

Coarse-grained models for macromolecular systems

V. Hugouvieux

INRA, UMR1083 SPO, 34060 Montpellier Cedex 1, France

Abstract. Neutron scattering experiments and simulations are often used as complementary tools in view of revealing the structure and dynamics of molecular and macromolecular systems. For polymeric and self-assembling systems, the simulation of large-scale structures and long-time processes is often achieved by using coarse-grained models which allow to gain some orders of magnitude in space and time scales. By discarding some details of the chains, they also allow a better understanding of the main features of the system that govern its behaviour, such as the sequence of a macromolecule or some interaction that leads to its self-assembly. After a brief introduction on coarse-grained models and associated representations, the approach is illustrated with the case of amphiphilic regularly alternating multiblock copolymers in dilute and semidilute solutions. In this context, a generic HP (H: hydrophobic ; P: hydrophilic, polar) model is used together with a lattice representation of the system and a Monte Carlo algorithm. The simulations give access to the various structures and phases of the system as a function of the energy of interaction between H monomers, the ratio of H monomers in the chain, the length of the blocks and the concentration. In dilute solution structures range from swollen coils in good solvent to chains of micelles and layered structures in poor solvent. In semidilute solution microphase separation and gelation are observed as a function of substitution ratio and concentration.

1. WHY USE COARSE-GRAINED OR MESOSCOPIC MODELS?

One of the main reasons why scientists use coarse-grained or mesoscopic descriptions of molecular systems is to achieve longer simulation times and larger spatial scales by discarding part of the chemical details of the system. In the case of polymers and self-assembling systems this scale issue is fundamental as the behaviour of macromolecules covers several orders of magnitude in space and time. Several review articles deal with the application of coarse-grained models to soft materials [1], proteins [2] and membranes [3].

For standard condensed matter systems such as simple liquids it is often sufficient to perform simulations of some 10^3 atoms interacting through realistic forces. Even with such small systems the structural features and correlations can be captured: as shown by the oscillations of their pair correlation function, these systems are usually homogeneous on scales of the order of 10 \AA . Quantum chemistry simulations are moreover available for these systems and enable relevant force fields on the atomic scale to be derived.

On the polymer side [4], the situation is quite different with relevant length scales encompassing several orders of magnitude. At the level of a single chain of some 10^3 monomers, bond lengths are of the order of 1 \AA , persistence length is some 10 \AA and the radius of gyration of the coil is about 100 \AA . For collective phenomena, cross-linked networks and phase separation typical length scales may even be larger and reach 10^3 \AA . Also the size of the simulation box has to be larger than the characteristic length scales of the system, which leads to very large simulated systems (10^6 atoms or more). Hence the atomic scale simulation of polymer systems is restricted to small values of the degree of polymerization N_p and systems far from any phase transitions.

The need for coarse-grained representations is emphasized by the dispersion of the time scales pertaining to macromolecular systems. While the characteristic time for a C-C bond length or angle vibration is around 10^{-13} s and the average time between two reorientational jumps in the torsional potential is around 10^{-11} s, the time for the relaxation of the coil configuration of a single chain is many orders of magnitude larger. This relaxation time τ_{N_p} scales with the chain length as $\tau_{N_p} \propto N_p^2$ for $N_p < 100$ (Rouse model) [5] while for longer chains the scaling becomes $\tau_{N_p} \propto N_p^3$ (reptation model) [5, 6]. This typically leads to relaxation times of the order of 10^{-8} to 10^{-5} s, to be compared with times for bond length vibrations or local reorientations.

To cope with the broad range of time and length scales, coarse-grained or mesoscopic representations of polymer or self-assembling systems (such as membranes) have to be envisaged. Such models are obtained by lumping groups of atoms together into single effective particles or interaction sites, thus reducing the number of degrees of freedom of the system and removing the faster vibrations. Also simplified and softer potentials are used for the interactions between these particles (electrostatics and torsions are often omitted). This leads to an appreciable speed-up of the simulations and gives access to collective phenomena such as the self-assembly of block copolymers in solution [7, 8] or the formation and thermodynamic stability of membranes [3], usually not available from atomistic simulations.

Obviously coarse-grained models are not able to capture all the atomic scale properties of the system under study. This is precisely the point of using such a representation: it enables the main features influencing the mesoscale behaviour to be elucidated. Therefore two main kinds of questions can be addressed through mesoscopic modelling and simulation. On the one hand insight can be gained into the interactions that are responsible for the mesoscale behaviour (phase separation, self-assembly). On the other hand the universal behaviour at a mesoscopic scale can be revealed (e.g. phase transitions in diblock copolymer melts, transitions between different morphologies of copolymers in solution).

2. HOW TO DESIGN A MESOSCALE MODEL?

When building a mesoscale model, one of the main issues is to determine which degrees of freedom or interactions of the system have to be retained in order to recover the behaviour of interest. Two main types of coarse-grained models can be distinguished: the minimal or generic ones, and those derived by systematic coarse-graining. A *generic polymer model* is able to capture the features common to a class of polymers with a given architecture and composition pattern. In contrast, two polymers with similar architectures but chemically different monomers should behave differently in a *systematically coarse-grained representation*.

Before describing coarse-grained models developed for macromolecules, let us first focus on the case of the solvent, which accounts for a substantial part of the degrees of freedom of the system. In coarse-grained simulations of polymers, explicit description of the solvent is often eliminated and replaced by effective interactions between the solute molecules. We talk about solvent-free or implicit solvent models. For instance, the radius R of a polymer chain consisting of N monomers scales like $R \propto N^{1/2}$ and the size L of the box should be of the order $L \sim R$. So the mass fraction of polymer, $N/L^3 \sim 1/L$, decreases with increasing length of the polymer. Hence discarding the explicit representation of the solvent enables substantial computer simulation time to be saved.

In the case of minimal models (also called generic or simplified models) only a very restricted set of features of molecules or macromolecules are taken into account. Besides excluded volume and connectivity constraints, some interactions are deemed essential for the behaviour of interest. These are retained in the model. This kind of models is widespread for the study of scaling properties of polymeric systems (both static and dynamic) and in the field of self-assembling amphiphiles. Among generic models, a distinction can be made between off-lattice and lattice representations.

The idea behind lattice representations is to describe the chains as self-avoiding walks on a lattice [9]. The simplest model consists of a regular cubic lattice where each effective particle of the chain sits on a lattice site. A bond connecting two beads of the chain is a link between two nearest-neighbour sites of

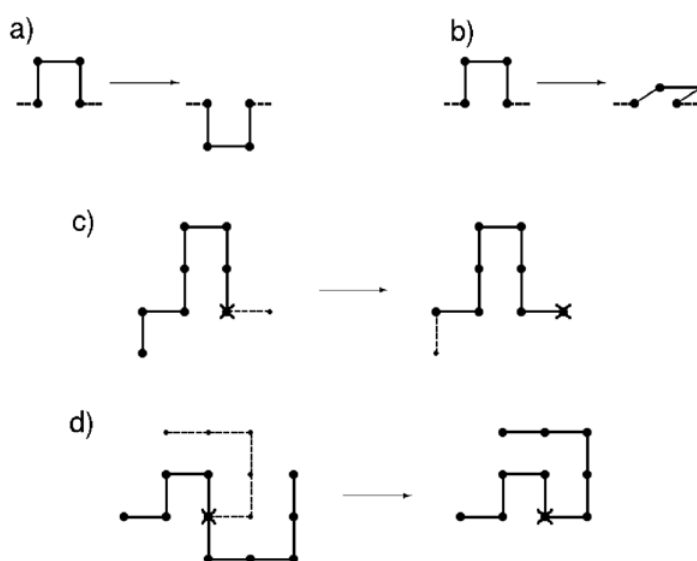


Figure 1. Some standard lattice moves: a) 180° crankshaft; b) 90° crankshaft ($d \geq 3$); c) slithering snake; d) pivot move (see [9] for other kinds of moves).

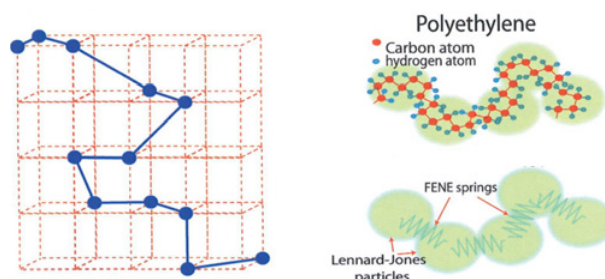


Figure 2. Schematic representation of the bond fluctuation model (*left*) and bead-spring model (*right*).

the lattice. Single occupation of the sites ensures excluded-volume constraints and non-crossing of the strands also has to be checked for. A variety of moves exist which are designed for different purposes, for instance end-bond, kink-jump, crankshaft, slithering snake and pivot moves (see Figure 1). The latter is more relevant when studying static properties, since the aim of pivot moves is to rapidly generate dynamically uncorrelated configurations of the chains.

The most widely used lattice algorithm is the bond-fluctuation model [10, 11] (see Figure 2, left). It can be seen as an intermediate between cubic lattice models and continuum models, since the bond between two monomers can take many different values (36 in 2d and 108 in 3d, compared to 4 in 2d and 6 in 3d for a simple cubic lattice). The constraints of excluded-volume and non-crossing of the strands are assured by restrictions in the allowed bond lengths.

In off-lattice representations, the positions of effective monomers are no more restricted to lattice sites. Again several models exist : freely jointed chain with rigid bonds, pearl-necklace model [4]. The bead-spring model is the most widely used off-lattice model. Bonds are represented using harmonic potentials with a constraint of finite extensibility (FENE potential) and effective particles interact through the Lennard-Jones potential (see Figure 2, right).

Such generic models have been widely used for the study of the static and dynamic properties of polymers. For instance, off-lattice models have been applied to study the phase diagrams of diblock copolymers [12] and the crossover from Rouse to reptation dynamics in polymer melts [13, 14]. Using bead-spring models, the formation of micelles has been investigated either in solution [15] or at surfaces [16]. Lattice models can also prove helpful in the study of the micellization and phase separation of surfactants and multiblock copolymers [17–19]. Protein folding and aggregation is another field of application of these models, where they are used in view of elucidating the generic mechanisms involved [20].

The other option is to derive a coarse-grained model for a specific atomic system [21]. The general principle of this method, usually called systematic coarse-graining, consists of:

1. choosing a set of key features of the system that the coarse-grained model should reproduce and
2. determining the coarse-grained interactions between the effective beads in order to reproduce the key characteristics of the system.

Two main kinds of such key features are presently used : structural and thermodynamic ones. When coarse-grained potentials are built using structural properties from all-atom molecular simulations, distributions of geometric quantities are used to compare the structure of the coarse-grained system and that of the all-atom configurations. These can be intramolecular (distances between two adjacent effective particles, bond angles between three consecutive effective beads, radius of gyration of the chain and eigenvalues of the corresponding tensor) or intermolecular (distances between the centers of mass of different chains) properties. The determination of the coarse-grained potentials is often based on Boltzmann inversion (for instance, in the case of non-bonded interaction potentials, $U^{CG}(r) = -kT \ln g(r)$ and variants thereof) of the target properties. When the target of the coarse-grained model are the thermodynamic properties, parameters of the coarse-grained potentials can be chosen so as to reproduce free energies of vaporization, hydration or partitioning, as was done in the case of the MARTINI coarse-grained lipid model [22]. A presently active field of research is to determine whether it is possible to derive coarse-grained potentials that would be consistent with both structural and thermodynamic properties of the corresponding all-atom description.

Systematic coarse-graining can for instance be used in conjunction with the method of dissipative particle dynamics [23].

Starting from atomistic descriptions and reaching coarser scales is the basic idea of multiscale simulations, where phenomena are studied at different levels of resolution. These multiscale descriptions enable a mapping to be done, from the high-resolution to the low-resolution scale, in order to reach large time and space scales. On the other hand, back-mapping can also be envisaged from the coarse-grained to the more-detailed level [24].

In the next section a case study is presented in details. It deals with the behaviour of amphiphilic multiblock copolymers in dilute and semidilute solutions. These systems are studied using a generic lattice model and a Metropolis Monte Carlo algorithm.

3. MULTIBLOCK COPOLYMERS IN SOLUTION: MODEL AND SIMULATION METHOD

In this section a more detailed account is given of the self-assembly and phase behaviour of linear multiblock copolymers in solution. The macromolecules are represented on a lattice at a coarse-grained level and a standard Monte Carlo algorithm is used for the simulations.

3.1 Motivation: the behaviour of methylcellulose

Methylcellulose is a water-soluble cellulose derivative which belongs to the class of associating polymers. It is obtained by the substitution of some hydroxyl groups of the anhydroglucose units of cellulose by the methoxide group O-CH₃ (see Figure 3). Depending on the degree of methylation of a glucose monomer, the monomer can be considered hydrophilic (no substitution) or hydrophobic

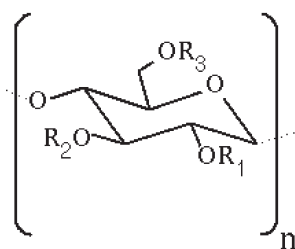


Figure 3. A generic monomer unit of methylcellulose with $R_i = \text{H}$ or $R_i = \text{CH}_3$.

(3 substituted groups). Methylcellulose is characterized by its average substitution degree which corresponds to the average number of substituted groups per monomer. The substitution pattern can vary strongly depending on the conditions for cellulose modification and it has been shown that commercial methylcellulose displays a heterogeneous distribution of the methyl groups along the cellulose backbone. In agreement with experimental results [25–27], methylcellulose is usually considered as a multiblock copolymer and not as a statistical one. The multiblock pattern gives an amphiphilic character to methylcellulose chains. Also this biopolymer derivative is known to perform thermoreversible gelation upon heating [27]. The interplay between amphiphilicity and gelation ability gives rise to a range of structures and properties which attract a lot of interest from food, biomedical, pharmaceutical and cosmetic industries. In particular amphiphilic block copolymers have attracted a great deal of interest from the biomedical field [28], with applications such as drug delivery vectors [29], nanoparticle stabilizers, nanoreservoirs, emulsion stabilizers, wetting agents, rheology modifiers [30, 31] or as injectable scaffold materials for tissue engineering [32]. Using a generic coarse-grained model and Monte Carlo simulations, we want to improve our understanding of the behaviour of amphiphilic multiblock copolymers in solution.

3.2 Defining a model

In view of understanding the impact of the substitution pattern and ratio on the behaviour of methylcellulose, copolymers are modelled as chains of N_m connected monomers of two types, either hydrophobic (H) or hydrophilic (P), depending on the number of substituted hydroxyl groups. We ignore the chemical details of the monomer units and each of them is represented as a single bead.

As stated before, experiments tend to show that methylcellulose has a blocky chemical structure, with highly substituted zones along the chain. This is why in our model copolymers are represented as regularly alternating blocks of hydrophobic and hydrophilic monomers. They are denoted as $(H_{B_H} P_{B_P})_n$, where B_H is the number of H monomers per hydrophobic block, B_P the number of P monomers per hydrophilic block and n the number of $H_{B_H} P_{B_P}$ patterns in the polymer. A given multiblock copolymer thus consists of a total of $N_m = n(B_H + B_P)$ monomers and its hydrophobic substitution rate, P_{sub} , is defined as:

$$P_{sub} = \frac{B_H}{B_H + B_P} \quad (1)$$

Both types of monomers are subjected to excluded-volume constraints. The solvent effect is simulated implicitly and the interactions between the monomers and the solvent are accounted for indirectly. An effective short-range attractive interaction between the H monomers mimics the effect of hydrophobicity which favours the association of the H monomers. In the model the interaction is restricted to the nearest H neighbours which corresponds to the range of the repulsion between the H monomers and the solvent. This so called hydrophobic interaction strongly depends on the quality of the solvent with respect to the H monomers. In practice this interaction is taken into account through the effective energy E_i between neighbouring H monomers: when $E_i = 0$, there is no interaction between H monomers and the quality

of the solvent is equally good for both kinds of monomers; for increasing $|E_i|$, the solvent becomes increasingly poor for H monomers which leads to the formation of H clusters. The notion of good and poor solvent always refers to H monomers, P monomers are always in good solvent conditions. In summary, the interaction is always equal to zero between two hydrophilic monomers and between a hydrophilic and a hydrophobic monomer while there is an attractive interaction of strength E_i between two hydrophobic monomers. Hence our model takes into account different solvent qualities and various sizes and ratios of hydrophobic and hydrophilic blocks.

3.3 Lattice model and Monte Carlo

The behaviour of the multiblock copolymers is simulated using the algorithm of coarse-grained cell polymer dynamics developed by Kolb and Axelos [33].

The basic idea of this model is to replace the complicated continuum dynamics of a polymer solution or melt by a truncated lattice version. Consider any continuum polymer system. To discretize this system we regularly divide space into compact cells. The centers of these cells form a regular lattice. The polymer conformation and its dynamics can then be uniquely discretized by placing all the monomers that lie in a given cell onto the corresponding lattice site. The cell size is a free parameter, for the present calculation we set it equal to two monomer volumes, i.e. each lattice site can be occupied by zero, one or two monomers. After projection of the conformation onto the lattice, the continuum dynamics becomes a nearest neighbour hopping dynamics. The bond lengths of the original model are also discretized to the values zero or one lattice spacings. This corresponds to a variation of the effective average bond length between one half and one lattice spacings, for a chain of doubly or singly occupied sites respectively.

The original polymer model and the lattice model have the same properties on scales larger than a lattice spacing. Computationally a lattice model is much more efficient, as the complicated continuum dynamics is replaced by a simple nearest neighbour hopping dynamics on the lattice. In order to reproduce statics and dynamics correctly, a lattice model must respect three essential features of polymer structure and dynamics: the monomer connectivity along the chain, the excluded volume interaction and the non-crossing of polymer strands. The connectivity requirement is assured by construction because neighbouring monomers along the chain always sit on the same or on nearest neighbour sites. The excluded volume restriction is respected automatically because each site can be occupied by at most two monomers. The condition that two polymer strands are not allowed to cross would require elaborate testing for the proposed discretization. To avoid this, we change the rules slightly: the double occupancy of a lattice site is restricted to monomers that are nearest neighbours along the same polymer chain (chemically bonded monomers). This is a minor change, which effectively corresponds to an increase of the range of the excluded volume interaction, but which allows to guarantee the strict non-crossing of strands in a simple way.

The key difference between cell polymer dynamics and standard lattice models is that in our model a lattice site can be occupied by two monomers. This is also its main advantage: pure reptation along the polymer chain is explicitly possible even at the highest densities, monomers moving between a doubly occupied and an empty site naturally create local chain length fluctuations on the lattice, there are no locked up conformations and hence no ergodicity problems as for standard lattice models. The model is computationally efficient as it has, under athermal conditions, no adjustable parameters and only one move type. The unique space and time scales of this move correspond to the cell size and its characteristic relaxation time of the corresponding real polymer system. The model is designed in a way such that the cited three basic conditions of polymer dynamics are guaranteed by construction for every move. It can therefore be implemented as easily and as efficiently as the simplest lattice models.

Figure 4 shows a two-dimensional (hexagonal) representation of the scheme. The choice of the fcc or hexagonal lattice is motivated by a greater number of nearest neighbours (12 resp. 6) and a wider and smoother range of bonding angles than a standard cubic (6) or square (4) lattice. For a precise operational definition of cell polymer dynamics the reader is referred to Ref. [34].

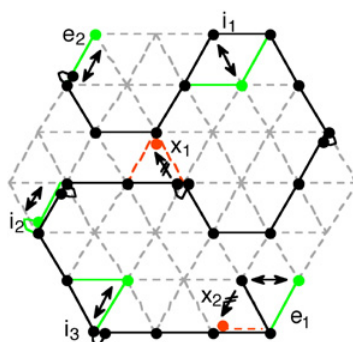


Figure 4. 2D representation of a polymer on the lattice and the possible monomer moves. Black dots and bold lines represent an allowed conformation of the chain and its monomers. Green lines and dots are acceptable moves of end monomer (e_1, e_2) resp. internal monomers (i_1, i_2, i_3). Red lines and dots (x_1, x_2) show forbidden moves (From [34]).

In this algorithm, Rouse type moves dominate the dynamics for dilute solutions whereas reptation moves dominate the dynamics at melt densities, just as expected for real polymers [5]. The shift from Rouse dominated to reptation dominated dynamics occurs naturally upon increasing the monomer density and without the tuning of a parameter.

Examples of moves that are rejected because they would lead to forbidden conformations are also shown in Figure 4 (internal monomer x_1 , end monomer x_2). Interactions are introduced by assigning every monomer to be H or P and by adding an energy contribution $E_i < 0$ to the total energy for every pair of H monomers that either occupy the same site or that occupy two nearest neighbour sites. When trying to move an interacting hydrophobic (H) monomer, the move is only accepted if it is energetically favourable, according to the Metropolis sampling scheme. The (dimensionless) energy difference ΔE of a conformation before and after a move is calculated and the move is accepted if all other constraints are satisfied and if $\min(1, \exp(-\Delta E)) < ran$, where $0 \leq ran \leq 1$ is a uniformly distributed random number. Time is measured in Monte Carlo steps (MCS) which corresponds to $N_m N_p$ trial moves, where N_m is the total number of monomers per chain and N_p the number of chains in the system.

4. DILUTE SOLUTION OF MULTIBLOCK COPOLYMERS

In the case of dilute solutions of copolymers, no intermolecular interactions are included and therefore we may consider the properties of a single copolymer. Simulations give access to different kinds of information. Direct visual inspection of the conformations and their evolution with time gives a first insight into the structure and dynamics of the mesostructures as a function of the parameters. A more quantitative analysis can be made from calculating structural properties of the simulated configurations: radii of gyration, size of H clusters and form factors. Both the radii of gyration and the form factors are also available from experiments. Confronting visual inspection and statistical properties allows one to identify the signature of each conformation that is expected in an experimental measurement (e.g. scattering experiment).

4.1 Effect of solvent quality

We first focus on the effect of the interaction energy E_i between H monomers on the overall size of the chain. The copolymer is characterized by its radius of gyration R_g , computed as:

$$R_g^2 = \frac{1}{N_m} \left\langle \sum_{i=1}^{N_m} (\vec{r}_i - \vec{r}_{cm})^2 \right\rangle \quad (2)$$

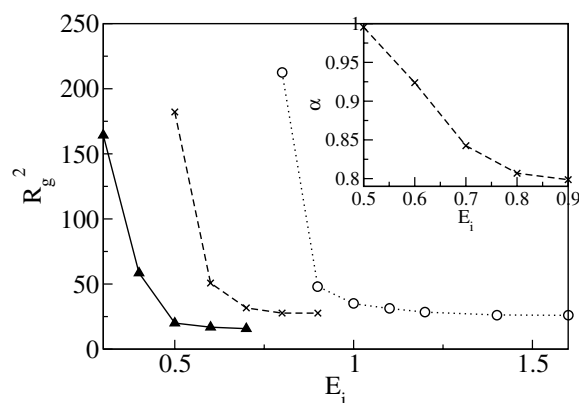


Figure 5. Mean squared radii of gyration as a function of E_i , for $N_m = 600$ and $B_H = 3$. Dotted line: $P_{sub} = 0.1$; Dashed line: $P_{sub} = 0.3$; Full line: $P_{sub} = 0.5$. Inset: Ratio $\alpha = R_{gH}^2 / R_{gP}^2$ versus E_i for $P_{sub} = 0.3$ (From [34]).

where \vec{r}_i is the position of monomer i and \vec{r}_{cm} is the position of the center of mass of the polymer at a given time, and $\langle \cdot \rangle$ denotes an average over all configurations. Figure 5 shows the evolution of the average value of the squared radius of gyration of the copolymer with the interaction energy E_i , for chain length ($N_m = 600$) and H block size ($B_H = 3$), for three values of P_{sub} ($P_{sub} = 0.1, 0.3$ and 0.5). We observe the expected behaviour, from a swollen state with large R_g^2 at low values of E_i (good solvent) to a more compact state with high solvent selectivity (poor solvent for H monomers). Decreasing the length of the non-interacting P blocks while keeping the H block length fixed (*i.e.* increasing P_{sub}) shifts the collapsed region to weaker energies. The same effect is observed when increasing simultaneously the lengths of interacting and non-interacting blocks while keeping P_{sub} fixed (not shown here), in agreement with previous findings [35].

Also form factors can be computed using the simulated configurations :

$$F(q) = \frac{1}{N_m} \left\langle \sum_{j,k=1}^{N_m} \exp [i\vec{q} \cdot (\vec{r}_j - \vec{r}_k)] \right\rangle \quad (3)$$

Since the system is isotropic, the form factor can be averaged over all \vec{q} vectors of equal length, using the scheme described in Ref. [36]. The average $\langle \cdot \rangle$ is also performed over all configurations. We limit the analysis of the form factor at large q to $q < 2\pi/a$, where a is the lattice spacing, as for $q \gtrsim 2\pi/a$ the scattering intensity is dominated by lattice effects. We assume here that the scattering lengths of both kinds of monomers are equal. Figure 6 shows the form factors of the chains under different solvent conditions and for two values of P_{sub} . At $E_i = 0$, the solvent is equally good for H and P monomers and the copolymer shows a fractal dimension of 1.7 in the intermediate q -range, typical of homopolymers in good solvent with $F(q) \propto q^{-1.7}$. For interaction energies typical of poor solvents (as determined from the energy dependence of R_g , see Figure 5), several features can be noticed. The low q region of $F(q)$, for $qR_g \ll 1$, gives access to the overall size of the copolymer, using the Guinier approximation. Figure 6 is in agreement with the observations of the evolution of R_g , with a compaction of the coil as E_i increases, for fixed P_{sub} : In a poor solvent the plateau is reached at larger q values (smaller values of R_g). In the intermediate q -range, for $qR_g \geq 1$, one probes the internal structure of the polymer. In particular, $F(q) \propto q^{-4}$ characterizes a compact spherical structure. A homopolymer ($P_{sub} = 1$) under poor solvent conditions is expected to collapse into such a compact spherical structure (see Figure 6). For P_{sub} values different from $P_{sub} = 1$ (red line in Figure 6), the behaviour is different and correlations appear at intermediate q values, which become more pronounced with increasing E_i at fixed P_{sub} (not shown).

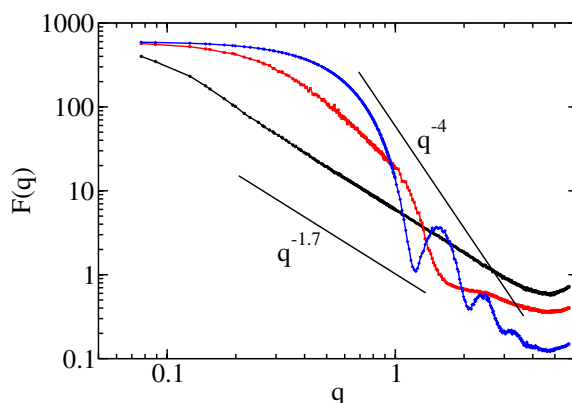


Figure 6. Total form factors for different values of P_{sub} with $N_m = 600$ and $B_H = 3$. Good solvent conditions: $E_i = 0.0$ (black line). Poor solvent conditions: $P_{sub} = 0.3$, $E_i = 0.9$ (red line); $P_{sub} = 1$, $E_i = 0.4$ (blue line) (From [34]).

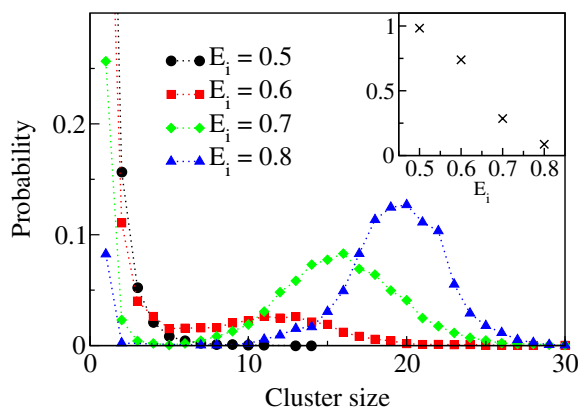


Figure 7. Probability distribution (number probability) of cluster sizes and their evolution with interaction energy, from $E_i = 0.5$ (good solvent) to $E_i = 0.8$ (poor solvent). The copolymer is defined as $(H_{B_H} P_{B_P})_{60}$ with $N_m = 600$, $B_H = 3$, $B_P = 7$ and $P_{sub} = 0.3$. The cluster size is expressed in numbers of hydrophobic blocks. The probability is negligible beyond $N_c \geq 30$. The inset shows the total fraction of small clusters ($N_c < 5$) as a function of E_i (From [34]).

The presence of correlation lengths in $F(q)$, as shown by the shoulder at $q \approx 1$, suggests the existence of mesostructures with characteristic lengths smaller than the overall chain size R_g in poor solvent. This observation is likely to be related to the existence of dense clusters of H monomers, as indicated by the decreasing ratios of hydrophobic to hydrophilic radii of gyration (Figure 5). To check this hypothesis we have identified hydrophobic clusters in the simulated conformations. For our purpose a cluster is defined as a set of interacting H blocks. Two blocks are said to interact if any two monomers of these blocks are nearest neighbours. This property is transitive, if block A interacts with block B and block B with block C then all three belong to the same cluster. The aggregation number N_C of a cluster is defined as the number of H blocks belonging to this cluster.

From the cluster size distributions for different values of P_{sub} and E_i , and for sufficiently long chains ($N_m = 600$), we distinguish two main types of behaviour. For low P_{sub} , in a poor solvent a distribution of cluster sizes appears around a well defined average value N_c . With increasing energy E_i the peak of the distribution becomes sharper and shifts to higher values. This is particularly clear for $P_{sub} = 0.3$ (see Figure 7). For $E_i = 0.8$ there are $\langle N_c \rangle \approx 20$ blocks in a cluster which corresponds to an

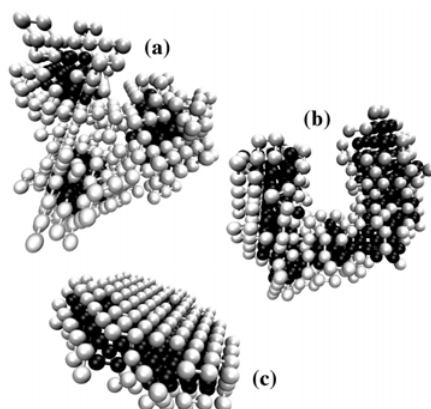


Figure 8. Typical configurations for different substitution rates: (a) $P_{sub} = 0.3$, chain of micelles; (b) $P_{sub} = 0.5$ with $B_H = 3$, tubular structure; (c) $P_{sub} = 0.5$ with $B_H = 2$, layered structure. H monomers are depicted in black, P monomers in light gray (From [34]).

average of three H clusters per chain. For $P_{sub} = 0.1$ the same general picture is observed, but with the interaction energies shifted to higher values. Simulations with $N_m = 600$ and 1200 at $P_{sub} = 0.1$ support the expectation that the number of clusters is growing with the chain length and the size distribution (aggregation number) is independent of chain length. For large values of P_{sub} , $P_{sub} = 0.5$, the cluster size distribution is qualitatively different from small P_{sub} . Up to $E_i = 0.4$ no well defined cluster is present, beyond $E_i = 0.5$ almost all H blocks collapse into a single cluster. Not surprisingly, in both cases a decrease of the probability of finding isolated H blocks with increasing interaction energy E_i is observed.

Calculation of the radii of gyration, form factors and cluster size distributions gives information about the influence of the solvent quality on the overall structure of the polymer coil but also on its internal correlations. For a full understanding of the intramolecular structures in poor solvent, further analyses have to be performed.

4.2 Intramolecular structures in poor solvent

A first insight into the intramolecular structures in poor solvent can be obtained from the direct visualization of copolymer conformations. We notice several well-defined features, as shown in Figure 8.

For $P_{sub} \leq 0.3$, we observe the formation of intramolecular chains of micelles (a single intramolecular micelle for short chains): The H blocks aggregate into clusters (see Figure 7) which are surrounded by P blocks (see Figure 8a). These clusters typically are spherical or ellipsoidal, and they are linked to each other by one or several P blocks. The number of such intramolecular clusters increases linearly with N_m .

For $P_{sub} \geq 0.5$, when decreasing the solvent quality, the copolymer undergoes a transition from a swollen chain to a chain containing H aggregates and then to a single cluster containing all H blocks, no matter how long the copolymer. There are two different types of structures observed in the very poor solvent single cluster regime with $P_{sub} = 0.5$. For $B_H = B_P = 3$, upon increasing the energy there first appears a tubular shaped single cluster (the tube may be open, torus-like or with its ends folded back anywhere onto the tube) and for stronger energies a layered structure with a hydrophobic inner layer (crystalline, possibly distorted and with defects), see Figure 8. The transition from tubular to layered structure with increasing E_i proceeds via the formation of a chain of growing pieces of the layered structure. The transition from the tubular to the layered structure is most clearly seen in simulations

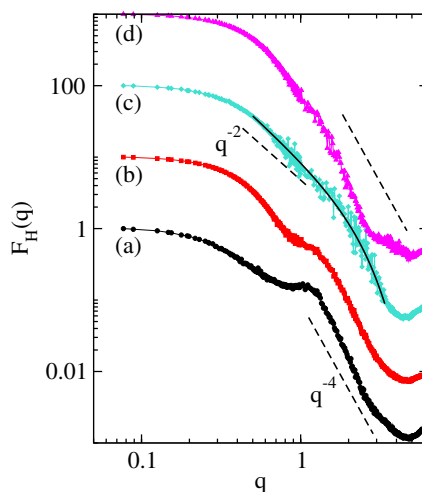


Figure 9. Partial form factor of hydrophobic monomers under poor solvent conditions. From bottom to top: (a) $P_{sub} = 0.3$, $N_m = 600$, $B_H = 3$, $B_P = 7$, $E_i = 0.9$ (chain of micelles); (b) $P_{sub} = 0.5$, $N_m = 600$, $B_H = 3$, $B_P = 3$, $E_i = 0.7$ (tubular structure); (c) $P_{sub} = 0.5$, $N_m = 600$, $B_H = 2$, $B_P = 2$, $E_i = 0.7$ (layered structure); (d) $P_{sub} = 0.625$, $N_m = 600$, $B_H = 5$, $B_P = 3$, $E_i = 0.8$ (layered structure with rough surfaces and defects). Solid line: analytical form factor of a 2d-platelet. Note that the curves are normalized by N_H and shifted vertically by one order of magnitude for clarity (From [34]).

using a slowly increasing energy ramp. The formation of a tubular and then a layered structure upon increasing the energy is also observed for $B_H = 5$ and $B_P = 3$. On the other hand, for $B_H = B_P = 2$, there is no tubular structure and the layered cluster is formed directly with poorer solvent quality. In this case the layered structure is much more regular and the surfaces are much smoother than the previous ones.

A more detailed analysis of the intramolecular structures can be obtained from the partial structure factors $F_H(q)$ (see Figure 9), $F_{cluster}(q)$ (not shown) and $F_P(q)$ (see Figure 10). $F_H(q)$ ($F_P(q)$) gives information about the correlations between all H (all P) monomers, whereas $F_{cluster}(q)$ correlates H monomers within the same cluster and can be used for characterizing the shapes and sizes of the H clusters (see Sec. 4.1). Let us first focus on the properties of $F_H(q)$ (Figure 9) and $F_{cluster}(q)$. For small q values, $F_H(q)$ reaches a plateau, which corresponds to the Guinier regime where $qR_g \ll 1$. For intermediate q values, $F_H(q)$ tends to behave as q^{-4} , which stands for a compact object with a sharp interface (well defined density drop). This feature is independent of P_{sub} and in this range of q values, $F_H(q)$ and $F_{cluster}(q)$ can be superimposed (adjusting the scales appropriately). Note that for $P_{sub} > 0.3$, $F_H(q) \approx F_{cluster}(q)$ as most H blocks belong to a single cluster. For $P_{sub} = 0.3$, $F_{cluster}(q)$ can be fitted to the analytical form factor of a polydisperse sphere of radius $R_s = 1.85$ (not shown, see [34]). For lower q values, the behaviours of $F_H(q)$ and $F_{cluster}(q)$ depart from each other for $P_{sub} = 0.3$: $F_H(q)$ shows a well-defined correlation distance at $q_c \approx 1.2$ (≈ 5.2 lattice units in real space), which becomes more pronounced for longer chains and disappears for short chains (single intramolecular micelle). Therefore this correlation peak reflects the average distance between H clusters: It is absent in the case of a single cluster and increases when the number of clusters and hence the number of inter-cluster distances increases. For $P_{sub} = 0.5$ and $B_H = 3$ (see Figure 9b), we observe an inflexion point around $q \approx 1.3$, which turns into a short plateau for longer chains ($N_m = 1200$, not shown here). As explained above (and detailed in Sec. 4.1), there is a single cluster of H monomers and this inflexion point might correspond to the mean distance between monomers on opposite sides of closed tubular conformations, predominant in this case. This hypothesis is supported by the fact that the intensity at which the inflexion occurs increases with increasing E_i , when the structure becomes better defined. For $P_{sub} = 0.5$ and

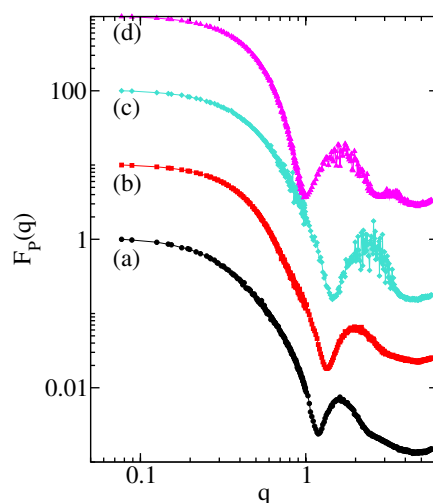


Figure 10. Partial form factors of hydrophilic monomers in poor solvent conditions. From bottom to top: (a) $P_{sub} = 0.3$, $N_m = 600$, $B_H = 3$, $B_P = 7$, $E_i = 0.9$ (chain of micelles); (b) $P_{sub} = 0.5$, $N_m = 600$, $B_H = 3$, $B_P = 3$, $E_i = 0.7$ (tubular shape); (c) $P_{sub} = 0.5$, $N_m = 600$, $B_H = 2$, $B_P = 2$, $E_i = 0.7$ (layered structure); (d) $P_{sub} = 0.625$, $N_m = 600$, $B_H = 5$, $B_P = 3$, $E_i = 0.8$ (layered structure with rough surfaces). Note that curves are normalized by N_H and shifted vertically for clarity (From [34]).

$B_H = 2$, $F_H(q)$ becomes quite different with the onset of a q^{-2} behaviour at intermediate q values, which is typical of two-dimensional structures. For higher q , the structure is compact with a q^{-4} slope. The H blocks form a compact two-dimensional structure. The same kind of behaviour is observed for $P_{sub} = 0.625$ with $B_H = 5$ and $B_P = 3$, however the onset of the q^{-2} behaviour is not observed for two reasons: For the same value of N_m , the layer is thicker and its surface area is smaller than for $B_H = 2$, which reduces the range of the q^{-2} behaviour. At the same time, the surface is much rougher and shows more defects for $P_{sub} = 0.625$, which leads to an effective surface fractal dimension $d_g \geq 2$.

If we now move on to the contribution from the P monomers (Figure 10), the salient feature of $F_P(q)$ is a peak at a q_P value between 1 and 2, depending on P_{sub} . The most straightforward case to interpret is the two-dimensional platelet of H blocks, as found for $P_{sub} = 0.5$ with $B_H = 2$ and $P_{sub} = 0.625$ with $B_H = 5$. Due to connectivity constraints, the P blocks must coat the surfaces of the H platelet and consequently, $F_P(q)$ has a scattering function similar to that of stacked lamellae. The position and widths of the peak is in this case related to the spacing d between two consecutive lamellae [37]. For the other two cases, the situation is quite different. The distances $d_p = \frac{2\pi}{q_P}$ can be associated with the tube diameter and the sphere diameter of the hole due to the presence of the H core (see [34] for more details).

4.3 Overview of the behaviour in dilute solution

Comparing the statistical properties extracted from the simulations with the direct visualization of the intramolecular structures, we are able to give an accurate description of the conformations of single multiblock copolymers as a function of solvent quality in terms of the form factors.

For weak interaction energies, the copolymer trivially has the properties of a swollen coil, as shown by $F(q) \propto q^{-1.7}$. The transition from good to poor solvent is accompanied by the formation of fluctuating aggregates due to the attractive interaction between H monomers. With increasing values of E_i , after the individual collapse of the hydrophobic blocks different intramolecular structures develop. This is shown in the phase diagram, Figure 11, as a function of energy E_i and hydrophilic block length

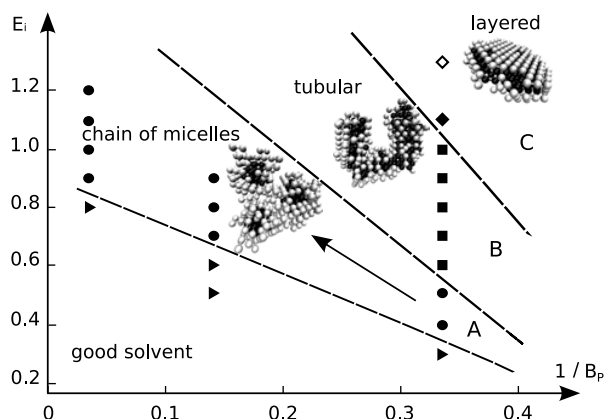


Figure 11. Phase diagram of the intramolecular structures of multiblock copolymers as a function of E_i and $\frac{1}{B_P}$, for $B_H = 3$. Symbols denote simulated systems, with equilibration at constant energy (solid symbols) or with an energy ramp (open symbol). Copolymers in good solvent are denoted as triangles, pearl necklaces of micelles as dots, tubular structures as squares and layered structures as diamonds. Dashed lines indicate the boundaries between the phases corresponding to the different structures. Snapshots show typical conformations observed in the simulations. The arrow points in the direction where a simultaneous increase of E_i and B_P maintains the energy-entropy balance necessary for stabilizing a given structure (From [34]).

B_P . With increasingly poorer solvent the qualitatively different conformations encountered are, in order, chains of micelles (see region A in Figure 11), tubular (region B) and layered structures (region C). The diagram corresponds to a given value of B_H ($B_H = 3$ in Figure 11), but it is similar for other values of B_H . The parameters E_i and B_P (or $\frac{1}{B_P}$ in the diagram) are used, because the observed structures result from a balance between energy and entropy: E_i governs the energy contribution of the interacting H monomers, while B_P is a measure for the configurational entropy of the P blocks. Different values of B_H change the size of the elementary H cores, but do not change the principle of the energy-entropy balance responsible for the different structures.

In region A of Figure 11, intramolecular pearl necklaces of core-shell micelles are formed, with well-defined core size of the micelles. The size of the pearls are stabilized by the P shells: Increasing the size of the H clusters, for a given E_i and B_P , would lead to an important entropy penalty due to the crowding of the P shell; decreasing the size of the H cores would induce an energy loss that is not compensated by a sufficient entropy gain. As a consequence, a simultaneous increase of E_i and B_P in region A (see the arrow) keeps the topology of the structure unchanged: strengthening E_i tends to increase the size of the H cores, increasing the length of the P blocks reinforces the protective corona of the micelles and tends to decrease the core size.

In region B, the formation of tubular-shaped structures with a hydrophobic core is observed. Such a tubular core is energetically more favourable than the spherical cores of region A but it has less surface area available for the hydrophobic corona. With respect to region A, region B corresponds either to copolymers with shorter P blocks (larger $\frac{1}{B_P}$) thus reducing the entropic contribution, or to larger energies E_i , or both. As in region A, the length of the copolymer has no influence on the structure of the tube. The energy E_i does have an effect on the detailed shape of the tube: for smaller values the tube and the corona have a spherical cross-section whereas for larger energies small parts of the tube locally begin to form pieces of layered domains, as a precursor of the third morphology.

In region C one observes two-dimensional disk-like layered structures with an inner H layer and two outer P layers. With respect to B this morphology requires a further increase of the energy E_i and/or a decrease of the length of the hydrophilic blocks B_H . Besides the data point for $B_H = 3$ shown in Figure 11 similar structures were obtained for $E_i \approx 0.7$ for $B_H = B_P = 2$ and for $B_H = 5, B_P = 3$. The

most perfect crystalline structures were obtained for $B_H = B_P = 2$ whereas for $B_H = 5$ and $B_P = 3$ the resulting crystal is less perfect and with a rougher surface. This difference can be explained by the very short H and P blocks for $B_H = B_P = 2$, which impose a strong restriction on the hydrophobic monomers during the formation of the layer.

Simulations with other values of B_H confirm the morphologies and their dependence on the parameters. The chain of transitions of aggregates \rightarrow tubular shape \rightarrow disk-like layered structure with increasing E_i is not only observed for $B_H = B_P = 3$ but also for $B_H = 3$ and $B_P = 5$, with energies $E_i = 0.5$ and $E_i = 0.7$, respectively. The different morphologies and their order in the phase diagram can be understood as a result of the competition between entropy and energy: starting from the origin, the same morphological changes are expected from an entropy decrease (horizontally) and from an energy increase (vertically).

The different morphologies found in our simulations can be compared with the theoretical work based on scaling arguments by Borisov and Zhulina [38] on amphiphilic graft copolymers with hydrophobic backbone: They predict the formation of analogous structures (necklace of star-like or crew-cut micelles, cylindrical worm-like micelle and lamellar structure) depending on the degree of branching of the copolymer and the solvent quality. Despite the different connectivity constraints for the two polymers, sparse grafting may be compared to high substitution rate and dense grafting corresponds to multiblock copolymers with few associating monomers. On the experimental side, different studies on multiblock linear copolymers and graft copolymers revealed the formation of intramolecular core-shell micelles, chains of micelles and rod-like micelles (see [34] for more details), in qualitative agreement with our findings.

5. SEMIDILUTE SOLUTION OF MULTIBLOCK COPOLYMERS

In semidilute solution both intra- and intermolecular interactions are accounted for. The concentration ϕ is defined as the occupied fraction of the total available volume which, because of the double occupancy, is twice the number of available lattice sites. Simulations were performed on monodisperse many-chain systems, for two substitution ratios, $P_{sub} = 0.2$ and $P_{sub} = 0.5$, for different concentrations ϕ ranging from 0.01 to 0.5 and for different interaction energies E_i . This enables a broad range of behaviours to be encompassed, from rather dilute and good solvent conditions, to concentrated and strongly interacting systems. Also different chain lengths N_m and numbers N_p are used as they give insight into their influence on phase behaviour (influence of N_m on percolation threshold) and finite-size effects (influence of the box size at a given concentration). For poorly substituted copolymers micelles are formed, which may be either intra- or intermolecular depending on concentration. For the higher substitution ratio ($P_{sub} = 0.5$), larger tubular structures form which grow with concentration. Also gelation is observed. The structure of the observed network strongly depends on P_{sub} : for poorly substituted copolymers, the connection is through the cross-linked micelles while for the highly substituted chains, the connection is through the extended hydrophobic cores. The interplay between gelation and phase separation of the hydrophobic monomers is observed in the case of poorly substituted copolymers. More details about this work can be found in [39].

Acknowledgements

The author would like to thank Drs. Max Kolb and Monique Axelos for fruitful collaboration while working on the behaviour of multiblock copolymers.

References

- [1] S. O Nielsen, C. F Lopez, G. Srinivas, and M. L Klein. Coarse grain models and the computer simulation of soft materials. *J. Phys. Condens. Mat.*, 16:481, 2004.

- [2] V. Tozzini. Multiscale Modeling of Proteins. *Accounts Chem. Res.*, 43:220, 2010.
- [3] M. Müller, K. Katsov, and M. Schick. Biological and synthetic membranes: What can be learned from a coarse-grained description? *Phys. Rep.*, 434(5-6):113–176, 2006.
- [4] K. Binder. *Monte Carlo and molecular dynamics simulations in polymer science*. Oxford University Press US, 1995.
- [5] M. Doi and S. F. Edwards. *The theory of polymer dynamics*. Oxford University Press, USA, 1986.
- [6] P. G. de Gennes. *Scaling concepts in polymer physics*. Cornell Univ. Press, 1979.
- [7] G. Srinivas, D. E. Discher, and M. L. Klein. Self-assembly and properties of diblock copolymers by coarse-grain molecular dynamics. *Nat. Mater.*, 3(9):638–644, 2004.
- [8] X. He and F. Schmid. Spontaneous formation of complex micelles from a homogeneous solution. *Phys. Rev. Lett.*, 100(13):137802, 2008.
- [9] A. D. Sokal. *Monte Carlo methods for the self-avoiding walk*, pages 47–124. 1995.
- [10] I. Carmesin and K. Kremer. The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions. *Macromolecules*, 21(9):2819–2823, 1988.
- [11] J. S. Shaffer. Effects of chain topology on polymer dynamics: Bulk melts. *J. Chem. Phys.*, 101(5):4205, 1994.
- [12] R. D. Groot and T. J. Madden. Dynamic simulation of diblock copolymer microphase separation. *J. Chem. Phys.*, 108:8713, 1998.
- [13] K. Kremer, G. S. Grest, and I. Carmesin. Crossover from rouse to reptation dynamics: A Molecular-Dynamics simulation. *Phys. Rev. Lett.*, 61(5):566, 1988.
- [14] K. Kremer and G. S. Grest. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.*, 92(8):5057, 1990.
- [15] D. Viduna, A. Milchev, and K. Binder. Monte carlo simulation of micelle formation in block copolymer solutions. *Macromol. Theor. Simul.*, 7(6):649–658, 1998.
- [16] A. Milchev and K. Binder. Formation of surface micelles from adsorbed asymmetric block copolymers: A monte carlo study. *Langmuir*, 15(9):3232–3241, 1999.
- [17] M. A. Floriano, E. Caponetti, and A. Z. Panagiotopoulos. Micellization in model surfactant systems. *Langmuir*, 15(9):3143–3151, 1999.
- [18] A. Z. Panagiotopoulos, M. A. Floriano, and S. K. Kumar. Micellization and phase separation of diblock and triblock model surfactants. *Langmuir*, 18(7):2940–2948, 2002.
- [19] M. E. Gindy, R. K. Prud'homme, and A. Z. Panagiotopoulos. Phase behavior and structure formation in linear multiblock copolymer solutions by monte carlo simulation. *J. Chem. Phys.*, 128(16):164906–13, April 2008.
- [20] S. Abeln and D. Frenkel. Disordered flanks prevent peptide aggregation. *PLoS Comput. Biol.*, 4(12):e1000241, 2008.
- [21] F. Müller-Plathe. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *ChemPhysChem*, 3:754–769, 2002.
- [22] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. The martini force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B*, 111:7812–7824, 2007.
- [23] P. J. Hoogerbrugge and J. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.*, 19:155, 1992.
- [24] C. Peter and K. Kremer. Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back. *Soft Matter*, 5(22):4357–4366, 2009.
- [25] M. Hirrien, C. Chevillard, J. Desbrières, M. A. V. Axelos, and M. Rinaudo. Thermogelation of methylcelluloses: new evidence for understanding the gelation mechanism. *Polymer*, 39:6251–6259, 1998.
- [26] P. W. Arisz, H. J. J. Kauw, and J. J. Boon. Substituent distribution along the cellulose backbone in o-methylcelluloses using GC and FAB-MS for monomer and oligomer analysis. *Carbohydr. Res.*, 271(1):1–14, 1995.

- [27] K. Kobayashi, C. Huang, and T. P. Lodge. Thermoreversible gelation of aqueous methylcellulose solutions. *Macromolecules*, 32(21):7070–7077, 1999.
- [28] Y. H. Bae, K. M. Huh, Y. Kim, and K.-H. Park. Biodegradable amphiphilic multiblock copolymers and their implications for biomedical applications. *J. Control. Release*, 64(1-3):3–13, 2000.
- [29] K. Kataoka, A. Harada, and Y. Nagasaki. Block copolymer micelles for drug delivery: design, characterization and biological significance. *Adv. Drug Delivery Rev.*, 47(1):113–131, 2001.
- [30] G. Riess. Micellization of block copolymers. *Prog. Polym. Sci.*, 28:1107–1170, 2003.
- [31] S. R. Bhatia, A. Mourchid, and M. Joanicot. Block copolymer assembly to control fluid rheology. *Curr. Opin. Colloid Interface Sci.*, 6(5-6):471–478, 2001.
- [32] A. Fatimi, J.-F. Tassin, S. Quillard, M. A. V. Axelos, and P. Weiss. The rheological properties of silylated hydroxypropylmethylcellulose tissue engineering matrices. *Biomaterials*, 29(5):533–543, 2008.
- [33] M. Kolb and M. A. V. Axelos. Rouse and reptation dynamics: a polymer lattice model. *XV Conference of the European Colloid and Interface Society, Coimbra, Portugal*, 2001.
- [34] V. Hugouvieux, M. A. V. Axelos, and M. Kolb. Amphiphilic multiblock copolymers: From intramolecular pearl necklace to layered structures. *Macromolecules*, 42(1):392–400, 2009.
- [35] J. M. P. van den Oever, F. A. M. Leermakers, G. J. Fleer, V. A. Ivanov, N. P. Shusharina, A. R. Khokhlov, and P. G. Khalatur. Coil-globule transition for regular, random, and specially designed copolymers: Monte carlo simulation and self-consistent field theory. *Phys. Rev. E*, 65(4):041708, 2002.
- [36] J. Baschnagel and K. Binder. Structural aspects of a three-dimensional lattice model for the glass transition of polymer melts: a monte carlo simulation. *Physica A*, 204(1-4):47–75, 1994.
- [37] F. Nallet. De l'intensité à la structure en physico-chimie des milieux dispersés. In J. P. Cotton and F. Nallet, editors, *Diffusion de neutrons aux petits angles*, volume 9 of *J. Phys. IV*, pages 147–195, France, 1999. EDP Sciences.
- [38] O. V. Borisov and E. B. Zhulina. Amphiphilic graft copolymer in a selective solvent: Intramolecular structures and conformational transitions. *Macromolecules*, 38(6):2506–2514, 2005.
- [39] V. Hugouvieux, M. A. V. Axelos, and M. Kolb. Micelle formation, gelation and phase separation of amphiphilic multiblock copolymers. *Soft Matter*, 7:2580, 2011.