT chains by means of single molecule fluorescence excitation polarization spectroscopy and by crosschecking their experimental results with molecular dynamics simulations. The experiments as well as the simulations were performed at room temperature. Both approaches

allow for the same interpretation. While regioregular P3HT molecules form highly ordered structures where the thiophene rings of the backbone align and the polymer forms hairpins or loops, depending on the length of the polymer, molecules which are regiorandom observe a disturbed or twisted alignment of the backbone. These ordered structures for rr-P3HT can be accounted to aligned side chains. This effect is already significantly visible for relatively short chains of 30 repeat units. Hence, the regioregularity of a conjugated polymer has a very strong influence on the structure of the single chain and thus its physical and electronic properties.

Another important factor that plays a role in structure formation is the rigidity of the polymer backbone. Hu et al. argue in reference [19] that the crystallization and the aggregation of conjugated polymers is largely governed by the single chain morphology. Contacts between polymer backbones in the melt or crystal significantly contribute to the overall properties of the material. To this end Hu et al. study four thiophene-based polymers including the aforementioned rr-P3HT, poly(2,5-bis(3-tetradecylthiophene-2-yl)thieno[3,2-b]thiophene) (PBTTT-14), poly(2,5-bis(3-tetradecylthiophene-2-yl)-thiophene-2-ylthiazolo[5,4-d]thiazole) (PTzQT-12) and poly(3,3-didocecylquater-thiophene) (PQT-12). The aim of this study was to use single-molecule spectroscopy (SMS) to comprehend the influence of the backbone rigidness on the folding characteristics of single chains and thus the crystallization of the given polymer. They find that the stiffness increases from rr-P3HT to PQT-12, PBTTT-14, and PTzQT-12, with P3HT being the most flexible and PTzQT-12 being the most rigid. P3HT is the one polymer out of the four that exhibits the most folded structures that aggregate strongly. Even though PQT-12 has a very similar backbone and also most of the chains in the experiments folded they do not aggregate as strongly as P3HT. Compared to P3HT the polymer PBTTT-14 shows less folding due to the stiffer backbone. This is contrary to findings of their nanostructures in crystals, where PQT-12 has mostly non-folding properties and PBTTT-14 has a high likelihood to fold. Although PTzQT-12 is the most rigid polymer in this study, a significant fraction of single PTzQT-12 polymer chains fold. Hu et al. conclude that the backbone stiffness does show an influence on the folding behavior of single chains, but the backbone properties alone do not allow to conclude on the crystal structure and effects of the side chains also have to be included to make reliable predictions.

3 Investigated polymer models

A few atomistic force fields for P3HT have been published [13–15,20]. A few coarse-grained models of P3HT [12,13,21] derived from atomistic molecular dynamics simulations exist as well. These have mainly been developed to simulate blends of P3HT and PCBM (phenyl-C61-butyric acid methyl ester) or fullerene for organic photovoltaic cells. The properties of PCBM and the blend with P3HT are not a matter of this work, we rather focus on the properties of the P3HT models only. Moreover, we focus on isolated molecules at fixed temperatures in vacuum, i.e., we use the canonical ensemble.

3.1 Model A

The first model we studied is due to Lee et al. [12] and we refer to it as Model A. In this model each repeat unit of P3HT is represented by one CG particle as depicted in Figure 1a. Atomistic structure information such as regioregularity is lost in this representation. The bonded interactions of the CG particles have been parameter-ized using an all-atom MD simulation of an isolated 13–mer P3HT molecule in the

canonical ensemble at $423\,\mathrm{K}$. The nonbonded interactions of the CG particles have been parameterized by an all-atom MD simulation of P3HT molecules in what the authors call a "particle bath" at $423\,\mathrm{K}$ and $1\,\mathrm{atm}$. The total potential energy U of the polymer is given by

$$U = U_{\mathrm{bond}} + U_{\mathrm{angle}} + U_{\mathrm{torsion}} + U_{\mathrm{nonbonded}}. \tag{1}$$

The different terms are defined as

$$U_{\mathrm{bond}} = \sum_{i=1}^{n_{\mathrm{b}}} U_{\mathrm{b}}(r_{\mathrm{b},i}), \qquad U_{\mathrm{angle}} = \sum_{i=1}^{n_{\mathrm{a}}} U_{\mathrm{a}}(\theta_i), \qquad U_{\mathrm{torsion}} = \sum_{i=1}^{n_{\mathrm{t}}} U_{\mathrm{t}}(\phi_i), \tag{2}$$

$$U_{\mathrm{nonbonded}} = \sum_{i=1}^{n_{\mathrm{nb}}} U_{\mathrm{nb}}(r_{\mathrm{nb},i}), \tag{3}$$

where n_{b} is the number of bonds, n_{a} is the number of bending interactions, n_{t} is the number of torsions and n_{nb} is the number of nonbonded interactions. Respectively, $U_{\mathrm{b}}(r)$ is the potential energy of two bonded particles at distance r, $U_{\mathrm{a}}(\theta)$ is the bending potential energy of three bonded particles at an angle θ, $U_{\mathrm{t}}(\phi)$ is the torsional potential energy of four bonded particles with a torsion angle ϕ and $U_{\mathrm{nb}}(r)$ is the nonbonded potential energy of two particles at distance r. The number of repeat units is denoted by N. The different types of interactions in Model A are

$$U_{\mathrm{b}}(r) = \frac{1}{2}k_{\mathrm{b}}(r - r_0)^2, \tag{4}$$

$$U_{\mathrm{a}}(\theta) = \frac{1}{2}k_{\theta}(\theta - \theta_0)^2, \tag{5}$$

$$U_{\mathrm{t}}(\phi) = \frac{1}{2}\left[V_1(1 + \cos\phi) + V_2(1 - \cos 2\phi) + V_3(1 + \cos 3\phi)\right], \tag{6}$$

$$U_{\mathrm{nb}}(r) = 4\epsilon\left[(\sigma/r)^{12} - (\sigma/r)^6\right]. \tag{7}$$

The parameters of Model A (r_0, θ_0, $k_{\mathrm{b}}, k_{\theta}, V_1, V_2, V_3, \epsilon$ and σ) are listed in reference [12].

3.2 Model B

The second model we studied is due to Huang et al. [13] and we refer to it as Model B. In this model each repeat unit of P3HT is represented by three CG particles labelled P1, P2 and P3 as depicted in Figure 1b. This model includes more structural details, as features like the regioregularity are parameterizable via the angular potential between P1-P1-P2 particles. Again, the total potential energy of the polymer is given by equation 1. There are three particle types in Model B, however, and therefore several sets of parameters for each interaction potential are used for interactions between different types of particles. The nonbonded potential $U_{\mathrm{nb}}(r)$ (including contributions from electrostatic interactions) is given in the form of a table and the other potentials are given in the form of power series,

$$U_{\mathrm{b}}(r) = \sum_j c_{\mathrm{b},j}(r - r_0)^j, \quad U_{\mathrm{a}}(\theta) = \sum_j c_{\mathrm{a},j}(\theta - \theta_0)^j, \quad U_{\mathrm{t}}(\phi) = \sum_j c_{\mathrm{t},j}\cos^j(\phi), \tag{8}$$

where c_j are different constants for each potential type. The parameters of Model B are listed in reference [22].

3.3 Model C

The last model we investigated is an all-atom model based on the work of Raos et al. [14,15] and Huang et al. [13]. This model assumes a regioregularity of 100%, meaning all monomers are connected via the head-to-tail approach as depicted in Figure 1c. The force field used for the simulation is the OPLS-AA model [23]. This force field includes harmonic bonds between atoms, harmonic bond angles, proper as well as improper dihedral angles and a non-bonded term in the form of Lennard-Jones and Coulomb interactions. The parameters we used can be found in reference [24].

4 Observables of interest

4.1 Squared radius of gyration and squared end-to-end distance

Two widely used quantities in polymer physics are the squared radius of gyration R_{g}^2 and the squared end-to-end distance R_{e}^2. The first one describes the dimensions of a polymer chain and is the trace of the gyration tensor Q,

$$R_{\mathrm{g}}^2 = \sum_{i=1}^{3} Q_{ii} \tag{9}$$

with

$$Q_{ij} = \frac{1}{N_0} \sum_{n=1}^{N_0} \left(r_i^{(n)} - \overline{r}_i \right) \left(r_j^{(n)} - \overline{r}_j \right), \tag{10}$$

where N_0 is the total number of particles ($N_0 = N$ in Model A, $N_0 = 3N$ in Model B and $N_0 = 52 + 25(N-2)$ in Model C), $r_i^{(n)}$ ($i = x, y, z$) is the i-th Cartesian coordinate of the n-th particle and $\overline{r}_i = \sum_{n=1}^{N_0} r_i^{(n)}/N_0$ is the center of mass along the axis i. The total number of particles along the backbone of the molecule is denoted by N.

The second characteristic quantity describing the chain extension is the squared end-to-end distance

$$R_{\mathrm{e}}^2 = |\mathbf{R}_{\mathrm{e}}|^2 = |\mathbf{r}_N - \mathbf{r}_1|^2 \tag{11}$$

measured on the backbone, i.e., $\mathbf{r}_1$ is the position of the first and $\mathbf{r}_N$ of the last backbone particle. Both quantities R_{g}^2 and R_{e}^2 can be computed from computer simulations and experiments.

4.2 Persistence length

There are several estimators of a persistence length l_{p} of a polymer chain [32]. One definition uses the projection of the end-to-end vector $\mathbf{R}_{\mathrm{e}}$ onto the bond vector $\mathbf{b}_k$ of bond k along the backbone

$$l_{\mathrm{p}}(k) = \mathbf{R}_{\mathrm{e}} \cdot \mathbf{b}_k / |\mathbf{b}_k|^2 \tag{12}$$

to compute a persistence-length estimate. As $l_{\mathrm{p}}(k)$ can assume negative values for instance when the molecule bends, a small expectation $\langle l_{\mathrm{p}}(k) \rangle$ can be either a result of averaging small numbers or averaging positive and negative numbers. To discriminate between both situations we define the complementary observable

$$l_{\mathrm{p}}^*(k) = \left| \mathbf{R}_{\mathrm{e}} \cdot \mathbf{b}_k / |\mathbf{b}_k|^2 \right| \tag{13}$$

to be strictly non-negative. Now for some value of k a small expectation $\langle l_{\mathrm{p}}(k) \rangle$ has either a small or a significantly larger expectation $\langle l_{\mathrm{p}}^{*}(k) \rangle$. Another observable used to compute a persistence-length estimate uses the orientational correlation function of unit vectors $\mathbf{e}_k = \mathbf{b}_k / |\mathbf{b}_k|$ which is

$$\cos\theta(s) = \frac{1}{N-2-s} \sum_{k=1}^{N-2-s} \mathbf{e}_k \cdot \mathbf{e}_{k+s}. \tag{14}$$

In the same spirit as before, as the summands of the expectation $\langle \cos\theta(s) \rangle$ can assume negative values we define the complementary observable

$$\cos\theta^{*}(s) = \frac{1}{N-2-s} \sum_{k=1}^{N-2-s} |\mathbf{e}_k \cdot \mathbf{e}_{k+s}| \tag{15}$$

to be strictly non-negative. Now, there are a few methods to compute an estimate of the persistence length of the polymer using the functions defined here, some of which are listed and compared in reference [32]. For instance, the expectation $\langle l_{\mathrm{p}}(k) \rangle$ is assumed to have an inverted parabola form and the maximum value at the plateau is taken, this estimate we denote as $l_{\mathrm{p}}^{(1)}$,

$$l_{\mathrm{p}}^{(1)} = \max \langle l_{\mathrm{p}}(k) \rangle. \tag{16}$$

The mean value of the expectation $\langle l_{\mathrm{p}}(k) \rangle$ we denote as $l_{\mathrm{p}}^{(2)}$,

$$l_{\mathrm{p}}^{(2)} = \frac{1}{N-1} \sum_{k=1}^{N-1} \langle l_{\mathrm{p}}(k) \rangle. \tag{17}$$

Another estimate assumes the expectation $\langle \cos\theta(s) \rangle$ to decay exponentially as $\exp\left(-s/l_{\mathrm{p}}\right)$. A fitting procedure should hence produce an estimate of the persistence length. We denote one estimate in the spirit of reference [32] to be $l_{\mathrm{p}}^{(3)}$,

$$l_{\mathrm{p}}^{(3)} = \sum_{s=0}^{N-2} \langle \cos\theta(s) \rangle. \tag{18}$$

The last estimate we define here, $l_{\mathrm{p}}^{(4)}$, is based on the extrapolation method used in the software package Gromacs [25–29]. The persistence length there is defined as number of bonds where the expectation $\langle \cos\theta(s) \rangle$ reaches a value of $1/e$. This point is determined by a linear interpolation of $\ln \langle \cos\theta(s) \rangle$.

We were not interested in obtaining a single estimate of a persistence length but were more interested in the shape of the functions as they contain structural information and do not necessarily conform with the required assumptions. Moreover, we show that the expected plateau form of $\langle l_{\mathrm{p}}(k) \rangle$ does not exist at low temperatures for the models studied.

4.3 Structural information

There is more information that can be extracted from $l_{\mathrm{p}}(k)$ and $\cos\theta(s)$ besides estimates of the persistence length. Both functions bear structural information for collapsed and crystalline molecules in the form of oscillations. Zero crossings in both
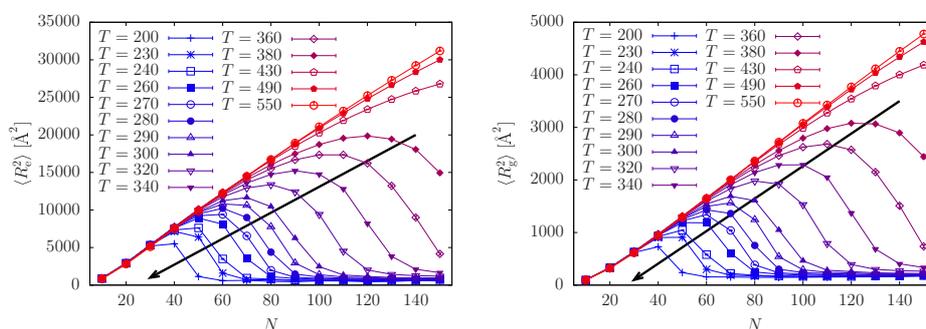
Fig. 2. Squared end-to-end distance R_e^2 and squared radius of gyration R_g^2 of an isolated molecule of Model A in bulk in the temperature range $T = 200\,\mathrm{K}\ldots550\,\mathrm{K}$ as a function of chain length N. The color gradient indicates the temperature change from cold (blue) to hot (red). The arrow points in the direction of decreasing temperature.

functions indicate folds in the chain. Due to chain symmetry $l_p(k)$ is symmetric around $(k-1)/(N-2) = 0.5$ for Model A. However, Model B and Model C represent an asymmetric molecule due to the side chain making a specific angle with the backbone, which is not $\pi/2$. This breaks the chain symmetry and the direction along the backbone which one uses to compute observables does matter.

5 Results

We present results on observables indicating the geometry of conformations for a range of temperatures $T = 200\,\mathrm{K}\ldots600\,\mathrm{K}$ for Model A and Model B. For the all-atom Model C we present results at room temperature $T = 300\,\mathrm{K}$ only and compare these with results from the other two models. The polymer chain length N we studied was in the range $N = 10\ldots150$ repeat units for Model A and $N = 10\ldots60$ repeat units for Models B and C. This was due to the huge computational requirements of the more detailed models (for concrete estimates see Ref. [16]). Model B has three times more particles per repeat unit than Model A, whereas Model C has about 25 more particles per repeat unit than Model A. The number of particle interactions is $O(N^2)$, therefore we restricted ourselves to room temperature $T = 300\,\mathrm{K}$ for the most complex model, Model C.

5.1 Isolated chains of Model A

Markov Chain Monte Carlo simulations of Model A in bulk in the canonical ensemble were conducted. An isolated polymer molecule was simulated in the temperature range $T = 200\,\mathrm{K}\ldots550\,\mathrm{K}$ for chain lengths $N = 10\ldots150$ repeat units using parallel tempering [30,31] in the temperature T. The simulation volume was chosen to be infinitely large. This choice inflicts no constraints on the conformations of the molecule. The results are plotted in Figures 2–5. Error bars are of the same size or smaller than the symbols on most plots in this work including Models B and C and are thus hardly visible.

The model shows a collapse transition for sufficiently long chains in the temperature range that was studied (see Fig. 2). Moreover, the collapse temperature T_Θ strongly depends on the chain length N.

The function $l_p(k)$ in Figure 3 shows strong dependence on the chain length N for short chains of length $N = 10$ and $N = 20$ at temperature $T = 300\,\mathrm{K}$ with a maximum in the middle of the chain. This is similar to the observations in reference [32].
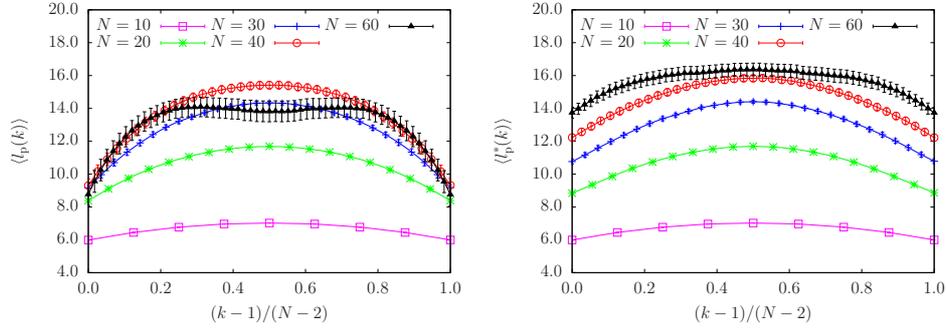
Fig. 3. Functions $l_p(k)$ and $l_p^*(k)$ of an isolated molecule of Model A in bulk at temperature $T = 300$ K for chain lengths $N = 10 \ldots 60$ repeat units.
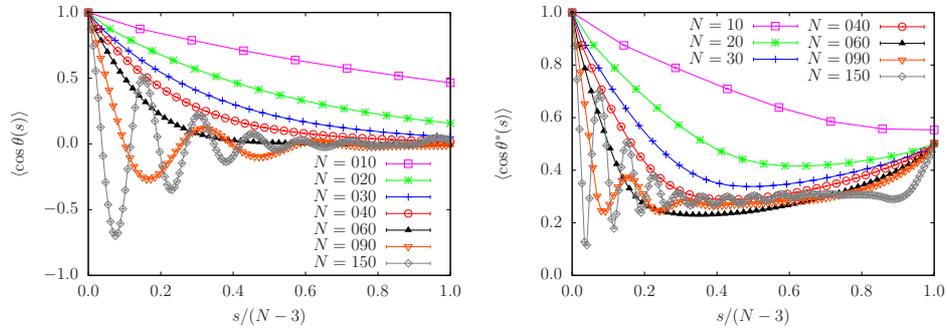
Fig. 4. Functions $\cos\theta(s)$ and $\cos\theta^*(s)$ of an isolated molecule of Model A in bulk at temperature $T = 300$ K for various chain lengths N.

However, longer chains collapse in the temperature range studied and the function $l_p(k)$ behaves differently in the collapsed phase. The expected plateau is not present for all chain lengths and temperatures. The function oscillates in the collapsed phase indicating characteristic folds in certain parts of the chain (see Fig. 5). The number of oscillations strongly depends on the chain length N. This effect is also reflected in the function $\cos\theta(s)$ shown in Figure 4. It is positive when bonds along the polymer chain are parallel, zero when they are perpendicular and negative when they become anti-parallel. An expected continuous exponential decay towards zero is not present for all molecules and temperatures. Collapsed conformations at $T = 300$ K appear for chain length $N = 90$ and more, and oscillations are visible in the $\cos\theta(s)$ function (see Fig. 4). Chain lengths $N = 60$ and less have no oscillations at this temperature. The transition between a random coil like and a structured collapsed conformation is illustrated in Figure 5 for a chain of length $N = 50$. A distinct number of folds appear at low temperatures, in particular four at $T = 200$ K for this chain length.

5.2 Isolated chains of Model B

Markov Chain Monte Carlo simulations of the Model B in bulk in the canonical ensemble were conducted. An isolated polymer molecule was simulated in the temperature range $T = 200$ K $\ldots 600$ K for chain lengths $N = 10 \ldots 60$ repeat units using parallel tempering in the temperature T. The simulation volume was again chosen to be infinitely large. The results are plotted in Figures 6–8.
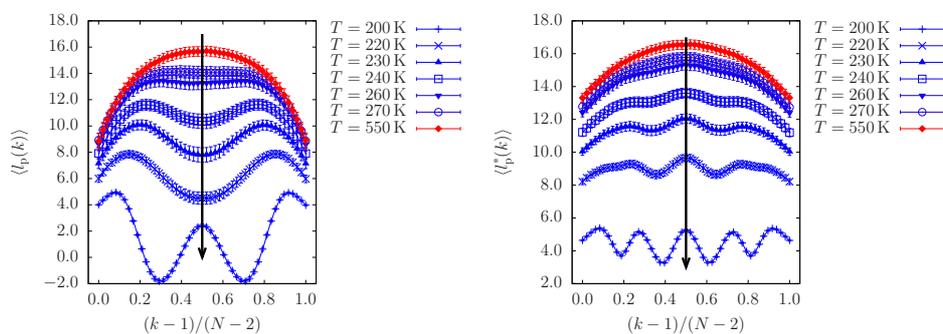
Fig. 5. Functions $l_\mathrm{p}(k)$ and l_p^* of an isolated molecule of Model A in bulk in the temperature range $T = 200\,\mathrm{K}\ldots550\,\mathrm{K}$ for chain length $N = 50$ repeat units. The color gradient indicates the temperature change from cold (blue) to hot (red). The arrow points in the direction of decreasing temperature.
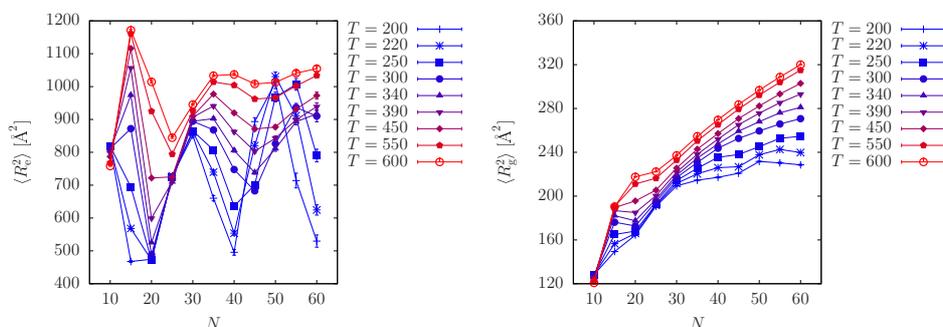
Fig. 6. Squared end-to-end distance R_e^2 and squared radius of gyration R_g^2 of an isolated polymer of Model B in bulk in the temperature range $T = 200\,\mathrm{K}\ldots600\,\mathrm{K}$ for chain lengths $N = 10\ldots60$ repeat units. The color gradient indicates the temperature change from cold (blue) to hot (red).

The model shows a rich chain-length and temperature dependence in the ranges studied. The squared end-to-end distance R_e^2 in Figure 6 for the chain length $N = 10$ repeat units indicates that the polymer has a rod like conformation at low temperatures ($T = 200\,\mathrm{K}$), whereas at high temperatures ($T = 600\,\mathrm{K}$) it assumes more random conformations. We refer to these two temperatures as low and high, respectively, throughout this section. At the next chain length $N = 15$ repeat units, the polymer is folded in a hairpin like conformation at low temperatures whereas it unfolds at high temperatures. For the chain lengths $N = 25, 30$ repeat units this effect is insignificant and sets on again for $N = 35, 40$ repeat units. However, the next distinct lengths are $N = 45, 50$ where both ends of the molecule are not close at low temperatures, however increasing the temperature brings them closer ($T = 300\,\mathrm{K}$) whereas further increasing the temperature separates them again as the molecule swells. The squared radius of gyration R_g^2 indicates a typical process of molecule swelling with increasing temperature for chain lengths of $N = 40$ repeat units and longer.

The function $l_\mathrm{p}(k)$ in Figure 7 shows a strong dependence on the chain length N at temperature $T = 300\,\mathrm{K}$. The short chain length $N = 10$ repeat units has a function with a plateau in the center of the chain. However, the chain length $N = 20$ repeat units indicates increasing number of conformations with a wide fold in the middle of the chain. A closer look at $l_\mathrm{p}^*(k)$ shows the onset of folds around the middle of the
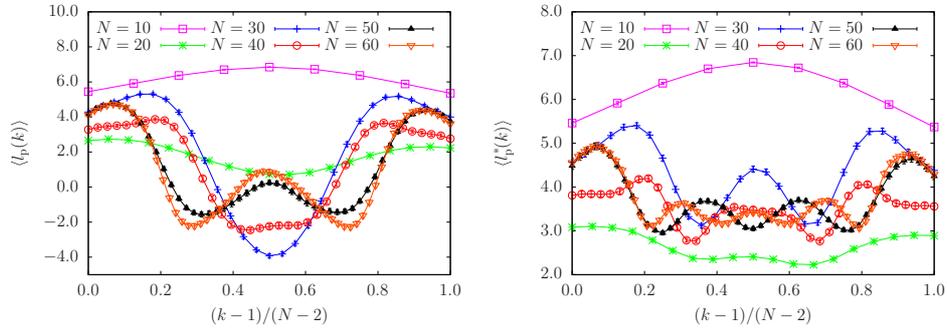
Fig. 7. Functions $l_p(k)$ and $l_p^*(k)$ of an isolated molecule of Model B in bulk at temperature $T = 300\,K$ for chain lengths $N = 10 \ldots 60$ repeat units.
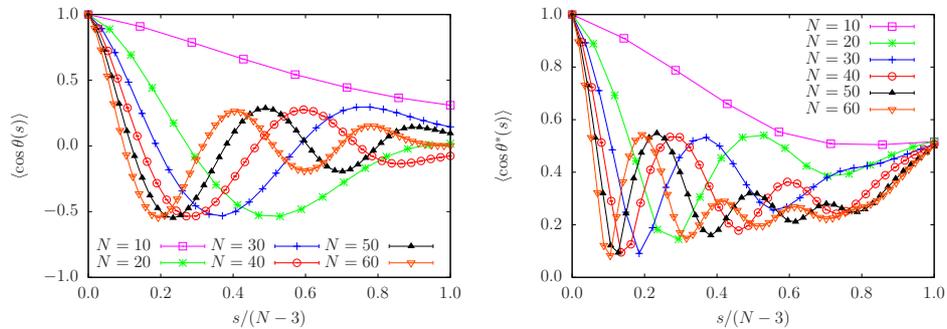
Fig. 8. Functions $\cos\theta(s)$ and $\cos\theta^*(s)$ of an isolated molecule of Model B in bulk at temperature $T = 300\,K$ for chain lengths $N = 10 \ldots 60$ repeat units.

chain as there are three local minima. Increasing the chain length to $N = 30$ repeat units gives a pronounced signal in both $l_p(k)$ and $l_p^*(k)$. The first has a negative minimum around the center of the chain indicating three distinct chain segments, two aligning parallel to the end-to-end vector $\mathbf{R}_e$ and one aligning antiparallel to it. The three distinct maxima in $l_p^*(k)$ support this. The function $\cos\theta(s)$ in Figure 8 crosses zero two times for the length $N = 30$ repeat units which also indicates three distinctly aligned segments of the polymer. Longer chains show even richer structure with more aligned segments.

5.3 Isolated chains of Model C

Molecular dynamics (MD) simulations with the software package LAMMPS [33] of Model C in bulk were conducted using the Nosé-Hoover thermostat and a timestep of 1 fs. Statistics where collected from 100 independent runs with a length of 2.5 ns each. Measurements were started after monitoring multiple observables to assure equilibrium was reached. An isolated polymer molecule was simulated at the temperature $T = 300\,K$ for chain lengths $N = 20 \ldots 60$ repeat units. The simulation volume was chosen again to be infinitely large. Lots of short simulation runs from the same initial extended polymer conformation were performed and the data presented is averaged over all runs. The results are plotted in Figures 9 and 10.

The molecules are in a compact state indicated by the small values of the squared end-to-end distance R_e^2 and the squared radius of gyration R_g^2 (see Fig. 11 in the
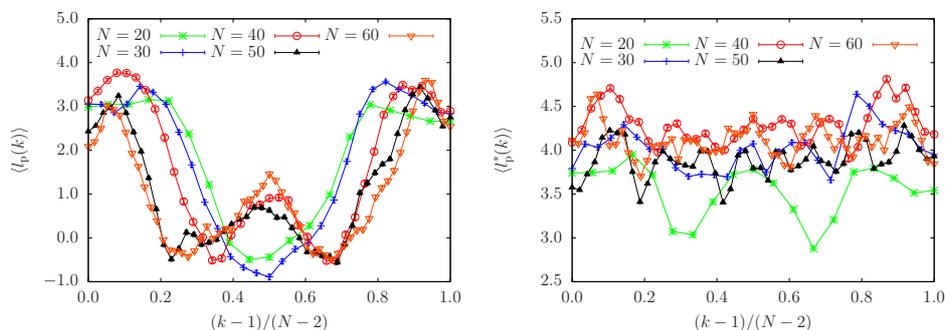
Fig. 9. Functions $l_{\mathrm{p}}(k)$ and $l_{\mathrm{p}}^{*}(k)$ of an isolated molecule of Model C in bulk at temperature $T = 300\,\mathrm{K}$ for chain lengths $N = 20 \ldots 60$ repeat units.
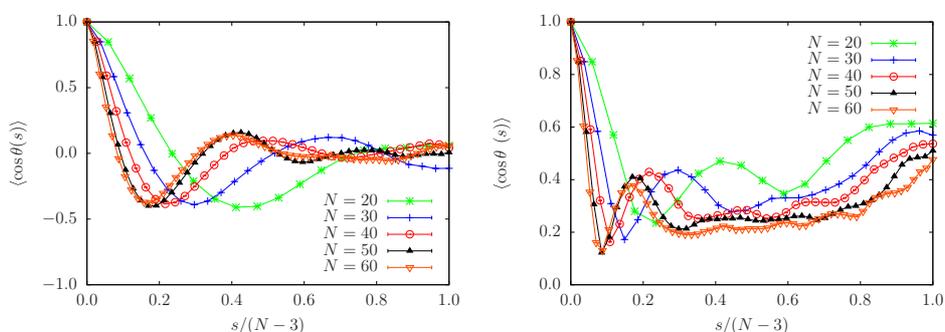
Fig. 10. Functions $\cos\theta(s)$ and $\cos\theta^{*}(s)$ of an isolated molecule of Model C in bulk at temperature $T = 300\,\mathrm{K}$ for chain lengths $N = 20 \ldots 60$ repeat units.

comparison with Models A and B). Both values do not change significantly for the studied polymer lengths.

The function $l_{\mathrm{p}}(k)$ in Figure 9 shows a clear dependence on the chain length N and has negative values for all of the studied lengths. The minima are around the middle of the chain for $N = 20, 30$ and indicate folds in that region. The longer chains have a local maximum in the middle, however, folds are present there as well. Unfortunately the statistics gathered is insufficient to produce smooth curves with small error bars, therefore a precise interpretation of the states is not yet possible. This is more pronounced in the function $l_{\mathrm{p}}^{*}(k)$. Moreover, the error bars are probably underestimated due to correlations in the individual MD trajectories. However, the data roughly indicates the types of conformations present for each length. The function $\cos\theta(s)$ in Figure 10 crosses zero for all lengths and thus supports the presence of several parallel and anti-parallel segments along the chain.

5.4 Comparison of all three models at T = 300 K

We compared the results from all three models at the temperature $T = 300\,\mathrm{K}$ and present these in Figures 11–13. The first obvious difference is between Model A and the other two by looking at the squared end-to-end distance R_{e}^{2} and the squared radius of gyration R_{g}^{2}. The first one indicates relatively stiff molecules in Model A for length up to $N = 60$ repeat units, whereas Model B and Model C describe by far more flexible molecules. In fact, rr-P3HT is known to be a semi-flexible polymer.
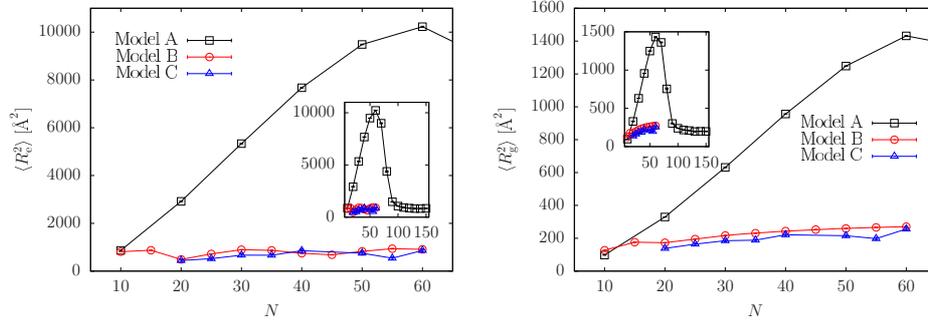
Fig. 11. Comparison of Model A, Model B and Model C. Squared end-to-end distance R_e^2 and squared radius of gyration R_g^2 of an isolated polymer in bulk at the temperature $T = 300\,\mathrm{K}$ versus chain length N. The inset contains data for chains of length up to $N = 150$ repeat units for Model A.

For chains of length $N = 100$ and longer, a Model A molecule collapses in a compact state and has similar values as a shorter molecule of the other two models. An interesting observation to note is that Models A and B agree to a great extent for the shortest chain studied, $N = 10$. This can probably be attributed to the fact that both coarse-grained models have been parameterized for short chain lengths with $O(10)$ repeat units. Our second observation is that Models B and C agree to a great extent on the order of magnitude as well as the trend in both observables.

Further comparison of the properties we did by means of the four different estimates $l_p^{(1...4)}$ of the persistence length along the polymer backbone. Those show good agreement between all three models for a chain of length $N = 10$ again, except for $l_p^{(4)}$ where the data on $N = 10$ is omitted (this chain is too short to observe the decay of correlations to $1/e$, cf. Figs. 4 and 8). These observables show again a rather stiff molecule of Model A, whereas more flexible ones for Models B and C. Moreover, a molecule of Model A of length $N = 100$ and more has the same order of persistence length as shorter molecules of Models B and C. Data on Models B and C agree to a great extent for all chain lengths.

We need to make a remark on the negative values of $l_p^{(3)}$ as negative persistence makes obviously no sense. Unfortunately, the computation of $l_p^{(3)}$ and $l_p^{(4)}$ assumes that the function $\langle \cos \theta(s) \rangle$ decays exponentially. This is, however, not the case as most of the curves we show here have negative values and even oscillate. As $l_p^{(3)}$ is just a sum over the values of $\langle \cos \theta(s) \rangle$, it can become negative. Moreover, the computation of $l_p^{(4)}$ assumes an exponential decay as well but uses an extrapolation method. Obviously, both estimates do not agree at all on the data we present, and they do not agree with the other two estimators as well.

The last remark we need to make is that both Model A and Model B have been parametrized at different conditions. It is generally believed that a coarse-grained model is valid only at the conditions it has been parametrized at. Use at other conditions, i.e., a range of temperatures, is thus not believed to guarantee reliable results. An atomistic model, i.e., Model C, is, however, generally used at a variety of temperatures, though being usually parametrized by density functional computations at $T = 0\,\mathrm{K}$ in combination with experimental results at finite temperatures and pressures. We followed this general practice of using molecular force fields at various temperatures and applied it to the other two models as well. Interestingly, Models B and C seem to be compatible at least at one typical temperature, i.e., $T = 300\,\mathrm{K}$, for all of the studied chain lengths.
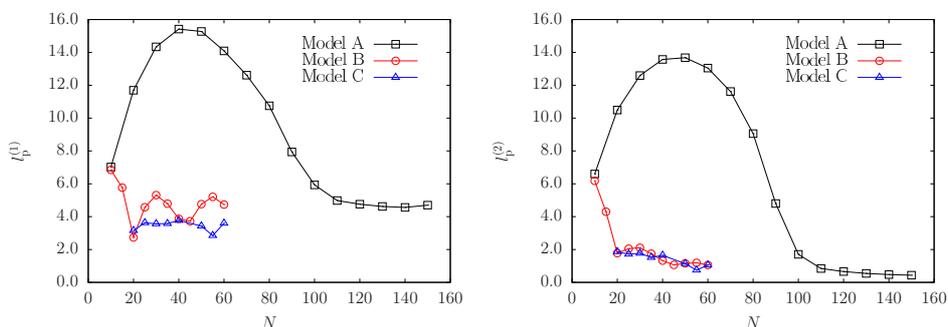
Fig. 12. Comparison of estimators $l_p^{(1)}$ and $l_p^{(2)}$ of the persistence length l_p of an isolated molecule of Model A, Model B and Model C in bulk at temperature $T = 300\,\mathrm{K}$ versus chain length N.
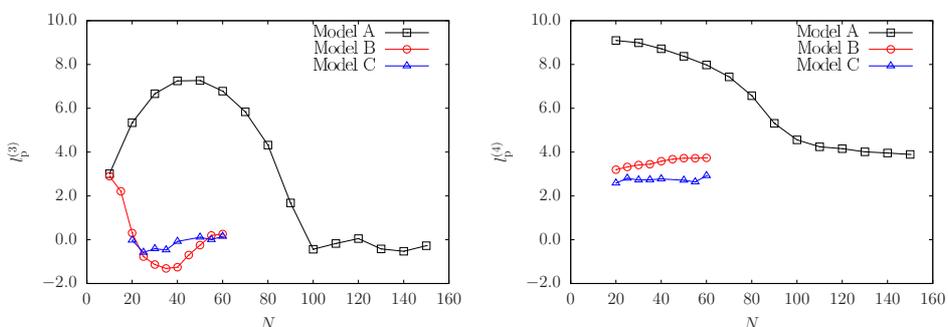
Fig. 13. Comparison of estimators $l_p^{(3)}$ and $l_p^{(4)}$ of the persistence length l_p of an isolated molecule of Model A, Model B and Model C in bulk at temperature $T = 300\,\mathrm{K}$ versus chain length N.

6 Conclusions

Mesoscopic and macroscopic morphologies in conjugated polymer melts and crystals such as in thiophene-based organic photovoltaics are greatly influenced by the properties of single chains. In this work we compare three models of P3HT, from a fully atomistic model using classical MD simulations, to a coarse-grained model which includes side-chain particles, and finally a very simplified homopolymer model in which each repeat unit is represented by one bead only. We look at structural information by comparing squared end-to-end distances, squared radii of gyration and also look at different estimators for the persistence length for all three models and various chain lengths. Both coarse-grained models were studied using Markov Chain Monte Carlo simulations. The agreement is best between the atomistic model (Model C) and the coarse-grained Model B with side-chain particles. The oversimplified Model A was found to be inadequate for our purposes, since it was mainly developed to study polymer melt structures instead of single-chain behavior. Future studies of P3HT on substrates other than the previously studied gold surface will be conducted using Models B and C, which should allow to compare simulational results to experiments, which are currently being set up by our experimental collaborators. Computer simulations provide a potent tool to systematically investigate the properties of different models for these polymer systems. Coarse-grained models in combination with Monte Carlo sampling allow for the efficient scanning of a large parameter space such as backbone stiffness or influence of the side chains.

Continuously comparing with atomistic molecular dynamics simulations and improvement of the underlying force fields enables cross checking the validity of the coarse-grained simulations and the predictions these simulations deliver. Even comparison with recent experiments of P3HT adsorption on gold surfaces is very much in the realm of possibility and encourages to further pursue the investigation of coarse-grained models for conjugated polymers.

References

1. W. Hu, editor, *Organic Optoelectronics* (Wiley-VCH, Weinheim, 2013)
2. G. Giri, E. Verploegen, S.C.B. Mannsfeld, S. Atahan-Evrenk, D.H. Kim, S.Y. Lee, H.A. Becerril, A. Aspuru-Guzik, M.F. Toney, Z. Bao, Nature **480**, 504 (2011)
3. Y.Z. Wang, Q. Wang, H.Y. Xie, L.P. Ho, D.M.F. Tan, Y.Y. Diao, W. Chen, X.N. Xie, Nanoscale **4**, 3725 (2012)
4. J.G. Labram, E.B. Domingo, N. Stingelin, D.D.C. Bradley, T.D. Anthopoulos, Adv. Funct. Mater. **21**, 356 (2011)
5. L.-W. Chong, Y.-N. Chou, Y.-L. Lee, T.-C. Wena, T.-F. Guo, Org. Electron. **10**, 1141 (2009)
6. R.J. Kline, M.D. Mcgehee, M.F. Toney, Nat. Mater. **5**, 222 (2006)
7. S. Wang, A. Kiersnowski, W. Pisula, K. Müllen, J. Am. Chem. Soc. **134**, 4015 (2012)
8. X. Ma, Y. Guo, T. Wang, Z. Su, J. Chem. Phys. **139**, 014701 (2013)
9. L. Scifo, M. Dubois, M. Brun, P. Rannou, S. Latil, A. Rubio, B. Grévin, Nano Lett. **6**, 1711 (2006)
10. S. Förster, W. Widdra, J. Chem. Phys. **141**, 054713 (2014)
11. S. Förster, E. Kohl, M. Ivanov, J. Gross, W. Widdra, W. Janke, J. Chem. Phys. **141**, 164701 (2014)
12. C.-K. Lee, C.-W. Pao, C.-W. Chu, Energy Environ. Sci. **4**, 4124 (2011)
13. D.M. Huang, R. Faller, K. Do, A.J. Moule, J. Chem. Theory Comput. **6**, 526 (2009)
14. V. Marcon, G. Raos, J. Am. Chem. Soc. **108**, 18053 (2004)
15. V. Marcon, G. Raos, J. Am. Chem. Soc. **128**, 1408 (2006)
16. J. Gross, M. Ivanov, W. Janke, J. Phys.: Conf. Ser. **750**, 012009 (2016)
17. A. Marrocci, D. Lanari, A. Facchetti, L. Vaccaro, Energy Environ. Sci. **5**, 8457 (2012)
18. T. Adachi, J. Brazard, R.J. Ono, B. Hanson, M.C. Traub, Z.-Q. Wu, Z. Li, J.C. Bolinger, V. Ganesan, C.W. Bielawski, D.A. Vanden Bout, P.F. Barbara, J. Phys. Chem. Lett. **2**, 1400 (2011)
19. Z. Hu, J. Liu, L. Simon-Bower, L. Zhai, A.J. Gesquiere, J. Phys. Chem. B. **117**, 4461 (2013)
20. T.T. To, S. Adams, Nanosci. Nanotechnol. Lett. **4**, 703 (2012)
21. T.T. To, S. Adams, Phys. Chem. Chem. Phys. **16**, 4653 (2014)
22. D.M. Huang, R. Faller, K. Do, A.J. Moule, Supporting Information, J. Chem. Theory Comput. **6**, 526 (2009)
23. W.L. Jorgensen, D.S. Maxwell, J. Tirando-Rivers, J. Am. Chem. Soc. **118**, 11225 (1996)
24. K.N. Schwarz, T.W. Kee, D.M. Huang, Nanoscale **5**, 2017 (2013)
25. H.J.C. Berendsen, D. van der Spoel, R. van Drunen, Comp. Phys. Comm. **91**, 43 (1995)
26. E. Lindahl, B. Hess, D. van der Spoel, J. Mol. Mod. **7**, 306 (2001)
27. D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, J. Comp. Chem. **26**, 1701 (2005)
28. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008)

29. S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, Bioinformatics **29**, 845 (2013)
30. K. Hukushima, K. Nemoto, J. Phys. Soc. Jpn **65**, 1604 (1996)
31. C.J. Geyer, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, American Statistical Association (New York, 1991), p. 156
32. H.-P. Hsu, W. Paul, S. Rathgeber, K. Binder, Macromolecules **43**, 1592 (2010)
33. S. Plimpton, J. Comp. Phys. **117**, 1 (1995)