CHAPTER 29

Modeling of Biologically Motivated Soft Matter Systems

Ilpo Vattulainen

Laboratory of Physics and Helsinki Institute of Physics, Helsinki University of Technology, P.O. Box 1100, FIN-02015 HUT, Finland

Mikko Karttunen

Biophysics and Statistical Mechanics Group, Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, FIN-02015 HUT, Finland

CONTENTS

1. INTRODUCTION

Up to about 10 years ago, computational physics was considered to be more or less a fringe discipline having very little fundamental value. This is clearly demonstrated by the extremely small number of universities that offered a proper curriculum in computational physics at that time. This attitude has changed drastically, however, and today computational physics has become part of the mainstream. It has been particularly successful in fields such as materials science and soft matter, in which the complexity of the systems poses formidable challenges for the understanding of their behavior.

The negative attitude may have had its origin in the desire to be able to solve physical problems either by exact analytical means or by well-established experimental techniques. As any computer simulation only provides numerical solutions to models (as compared to exact solutions), it is likely that simulation was not considered to have the same seriousness and fundamental value as other methods. A related dilemma is known by most of those who teach computational methods in interdisciplinary programs: before they learn what computational modeling is about, there are always a few opportunistic students who have the idea that computational modeling is an easy way out without having to know much of the underlying theory or having to deal with complicated equipment. Yet in many ways computational modelers are like experimentalists. Instead of microscopes or accelerators, however, they use computers, and instead of designing experiments, they try to build accurate predictive models—for that they have to have solid knowledge of both the underlying theory and possible experiments.

Today, there is no doubt about the value of computer simulations. It has become very clear that advances in theory, experiments, and computational modeling go hand in hand. This is particularly so in interdisciplinary fields such as soft matter and biophysics [1–4], in which one needs to combine both knowledge and methodologies from very different fields. Computer simulations have also become, or are becoming, standard tools in particle physics, materials science, drug design, and biology, and the list could be easily continued to include almost any field in science and engineering.

In this review we focus on recent developments and advances in modeling soft matter and biological systems. Because of both fundamental and technological reasons, there is strong desire to understand the overall behavior of complex macromolecular and biological soft matter systems. From a theoretical point of view, this task is highly nontrivial because processes in these systems take place over a huge range of length and timescales, whereas current modeling and analytical techniques are feasible over relatively limited scales only.

These systems often belong to the class of so-called complex fluids, which are characterized by their multicomponent nature and simultaneous presence of several, often unseparable, time- and length scales. The presence of many time- and length scales is the crux of soft matter and biological modeling (see Fig. 1).

The time- and length scale problem is best illustrated by an example: water is always present in biological systems. The size of a water molecule is approximately 10^{-10} m. The relevant timescale is defined by molecular vibrations having characteristic times on the order of 10^{-15} s. However, the biologically relevant problem of protein folding can take anything from 1 μs up to about 1000 s, depending on the size of the protein. As for length scales, proteins are long polyampholytes (i.e., polymers carrying both positive and negative charges, and possessing a whole hierarchy of spatial ordering in different interdependent scales; see e.g., Ref. [5]). The other "classical" problem in biophysics involves DNA. The double stranded DNA has a length of 1 m and a persistence length of approximately 50–100 nm (having some sequence dependence). To make things even more complicated, one should remember that proteins, DNA, and lipids operate in cells, which from a biological point of view are the functional entities of interest. Furthermore, cells consist of a large variety of different components (proteins, different lipids, etc.) and have typical sizes around a few tens of micrometers. As if the above would not be problematic enough, in living matter they are actually dissipative structures operating under nonequilibrium conditions. Deciphering the physical mechanisms of living matter is guaranteed to provide excitement and challenges for generations of scientists.

Before moving on to more specific issues of computational modeling, another example is provided by polymers. Although many biologically relevant molecules such as proteins and DNA can be considered as biopolymers, there is also an enormously large field of polymer science not related to living matter. Technologically, polymers are one of the central materials in modern-day society, with applications ranging from sophisticated medical materials to the plastic bags we encounter while grocery shopping. From the modeling point of view, it is really the field of polymers in which the ideas of linking many time- and length scales have developed the fastest. This is easy to understand through the following simple example: The timescales associated with bond vibrations are roughly 10^{-15} s, whereas conformational transitions associated with individual bonds occur typically in timescales of 10^{-11} s. The related changes taking place along the chain take orders of magnitude longer than these
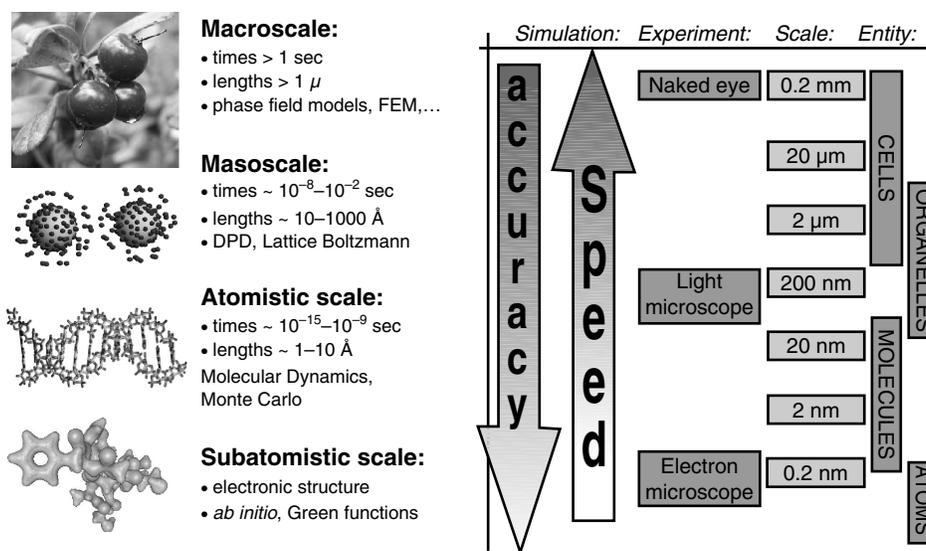
Figure 1. Different time- and length scales, typical computational methods used to study them, and biological entities related to various length scales. The dilemma between speed and accuracy is always present in simulations: it is always a trade-off between the two. In analogy to various simulation methods at different time- and length scales, it is not possible to use single experimental methods to cover all properties. The right-hand side of the figure is adapted from Ref. [4].

timescales. Furthermore, industrially important processes such as spinodal decomposition, or phase separation in general, have characteristic times of at least seconds.

What does the above mean from the point of view of computer simulations? There is a whole plethora of different methods, all having both advantages and limitations. In *ab initio* simulations, quantum mechanical details are taken into account, limiting the obtainable system sizes to a few hundred atoms as a result to the large number of electronic degrees of freedom—this is obviously not enough for polymeric or biological systems, is it? One can make an estimate that using the most powerful central processing units (CPUs) available at the moment for simulating the folding of the smallest proteins by using accurate *ab initio* methods, it would take about 10 000 CPU-years—and this is a fairly optimistic scenario.

As a consequence, quantum mechanical approaches are appropriate only for issues in which electronic degrees of freedom cannot be neglected. In addition, they can be used to develop force fields for classical molecular dynamics (MD) simulations that are able to reach time scales of the order of 100 ns and linear system sizes of some tens of nanometers. Classical MD thus provides a unique tool to study various properties related to many biologically relevant soft matter systems, including small biomolecules and systems composed of these systems. The MD technique is indeed the workhorse in the modeling of biophysical systems, and it will be described in detail in the following sections.

The above very clearly expresses the fact that the range of problems in which quantum mechanics is appropriate is very limited. Further, the realistic time- and length scales accessible through MD simulations are also rather modest and far below those important in, say, polymer dynamics or protein folding. What this suggests is that one should think of building a hierarchy of different methods, each method being valid and useful over a certain well-defined length and timescale. By a proper design, the different methods can be made, ideally at least, rigorously related to each other through systematic coarse-graining approaches in which the detailed atomic or molecular description is replaced with a more simple one. This is the idea behind multiscale modeling which has recently attracted a rapidly increasing amount of attention in computational materials research [6–16]. The approach where classical force fields are obtained from quantum MD and then used in classical MD can, as a matter of fact, be seen as an example of a coarse-graining procedure.

The multiscale approach sounds appealing, but it includes a fundamental problem: There is no unique way to perform coarse-graining. There are many aspects to that problem, but one can immediately get a grasp of it by considering a simple-minded spatial coarse-graining of a cholesterol molecule depicted in Fig. 2: How should one select the new "superatoms"? What are the new interactions in the coarse-grained description? Is the new model consistent with the microscopic one (i.e., does it provide correct static and dynamic properties)? Are there alternative methods? And so on.

Because of the above reasons, and despite the difficulties, coarse-graining has become an increasingly active field of research. A variety of different approaches, some more and others less rigorous, have been presented. We will discuss some of them in the following sections but, to give a flavor, we mention a few here. As already discussed, finding a
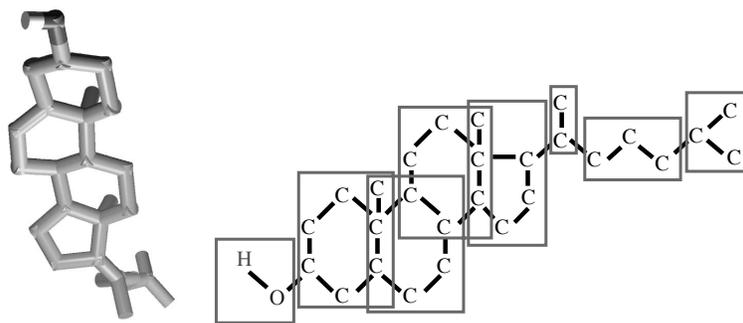
Figure 2. A model of cholesterol. Left: an atomic-level representation with the OH group shown with red and white at the top. Right: the same molecule with a possible definition of new coarse-grained interaction sites, or "superatoms," shown by the rectangular blocks.

systematic hierarchy is a major problem. Another important aspect concerns the fact that multiscale modeling involves two possible processes. First of all is coarse-graining or transforming detailed models to simplified descriptions with less degrees of freedom. This means averaging over some chosen properties of microscopic entities to form larger basic units for modeling. Second, fine-graining which is the opposite of coarse-graining. In this process, configurational properties are typically the key quantities as one maps a coarser model to a more detailed one.

In general, one tries to approach the problem via one of the following routes: by phenomenological methods based on (e.g., Ginzburg–Landau-type approach); by analytical approaches based on the operator projection formalism; by analyzing static properties and constructing the of potential of mean force; by free-energy methods (i.e., configurational analysis); or by analyzing the occurrence rates of different processes.

The phenomenological approaches include methods such as dissipative particle dynamics (DPD; see, e.g., [17]), possibly the most used mesoscale modeling method for soft matter systems. The main idea there is to use "soft" phenomenological interparticle potentials and a momentum-conserving thermostat to ensure hydrodynamics. The intrinsic timescale in DPD is typically 10–20 ps, and the length scale is around 1 nm. DPD was introduced about a decade ago [18–20], and since then various generalizations have been introduced, providing the method a more rigorous theoretical background [21]. We will discuss DPD in detail in the following sections.

A more macroscopic phenomenological approach is based on Ginzburg–Landau-type formalism by introducing order parameters and free-energy and density functionals for the system. This approach is based on finding the relevant (slow) variables and general symmetry principles, and as such is close to a macroscopic description. It is particularly fruitful in dealing with phenomena involving spinodal decomposition or nucleation processes [22]. The review of Kawakatsu et al. [23] puts these models in perspective with more microscopic approaches, and a very detailed discussion of the mathematical aspects can be found in Ref. [24].

The most successful analytical approach is the so-called GENERIC, or general equation for non-equilibrium reversible–irreversible coupling, introduced by Grmela and Öttinger [25–27]. The advantage of this method is that it provides a physically sound approach based on statistical mechanics. It uses projection operator formalism in building up hierarchies between different levels of description, thus ensuring self-consistency. Despite its rather mathematical formalism, it can be used in connection with actual numerical simulations. An excellent review of GENERIC and other analytical approaches is provided by Español in Ref. [28].

Using static properties to obtain potentials of mean force has been a common approach in polymer science. One of the most interesting efforts has been the activity of the Mainz/Bremen region on optimized methods for generating potentials from structural information. The particularly interesting part of that is the inverse procedure (i.e., fine-graining alongside with the coarse-graining efforts). The problems in polymer science are challenging, and it is easy to appreciate them if one remembers that for polymer melts of low molecular weight, the conformational relaxation time (caused by diffusion) scales as N^2 (Rouse scaling), where N is the number of monomers. This said, the scaling is much worse for high-molecular-weight melt. In real-world applications, the relaxation times are from milliseconds upward. It is obvious that coarse-grained methods are needed and that this is really the field in which there has been a lot of effort. For reviews, see Refs. [29–32].

As another example, let us note that there have been only a few published coarse-grained lipid simulations, and to the authors' knowledge, they are all based on interaction site definitions (i.e., selecting interaction sites on some physical principles and then, often phenomenologically, choosing the relevant inter- and intramolecular potentials). That approach is, to a degree, always ad hoc, but the results of the Philadelphia group on lipid membranes [15, 33, 34] and the Groningen group on vesicle dynamics [35] have been very encouraging.

Thus, there is a timely practical problem of finding generally applicable and computationally efficient methods to coarse-grain molecular systems and of establishing a systematic link between the microscopic and macroscopic regimes by the means of statistical mechanics.

We will discuss many of these issues in detail, but we would also like to refer the reader to some of the existing reviews on specialized topics: Kröger [36] discusses practical issues related to the rheology of polymers in detail, including aspects related to GENERIC and projection operator methods. Baschnagl et al. [29] provide an extensive review on coarse-graining in polymer science. In addition, the lecture notes of the summer school Soft-Simu2002 [28] and the lecture notes of the conference "Bridging the Time-Scale Gap" [37] also address issues not covered in this review at length, including hydrodynamics, lattice Boltzmann, granular materials, and quantum-classical models.

In general, multiscale modeling has been a well-adopted and accepted field in Europe, and in particular by the European Science Foundation SIMU Research Network [38]. This is also the case in Japan, where the efforts are highlighted by Masao Doi's ambitious seamless zooming project, which produced the OCTA package for multiscale simulations of soft matter systems [39, 40]. Hence, it is rather surprising that there has been very little activity in North America. It is not difficult to predict that this situation will change.

The rest of the article is organized as follows. In Section 2 we discuss the field of molecular simulations in general, focusing on modeling techniques from the microscopic point of view. That is followed in Section 3 by a discussion of some of the novel and most used mesoscopic simulation methods, including DPD and its extensions, phase field modeling, lattice Boltzmann, and the GENERIC hierarchy. In Section 4 we discuss a number of methodological issues that are crucial in molecular modeling, and we continue in Section 5 by presenting a few applications of atomic molecular dynamics simulations in the context of biophysical systems. Finally, we present an example of coarse graining and mesoscale simulations in Section 6 and close this work by a brief summary in Section 7.

2. MOLECULAR SIMULATIONS

The field of molecular simulations is extremely wide, covering a variety of different methods for different purposes in the quantum, atomic-scale classical, and mesoscopic and continuum regimes. In this section, we briefly discuss the main ideas of atomic-level molecular simulations, starting from a historical perspective and proceeding to a few atomic-scale simulation techniques. As the topic is very wide, we have decided to focus on methods that are most relevant for the applications discussed later in this chapter.

Because of their importance for studies of large-scale properties of soft matter systems, coarse-graining and mesoscale simulation methods are discussed separately in Section 3.

2.1. Historical Perspective from ENIAC to GRID

The beginning of the era of computational sciences can be traced back to the early 1950s, after which the field went through a rapid increase in activities. The progress has taken place hand in hand with the development of computers and computational methods that are the basis of computational modeling of physical systems. Although the former of these two fields is related to microelectronics and semiconductor physics, and the latter is closely associated with applied mathematics and computational sciences, they are both closely coupled to molecular simulations. As a matter of fact, the history of these two fields of science is closely related to each other, too. Since the development of the first electronic computer, ENIAC [41], many of the people who took part in the pioneering work of developing computers have also had a strong effect on the development of computational methods and algorithms. For example, N. Metropolis originally suggested [42] an obvious name for the Monte Carlo (MC) simulation method [43, 44], as the method is largely based on the use of random numbers. Also, with J. von Neumann, Metropolis studied the randomness of the decimals of π and e [45] and developed the first algorithm for generating pseudorandom numbers (the so-called midsquare method [46, 47]).

Ever since, development of computers and simulation methods have taken steps hand in hand. From the pioneering computer simulations of simple liquids on MANIAC, we have come to a situation in which molecular simulations of complex biophysical systems such as DNA condensation and channel proteins in lipid membranes are almost daily routine.

On one hand, this is a results of computing resources that at present are tremendous compared to those that were available in the 1950s. On the other hand, this is also a result of major progress in the development of computer algorithms and the methodology in general.

How have we arrived in the present situation? How the modeling of molecular systems started in the first place? One is tempted to think that the history of molecular models is related to the development of computers in the 1950s, but actually modeling of molecular systems started long before computers were invented. To our knowledge, the first structural models for molecules were suggested in the end of the ninteenth and in the beginning of the twentieth centuries. Despite their rather crude nature, these models allowed one to better understand the three-dimensional structure and related physicochemical properties of liquids. The actual pioneer, however, is probably van der Waals, whose work lead to the well-known equation of state that predicted a first-order gas–liquid phase transition.

Mathematical models in terms of computer simulations joined the game 50 years ago, when the first computer simulation of simple liquids was carried out by Metropolis et al. [44]. The MC method used at that time allowed one to consider structural properties of liquids, and although the advances in computing power have been rapid, the method itself is still alive and well. As a matter of fact, it is nowadays used with success on an ever-increasing number of problems. Yet, as the MC method is not particularly appropriate for studies of dynamical properties, other techniques have been designed for this purpose. The most appropriate approach in this regard is, of course, the MD technique. In MD, a system is described as an ensemble of interacting particles whose evolution in time is found through an integration of Newton's equations of motion. This approach was first used for a system of hard spheres by Alder and Wainwright [48, 49]. A more realistic description of a liquid in terms of Lennard–Jones particles was later accomplished by Rahman in the 1960s [50].

After this groundwork, the progress has been rapid. Barker and Watts suggested [51] the first model for molecular liquid water using MC, and Rahman and Stillinger [52] followed using MD. Nevertheless, despite major activities that have extended for over 30 years, water has remained as one of the most difficult problems in the field of molecular simulations. Equally difficult is the description of proteins that originated from the early work by McCammon in 1977 [53]. The progress ever since has been remarkable, and it is rather confusing to try to understand how tedious biophysical systems can be tackled today despite their highly complex nature. Although the small proteins studied by McCammon in 1977 were considered in vacuum without electrostatics for a timescale of picoseconds, today one is able to simulate transmembrane water channels and other integral proteins in fully hydrated lipid membranes with full electrostatic interactions over a timescale of several nanoseconds [54, 55].

Although the first computer simulations were based on the use of punch cards [56], our work today is based on highly advanced computing architectures together with modern programming languages. Also, starting from the first computers such as ENIAC and MANIAC, we have come to a situation in which computer simulations are carried out in massively parallel supercomputers composed of thousands of central processing units (CPUs). Most recently, initiatives have been made to establish highly parallel and connected networks of computer clusters (GRIDs) to obtain similar capacity across the Internet. Although it is easy to look into the past, the rapid development in the field of molecular simulations does not make it too easy to make any conclusive predictions about the future. Anyhow, in the present and in the next section, we will find that the advances in simulation techniques have been substantial during the last decades. On this basis, it is easy to predict that the future of molecular simulations is both challenging and very positive.

2.2. Main Principles of Molecular Modeling

What is molecular modeling? To clarify this obvious matter, let us first quote the view by Tamar Schlick [57]: "Molecular modeling is the science and art of studying molecular structure and function through model building and computation."

Model building can be as simple as the Lego bricks that the authors liked a lot some 25 years ago—and still do. This approach can provide one with important insight into the

structure of liquids and more complex molecules. The structure of DNA, for example, was first discovered using this approach (though, to our knowledge, not with Lego bricks). The discovery of C_{60} Buckminsterfullerene provides another example, as the first prototype of its structure made by Richard E. Smalley was made of paper sheets [58]. On the computational side, model building starts from building the topology and initial structures of molecules that typically need input from experiments such as nuclear magnetic resonance (NMR) and various scattering techniques. Having done this, one needs to write down the Hamiltonian operator $\mathcal{H}$, which includes all interactions present in the model system. The Hamiltonian operator (often called "force field" in classical simulations) plays the key role in molecular simulations.

In principle, if we were clever enough, we could simply use $\mathcal{H}$ to solve the structure and the dynamics of the system using just paper, pen, and our brains. However, as we are often dealing with a many-body problem having much more than just two interacting components, an analytical exact solution is rarely possible [59]. Alternatively, one can simplify the problem and try to find approximate descriptions of the system, but then the validity of the approximations remains elusive. This barrier can be crossed by computer simulations, where the aim is to study the dynamics of a given system with the corresponding Hamiltonian through computers. Computer simulations allow us to do theoretical experiments without any approximations, thus providing a bridge between theory and experiments.

In the field of molecular simulations, this underlying idea has lead to the introduction of a number of useful simulation techniques, including *ab initio*, molecular mechanics, classical molecular dynamics, MC, free-energy calculation techniques, dissipative particle dynamics, and many many more. Below, we discuss the key issues of the main approaches designed for the atomic regime.

The topology of the molecules is here taken for granted. Thus we concentrate on the Hamiltonian operator and on the related issues of finding a proper force field for a given system. For a more thorough discussion on these techniques, see Refs. [57, 60–62].

2.3. Molecular Modeling in the Atomic Regime: From Quantum Mechanics to Classical Molecular Dynamics

2.3.1. Quantum Mechanical Techniques

Quantum mechanical (QM) simulation methods are based on the solution of the Schrödinger equation. The Schrödinger equation

$$\mathcal{H}\Psi_n = E_n \Psi_n \tag{1}$$

is written in terms of the Hamiltonian operator $\mathcal{H}$, which acts on the eigenfunctions Ψ_n whose corresponding energy eigenvalues (quantum states) are given by E_n. The Schrödinger equation therefore describes the spatial probability distributions of energy states in a stationary case.

As the system in this case is composed of both nuclear and electronic degrees of freedom, one often makes the assumption that the motions of electrons and nuclei can be separated from each other. This so-called Born–Oppenheimer approximation is almost without exceptions in use in present-day QM methods and is well justified because the nuclei are usually much heavier than the electrons, which implies that the positions of nuclei are essentially fixed on the timescale of electronic motion. This assumption reduces the computational load considerably compared to traditional electronic structure methods. Yet, further approximations are often crucial to allow studies of reasonable system sizes. In practice, there are two commonly adapted approaches that allow that.

Ab initio ("from the beginning") is based on the Born–Oppenheimer approximation of the time-independent Schrödinger equation, ignoring relativistic effects and using a linear combination of atomic orbitals that is specific to a chosen basis set. The energy of the system is calculated iteratively (self-consistently), which makes the calculation computationally expensive for reasonable system sizes. Another possibility is to use density-functional theory

(DFT) in which the method is constructed to be based on the use of the electron density function. Different DFT schemes [63] differ by their treatment of the exchange/correlation energy, though in general this class of methods is rather efficient and widely used. One popular technique, for example, is the approach suggested by Car and Parrinello. In this scheme, electronic DFT is used to calculate the energies and densities of the valence electrons "on the fly" [64]. Because the technique is basically MD combined with DFT in the adiabatic case using an extended Lagrangian approach, it is both efficient and highly useful in QM (as well as in classical) problems. Another issue that has attracted plenty of attention is the use of DFT for time-dependent problems. Because of recent progress, time-dependent DFT is nowadays the method of choice for calculating, for example, excitation energies of complex molecules, and it is becoming more and more popular for studies of spectroscopic properties of nanostructures.

Semiempirical QM methods provide an alternative approach. They are partly based on empirical information to replace a full QM description of parameters that define the forms and energies of atomic orbitals. As expected, this reduces the computational load considerably, possibly at the expense of accuracy.

In practice, QM approaches are applied in soft matter systems only to processes and phenomena for which quantum mechanics cannot be neglected. The action of enzymes close to their active site and other chemically activated processes are examples of such cases. Yet the computational burden is major and limits the timescales that can be studied with QM methods up to tens of picoseconds. As a consequence, recent work has focused on combining QM methods with classical approaches such that, for example, the active site of an enzyme is treated quantum mechanically, whereas the remaining part of the system is classical. The problem then is to treat the boundary between QM and classical regimes [65–68].

As we will not present applications based on QM approaches, we prefer not to discuss this issue in more depth. An interested reader is referred to Refs. [57, 62, 69, 70] for more information on QM techniques, and to Refs. [66, 67, 71–73] for recent applications.

2.3.2. Interactions in Classical Methods

Classical simulation techniques are obviously based on QM, but they describe a given system in a considerably more simplified fashion. In essence, in classical methods the quantum degrees of freedom are integrated over—"thrown under the carpet"—such that the QM effects are incorporated into the classical intramolecular and intermolecular interactions. For example, although electronic degrees of freedom are not included in classical molecular simulations, the exclusion rules of QM are taken into account in an effective sense: They are manifested in classical simulations through strongly repulsive interatomic interactions at short distances, which implies that two atoms cannot occupy the same space at the same time.

The absence of electronic degrees of freedom basically implies that classical simulation techniques are based on the assumption that the Born–Oppenheimer approximation is valid. Assuming that this assumption holds true (i.e., we can write the energy of the system as a function of nuclear coordinates, and the computational load of doing molecular simulations is reduced considerably). Instead of worrying about interactions in a QM level, one can then describe the energy of the system as a sum of contributions resulting from different kinds of processes, such as bond stretching and bond bending. The key word here is force field.

The force field is the heart of molecular computer simulations. It describes the internal energy of the system at any moment of time, and it obviously consists of a number of terms, each of which describes some physical process. Generally speaking, for a system of N particles whose coordinates are given by $\{\vec{r}_i\}$, one can write down the force field as follows:

$$V = \sum_i v_1(\vec{r}_i) + \sum_i \sum_{j>i} v_2(\vec{r}_i, \vec{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} v_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) + \cdots \qquad (2)$$

The potential energy of the system is therefore divided into terms describing one-body, two-body, and three-body interactions, as well as other higher-order interactions that have not

been listed here. The one-body interaction (v_1) can be as simple as the action of a uniform electric field on a single atom. Two-body interactions (v_2) are obviously those between two interacting particles, and they are usually divided into bonded and nonbonded interactions. Bonded two-body interactions such as bond stretching are usually those between particles that are bonded by a covalent bond, and nonbonded interactions such as Coulombic interactions are between those atoms that are not bonded, such as atoms in different molecules. Higher-order interactions such as three- (v_3) and four-body interactions include bond bending, torsional interactions, and so forth.

In practice, a simple force field typically used for biologically relevant molecules can look like the following:

$$V = \sum_{\text{bonds}} \frac{k_i^{\text{b}}}{2}(l_i - l_i^{\text{ref}})^2 + \sum_{\text{angles}} \frac{k_i^{\text{a}}}{2}(\theta_i - \theta_i^{\text{ref}})^2 + \sum_{\text{torsions}} \frac{V_T}{2}[1 + \cos(n\omega - \gamma)]$$
$$+ \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left\{ 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \tag{3}$$

Here the first term on the right-hand side describes the interaction resulting from bond stretching for a pair of bonded particles, whose reference distance is chosen to be l_i^{ref}. The interaction in this case is harmonic, and thus it is best justified for small fluctuations. As a remark, note that l_i^{ref} is not the equilibrium distance because that depends on thermodynamic conditions as well as the composition of the system. Rather, l_i^{ref} should be regarded as the distance that two particles will adapt to in the absence of any other interactions in a system. The second term in Eq. (3) is the (harmonic) bending interaction for three consecutive particles (A–B–C) in the same chain-like piece of a molecule, with the reference value of the valence angle being θ_i^{ref}. The torsional interaction presented third is a four-body term, whereas a short-range Lennard–Jones 6–12 potential is often used to describe the steep repulsion caused by the Pauli-exclusion principle as well as van der Waals interactions resulting from dispersion forces. Finally, there is an electrostatic term for long-range Coulombic interactions for a pair of charged particles whose distance from each other is r_{ij}.

For clarity, let us emphasize here that, to define a force field, one must specify not only the set of potential energy functions in a force field but also the force-field parameters (as well as other practical details such as cutoff distances used in truncating long-range interactions) used in the calculations. If one of the parameters is changed, then the force field is also changed.

Although Eq. (3) provides us with a simple but typical example of a force field used for biomolecules, it allows us to pose a number of relevant questions. How are the force fields determined? Where do the parameters in the force fields come from? How can one validate a force field?

Using Eq. (3) as an example, each of the terms has a clear physical meaning, and the parameters associated with a given term can be determined either from theory or from experiments. The theoretical approach is essentially based on quantum mechanical calculations and is often used to, for example, determine the distribution of partial charges within a molecule. Unfortunately, however, this approach is of limited use because the fitting is not easy to perform and depends critically on the quality of the QM computations, such as the basis set used in calculations. Thus, it is common to calibrate parameters in force fields by fitting them to empirical data.

The parameters in the bond stretching and bending interactions, for example, can be determined by spectroscopic techniques, and the parameters in the van der Waals potential can be derived by fitting parameters to crystal structures. Anyhow, even if the parameter fitting process is done very carefully, there is no point in claiming that the obtained force field is perfect. The number of different combinations of parameters is huge, and slight weaknesses in some of them can be compensated for by adjustments in others. Therefore, it should not be surprising to realize that there are many different force fields that are commonly employed in the field of biomolecular modeling and simulation. For one reason or another, each of them seems to have pros and cons that limits their use. To our knowledge,

CHARMM (chemistry at Harvard macromolecular mechanics) force fields, for example, serve well for simulations of a number of DNA structures [74]. The force fields used in Gromacs [75] (see: http://www.gromacs.org), in turn, are likely a better choice for modeling of lipids, whereas their performance in DNA simulations could be better. In addition to these, many other force fields are in common use. AMBER, OPLS, NAMD, and MMFF are just some of the many force fields available (see Ref. [57]).

Hence, we think it is fair to say that all force fields have certain deficiencies. They are largely based on intuition and the validity of QM and experimental results used to fit the force-field parameters. Further, all force fields are based on some set of potential energy functions, whose validity can be questioned. The functional forms of the potential energy functions can be very complicated, as is the case in the MM4 force field, allowing, for example, accurate predictions of molecular structures. In contrast, one can employ simpler functional forms, which makes it more feasible to model large molecules such as proteins or DNA. Typical force fields in this category are AMBER and CHARMM. There is an endless need for improvements of force fields as comparisons among their results with respect to experimental observations reveal deviations. In the meantime, we can use current force fields rather safely. Even if their predictions are not completely correct quantitatively, they can provide us with plenty of qualitative insight into both structural and dynamic properties of complex biological systems.

Although the above discussion is definitely short, an extensive discussion of the topic can be found from a number of textbooks and recent review articles (see Refs. [57, 62]).

2.3.3. Dynamics in Classical Techniques

Although a force field is a crucial part of any molecular model, we need some means to find the equilibrium structure of the system. This can be done in various ways. One can apply energy minimization techniques or the MC simulation technique [60, 61], for example. These methods use the force field as an input and eventually provide one with structures corresponding to equilibrium conditions. However, as they cannot tell much of the dynamics in a system, we shall not consider them any further in this chapter. Instead, we discuss an approach that is the one most commonly used in (classical) atomic-level studies of biomolecular systems: the MD technique [60, 61].

MD is highly popular because of its simplicity and physical appeal. In MD, one essentially calculates forces acting on particles and then uses Newton's equations of motion to update their positions and velocities.

For simplicity's sake, let us consider a system of N particles in the microcanonical *NVE* ensemble (constant volume and constant energy without dissipation). Assuming that we know the force field of the system, we can calculate the forces acting on the particles. Let the force acting on particle i at time t be $\vec{F}_i(t)$. Then, in the spirit of Newton, we know that $\vec{F}_i(t) = m_i \vec{a}_i(t)$, where m_i is the particle's mass and $\vec{a}_i(t)$ is the acceleration resulting from the force. In the simplest possible level, one can now calculate the new position and the velocity of the particle at time $t + \delta t$ using a Taylor expansion about time t:

$$\vec{v}_i(t + \delta t) = \vec{v}_i(t) + \delta t \, \vec{a}(t) \tag{4}$$

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \delta t \, \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) \tag{5}$$

In principle, that is it. Once we have written down the force field, we can construct a system of, say, DNA attached to a lipid membrane surrounded by water and consider its dynamics through MD simulations by repeatedly solving the Newton's equations of motion.

However, as expected, there are a number of issues that might lead to problems. In addition to the trouble of finding a reliable force field for the system in question, one is bothered by various practical problems that are an inherent part of MD simulations. First of all is the integration of equations of motion. Instead of using Eq. (4), which is far too simple for practical purposes, one should apply an integration scheme that is both accurate and efficient. In practice, we need an accurate scheme that also allows the use of a large time

step δt. Although there is a wide variety of different integration schemes available, so-called symplectic integrators have turned out to be appropriate in many cases [70].

Second is how to deal with long-range interactions in a small (periodic) system. This question is very relevant because it has been shown that certain schemes for long-range interactions can lead to major artifacts in MD simulations of soft matter systems (see discussion in Section 4.1). At present, it seems like the Ewald-type summation schemes [60, 76] are most appropriate for this purpose.

Third, as experiments are typically done under conditions in which the temperature and some other thermodynamic variables are fixed, we have to ask ourselves how we can establish such conditions in MD simulations. This question with regard to different ensembles for different physical conditions and as regards ways to handle this issue in practice is discussed thoroughly (e.g., in Refs. [60, 61]). Further, as many soft matter systems are characterized by a wide range of different timescales such as the motion of bond stretching compared to the relative motion of two particles interacting through a long-range Coulombic potential, one should ask how to efficiently deal with this situation. The multiple-time step schemes discussed in Refs. [60, 77] provide one solution to this issue. Another and even more important issue is related to living systems in general, as they are basically never in equilibrium. If they are, they are dead. Thus, we need reliable ways to model biological soft matter systems under nonequilibrium conditions. The field of nonequilibrium MD aims to address related questions, and thanks to major activities in this challenging field [78], we are today at a stage at which many of the interesting research problems have been clarified, at least in part.

Processes in biologically relevant soft matter systems often take place over very large length and timescales, which puts some pressure on computational resources. Simulations in a parallel environment are therefore a rule rather than an exception, which leads to a number of practical questions, including, at the moment, how to deal with long-range interactions and random number generation in a parallel environment. Some of these issues will be discussed in more detail in Section 4.

For the time being, let us stress that the above issues can be a burden, but they can be (and in many cases have been) solved. Thus, classical atomic-level MD has reached a level at which it can be considered as a mature technique and can be applied to numerous intriguing problems in a wide range of fields. MD simulations of DNA [79, 80], protein systems [80, 81], protein folding [82, 83], and lipid membranes [84–88] are examples of this activity. Further examples will be discussed in Section 5.

3. COARSE-GRAINING: MOVING BETWEEN DIFFERENT TIME- AND LENGTH SCALES

The previous section summarized recent developments in molecular-level modeling of soft matter and biological systems. As mentioned in Section 1, the Holy Grail of modeling is to develop rigorous methods for linking different time- and length scales, such that both coarse-graining and fine-graining are well defined. In the following text, we will briefly review some of the new methods and provide a quick summary of their merits and problems. Of the specific methods, we provide a more detailed look into the dissipative particle dynamics in Section 3.2 and field theoretical methods in Section 3.3, as they represent two different philosophies (i.e., particle-based vs. continuum approaches; see Fig. 1).

As discussed earlier, classical MD simulations of biologically motivated systems are typically limited to system sizes not much larger than 10^4 atoms and timescales spanning up to around 100 ns in the best cases. The restrictions may be even worse if electrostatic interactions must be properly accounted for. In addition, for systems such as polymer melts, the computationally accessible time- and length scales are simply far too short for the system to be able to reach equilibrium as dynamic processes during equilibration occur under hydrodynamic conditions.

In the Section 1, we mentioned that polymer research has been one of the leading fields in multiscale modeling in both practical and theoretical aspects. This is indeed the case, and there are a lot of different approaches that have been used. On the analytical side the projection operator-based formalism of Akkermans and Briels [6] provides an interesting

approach to study the emergence of fluctuating forces in a more coarse-grained description. The projection operator formalism has also been used successfully (although in a very different manner) in the GENERIC approach [25–27], which provides an analytically rigorous method for coarse-graining. In addition to the projector operator formalism that has proven to be fruitful, another commonly used analytical approach relies on the Ornstein–Zernike equation [89] and the hypernetted chain closure [90, 91]. We will not discuss those approaches here any further, but refer the reader to the above references and the references therein.

Inverting the radial distribution functions $g(r)$ to obtain pair potentials is another possible starting point. The theoretical basis for that was given by Henderson [92], who stipulated that under fixed conditions two pair potentials that give rise to the same $g(r)$ cannot differ by more than a constant. This constant is determined by the condition for the pair potential that

$$V(r \to \infty) \to 0, \tag{6}$$

where r is the interparticle distance. In this respect, this approach is analogous to the Hohenberg–Kohn theorem [93], which states that all ground-state properties are determined by the electron density. This theorem lies at the heart of modern DFT. It is important to notice that the radial distribution function obtained from a simulation includes effects from the many-body interactions. Furthermore, this way it is possible to define new interaction sites and to compute the radial distribution function between them, and thus readily obtain new coarse-grained models at different levels of description.

There are new simplex-algorithm based optimization procedures developed by Müller-Plathe et al. [94–96] that exploit the above described relation. For an overview of the current coarse-graining approaches in the field of polymer science, see, for example, Ref. [29, 32] and references therein.

As a conceptually simple approach, the DPD [18–20] have recently gained lot of popularity in soft matter simulations. We will discuss that approach in detail in the following text. Other coarse-grained approaches that take hydrodynamics properly into account include the lattice Boltzmann method (see, e.g., the article of Pagonabarraga in Ref. [28]) and a novel and very promising approach developed very recently by Malevanets and Kapral (MK) [12, 13]. The MK method couples a molecular-level description with a mesoscale treatment of solvent-conserving hydrodynamics. This approach is particularly appealing for studying dilute systems with hydrodynamics. Concurrently, it has been applied to a number of problems, including studies of dilute polymer systems [97], individual colloids under flow [13], diffusion of colloids [98], and crystalline clusters [99]. Malevanets and Yeomans have further developed a variant of the MK method and applied that to study structural and dynamical properties of individual polymer chains in a hydrodynamic medium [100, 101]. Recent developments of the technique are discussed in Refs. [102–107].

Before moving on to describe a few of the methods in detail, we would like to mention, as a detailed discussion is out of scope of this review, that there are also other types of developments. The coupling of quantum and classical systems is a new and exciting field and can open entirely new avenues in biomolecular simulations. We refer the reader to the article by Kalibaeva and Ciccotti in Ref. [28] for an in-depth discussion. Algorithmic developments are an important issue, and we refer the reader to one of the new textbooks in the field (e.g., the book by Leach [62]).

3.1. Effective Interactions: Inverse Monte Carlo

Conservative forces can consist of various kinds of forces [108] resulting from, for example, electrostatic interactions, as well as descriptions of detailed intermolecular interactions such as van der Waals forces. As suggested earlier, it is possible to obtain conservative forces by coarse-graining the pairwise interactions from an MD simulation using, for example, the Inverse MC (IMC) procedure [109] or the other procedures [31, 32] that are similar in spirit.

The IMC method bears a close resemblance to the Renormalization Group MC method (see, e.g., the book by Goldenfeld [110]), in which one iteratively solves the renormalization

constants. In IMC, one inverts the radial distribution functions, calculated in atomistic MD simulations, to obtain the effective potentials for a coarse-grained model with a fewer number of degrees of freedom. The essential features of the IMC procedure were introduced by Lyubartsev and Laaksonen in 1995; Ref. [109] contains the relevant computational details.

There are two aspects that we wish to point out. First, the effective potential includes corrections from many-body interactions to the well-defined potential of mean force (PMF) [89], which is defined as

$$V^{\mathrm{pmf}}(r) \equiv -k_{\mathrm{B}}T \ln g(r) \tag{7}$$

where r is the interparticle distance and $g(r)$ the pair correlation function. Therefore, the effective potential considered here is not the same as the PMF as defined by Eq. (7). Inclusion of the many-body corrections is the reason why the iterative IMC scheme is needed. Second, this procedure guarantees self-consistency; that is, the effective potentials, when used in mesoscale (say, DPD simulations), lead to the same pair correlation behavior as the underlying MD simulations. Thus, one can expect that certain response functions such as compressibility are properly described by the effective potentials. This method can be applied to studies of soft matter systems with full hydrodynamics.

3.1.1. The IMC: An Example of Inversion Methods

Next we will sketch the outline of the IMC procedure. As mentioned, there are other methods for using structural information to obtain interaction potentials. We have chosen to use the IMC because it will be used to discuss a specific example later in Section 6.

Let us assume a simple system consisting of identical particles with pairwise interactions only. Then we can write the Hamiltonian as

$$H = \sum_{i<j} V(r_{ij}) \tag{8}$$

where $V(r_{ij})$ is the pair potential and r_{ij} is the distance between the interaction sites i and j. Let us further assume that we do not know $V(r_{ij})$ but we know the radial distribution function $g(r_{ij})$, for example, from another simulation or experiments.

To obtain the interaction potential $V(r_{ij})$, we discretize the Hamiltonian such that the interval $[0, r_{\mathrm{cutoff}}]$ is divided into N grid points. We denote this discretized potential by V_α such that $\alpha = 1, \ldots, N$. With this discretization, we can approximate the the Hamiltonian in Eq. (8) by

$$H_{\mathrm{discr}} = \sum_\alpha V_\alpha S_\alpha \tag{9}$$

Here, we have introduced S_α as the number of pairs that have their interparticle distance within the interval α. Recalling the definition of the radial distribution function immediately tells us that S_α is related to it through

$$\langle S_\alpha \rangle \sim 4\pi r^2 g(r_\alpha) N^2 \tag{10}$$

Because S_α is related to the radial distribution function, and thus also the discretized potential V_α, it is possible to write an expansion for S_α in terms of the discretized potential,

$$\Delta \langle S_\alpha \rangle = \sum_\gamma \frac{\partial \langle S_\alpha \rangle}{\partial V_\gamma} \Delta V_\gamma + \mathcal{O}(\Delta V^2) \tag{11}$$

Lyubartsev and Laaksonen [109] showed that the partial derivatives can be written in a compact and computationally tractable way, thus allowing one to use the above expansion for iteratively solving the discretized potential V_α.

To solve for V_α in practice, one has to perform regular MC simulations of the system to obtain $\langle S_\alpha \rangle$s. Then, solving V_α iteratively using Eq. (11), one obtains the interaction potential. The procedure resembles solving a multidimensional nonlinear equation using the Newton–Raphson method. The most natural starting point for the procedure is to use the potential

of mean force as defined above. It is also worth noticing that this procedure provides us with estimates of the uncertainty in $V(r)$. That can be formally done by eigenvalue analysis.

In Section 6, together with the DPD method, we will apply the IMC procedure to an aqueous sodium chloride solution.

3.2. Dissipative Particle Dynamics (DPD)

DPD was originally developed by Hoogerbrugge and Koelman to solve hydrodynamic flow problems in porous media. Since its introduction in 1992 [19, 20], DPD has become the most used coarse-grained simulation method in soft matter research, and several extensions and generalizations of DPD have been developed [21, 111–114]. Instead of porous media, DPD has gained most of its merits in simulations of polymers and surfactant systems [115–117] and, very recently, also in simulations of biological systems [10, 118–122]. To mention some examples, DPD has been applied to problems ranging from liquid-gas phase diagrams [123], to rupture of bacterial membranes [10], to self-assembly in Huntington's disease [122], and to packing of fillers in composites [124].

DPD is a coarse-grained method, and this is reflected in its interactions potentials, which are "soft" in contrast to the Lennard–Jones-type potentials used in classical MD. This is also where the strength and weakness of the standard (and the most used) version of DPD lies. By "softness" in potentials, we mean that the DPD potential has a finite value at zero particle separation (i.e., the Fermi exclusion principle is not accounted for). The standard DPD potential is shown in Fig. 3.

As can be seen from Fig. 3, the only parameters concerning the potential are the cutoff radius r_c and the amplitude of the potential, α_{ij}. That this indeed is a reasonable approximation was shown, although in a slightly different context, by Forrest and Suter [9], who studied polymeric systems and explicitly averaged over fluctuations to obtain effective interaction potentials between monomers. They demonstrated that when averaged over sufficiently long timescales, the averaged, or coarse-grained, potentials approach a quadratic form: this is the form used virtually exclusively in DPD simulations. The underlying microscopic potentials were of the usual Lennard–Jones type. We will present the functional form of the DPD potential in the following subsection.

Despite its success, DPD does not come without problems and open questions. We will discuss them in the following sections. We also briefly review the latest developments concerning the inclusion of electrostatic interactions in DPD as well as novel and appealing coarse-graining procedures, providing a more solid basis for DPD. The latter are an example of true systematic coarse-graining linking MD to continuum representation.
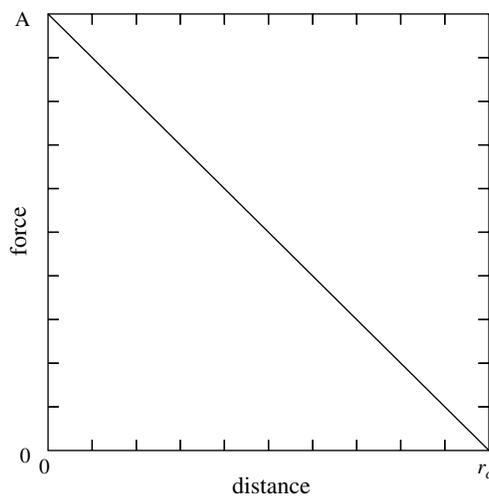
Figure 3. The conservative force in the "standard" DPD is linear with the cutoff distance r_c and the amplitudes of the pairwise interactions being the only parameters.

3.2.1. DPD Equations of Motion

Here we summarize the DPD equations of motion and discuss some aspects related to it. After that, we put the method in a more general context. For algorithmic details, a comprehensive discussion of the performance, implementation, and quality of various DPD integrators, we refer the reader to Refs. [125–127] and the references therein. For a more application-oriented and detailed review by one of the pioneers of the field, please see the article by Groot in Ref. [28].

One of the key ideas and motivations behind DPD is that it conserves hydrodynamics (i.e., all the interactions are pairwise conservative). In DPD the total pair force consists of three components; dissipative, random, and conservative. Thus, the force exerted on particle i by particle j can be written simply as

$$\vec{F}_{ij} = \vec{F}_{ij}^D + \vec{F}_{ij}^R + \vec{F}_{ij}^C \tag{12}$$

We start the discussion from the dissipative and random forces. In the standard notation, they are given as

$$\begin{aligned}
\vec{F}_{ij}^D &= -\gamma \omega^D(r_{ij})(\vec{v}_{ij} \cdot \vec{e}_{ij})\vec{e}_{ij} \quad \text{and} \\
\vec{F}_{ij}^R &= \sigma \omega^R(r_{ij})\xi_{ij}\vec{e}_{ij},
\end{aligned} \tag{13}$$

where $\vec{r}_{ij} \equiv \vec{r}_i - \vec{r}_j$, $r_{ij} \equiv |\vec{r}_{ij}|$, $\vec{e}_{ij} \equiv \vec{r}_{ij}/r_{ij}$, and $\vec{v}_{ij} \equiv \vec{v}_i - \vec{v}_j$ for particles with positions $\vec{r}_i$ and velocities $\vec{v}_i$. The ξ_{ij} are symmetric Gaussian random variables with zero mean and unit variance and are independent for different pairs of particles and different times. The pairwise conserved fluctuations can be easily justified, as two particles being close enough to each other experience the same fluctuations.

Equations (13) contain two amplitudes γ and σ and two weight functions $\omega^D(r_{ij})$ and $\omega^R(r_{ij})$. The amplitudes define the strengths, and the weight functions the spatial extent, of the dissipative and random forces. The specific functional form of the weight functions is not specified, and any physically motivated choice can be used. In contrast to this freedom concerning the functional form, the two weight functions, as well as the amplitudes, are coupled via a fluctuation–dissipation relation.

The coupling of the dissipative and random forces $\vec{F}_{ij}^D$ and $\vec{F}_{ij}^R$ is the result of the fact that the thermal heat generated by the random force must be balanced locally by dissipation. The precise relationship between these two forces is determined by the fluctuation–dissipation theorem [18], which sets the following two conditions for both the weight functions and the amplitudes of the forces in Eqs. (13),

$$\omega^D(r_{ij}) = [\omega^R(r_{ij})]^2 \quad \text{and} \quad \sigma^2 = 2\gamma k_B T^*, \tag{14}$$

where T^* is the canonical temperature of the system.

The most common choice for the weight functions ω^D and ω^R is the soft-repulsive form

$$\omega^R(r_{ij}) = \begin{cases} 1 - r_{ij}/r_c & \text{for } r_{ij} < r_c; \\ 0 & \text{for } r_{ij} > r_c, \end{cases} \tag{15}$$

where r_c is the cutoff distance and $\omega^D(r_{ij})$ is given by the fluctuation–dissipation relation above. It is somewhat surprising that, at least to the authors' knowledge, no other forms for the weight function have been used in published literature.

Thus far we have not said anything quantitative about the conservative force. To be precise, the DPD formulation does not specify its functional form. In principle, one is free to use any physically motivated choice, van der Waals forces, electrostatic interactions, or some coarse-grained interactions. The term DPD often refers to above formulation for random and dissipative forces, combined with the soft repulsive potential given as

$$\vec{F}_{ij}^C = F_{ij}^{(c)}(r_{ij})\vec{e}_{ij}, \tag{16}$$

where $F_{ij}^{(c)}(r_{ij}) = \mathscr{A}\omega(r_{ij})$. The parameter $\mathscr{A}$ is the amplitude of the force, and it can be defined via compressibilities or using Flory theory for polymers. For a detailed discussion, see the article by Groot in Ref. [28]. The force and the potential are plotted in Fig. 3.

What is usually referred to DPD is defined by Eqs. (13), (15), and (16) together with fluctuation–dissipation in Eq. (14). This definition is rather restrictive, as the dissipative and random forces are independent of the conservative force. This fact has been recognized by many authors, and it is appealing to think of DPD as a momentum-conserving thermostat [11, 111, 126, 128], as we will discuss in the next section.

The above formulation is the universally accepted and exclusively used form of DPD derived by Español and Warren [18] using a Fokker–Planck equation. It is also worth pointing out that the original DPD formulation by Hoogerbrugge and Koelman does not obey the fluctuation–dissipation theorem and hence does not relax to the canonical distribution, as required by thermodynamical equilibrium.

3.2.2. DPD as a Thermostat

As already mentioned in the previous section, DPD can be defined as a momentum-conserving thermostat for MD simulations. Earlier we defined DPD as a momentum-conserving coarse-grained MD using conservative potentials as given in Eq. (16). That is indeed the standard phenomenology, but it would be more appropriate to use the term DPD for a momentum thermostat defined by Eqs. (13) and (14) and the weight functions, for example, by Eq. (13).

These aspects have been recently studied in detail by Vattulainen et al. [126, 127] for general Lennard–Jones systems, and by Soddemann et al. [128], who showed the usefulness of the DPD thermostat also in nonequilibrium MD simulations.

Another problematic aspect with DPD has been the treatment of electrostatics. As it is easy to understand from the softness of the potentials, electrostatics needs special attention. This issue has been resolved by Groot [129, 130]. The idea is that the charges are smeared out and then the electrostatic field is solved in a grid. From the computational point of view, this approach is close to the multigrid method introduced by Sagui and Darden [131, 132]. Groot was also able to show that his approach is consistent with the Stillinger–Lovett sum rules, and thus it provides a sold basis for treating the electrostatics properly. It is likely that this will increase the usefulness of DPD in biomolecular simulations.

3.2.3. An Alternative: The Lowe–Andersen Thermostat

A different formulation for DPD was given by Lowe [133]. His approach is based on the thermostat developed in 1980 by Andersen [134]. The Andersen thermostat is different from other MD thermostats, as in this approach the velocities are periodically drawn from a Maxwell distribution. This formulation is computationally quite efficient, although it does not conserve momentum. Lowe [133] used the Andersen thermostat as a starting point and devised a momentum-conserving thermostat out of it. The idea here is that one chooses a parameter that describes the frequency (Γ in Lowe's original notation) for a particle to interact with the heat bath. Then, $\Gamma\delta t$, where δt is the time step, defines a probability for thermalization. This thermalization is done in a pairwise fashion, such that the method is momentum conserving, and it also produces the correct canonical distribution. One of the main advantages of this method is that it enhances viscosity and allows proper tuning of the Schmidt number, which is the ratio between kinematic viscosity and the diffusion coefficient.

Lowe's approach is appealing: There are no dissipative forces, and thus Lowe's approach is easier to use and performs well. The thermalization rate may be varied over a wide range, which implies that the dynamical properties of the system can be tuned in a controlled fashion. It is rather surprising, though, that in addition to Lowe's original article, there is only one other published work using this method [126].

3.2.4. A DPD Example: Liposome Formation

To give a brief illustration of a DPD simulation using the soft potentials as defined in Eq. (16), we have studied vesicle formation and dynamics [135]. The lipid molecules consist

of eight DPD particles, of which two belong to the headgroup, and the rest are divided equally between the two tails. The system illustrated in Fig. 4 has a total of 500 lipids and 188 000 waters (i.e, 228 000 DPD particles in total).

The amplitudes of the conservative force [i.e., the parameters $\mathscr{A}$ in Eq. (16)] between different species are shown in Table 1. In addition, the description of lipids included a harmonic freely jointed spring.

The simulation started with the formation of a stable bilayer. After that, the simulation was halted, allowing us to add a major amount of excess water around the bilayer. On this happening, the system was simulated under *NVT* conditions such the bilayer had to deform to a vesicular shape. This process and the formation of a spherical bilayer structure is clearly visible in Fig. 4. We are currently studying the dynamics of related vesicle systems.

3.2.5. Analytical Approaches: Systematic Derivation of DPD

The standard DPD as presented above is phenomenological. A very interesting formal approach has recently been presented by Flekkøy et al. [8, 21], who were able to formally link DPD to molecular level properties. They use a Voronoi tessellation–based technique to establish a systematic coarse-graining procedure that couples molecular dynamics to DPD and provides a systematic connection between continuum and particle representations. The great advantage that this method has is that it can be used to resolve different length scales simultaneously. The method is formally akin to the well-known renormalization group procedure extensively used in analytical treatment of critical phenomena. As such, this coarse-graining procedure provides a true systematic and tunable method. For details about the renormalization group procedure, see, for example, the book by Goldenfeld [110]. The mathematical details can be found from the original articles as cited above.

The approach of Flekkøy et al. provides a systematic and computational tractable, although not straightforward, method to coarse-grain molecular systems. From a computational point of view, their method is more demanding than particle-based methods such as the standard DPD, as the Voronoi tessellation procedure requires $N \log N$ operations and it comes with fairly large additional prefactor.

Although not connected directly to DPD, we mention here another general method called GENERIC, introduced by Grmela and Öttinger [25–27]. The idea is that there is a general form for the time-evolution of nonequilibrium systems and that it can be written in a general form

$$\frac{dx}{dt} = L(x)\frac{\delta E(x)}{\delta x} + M(x)\frac{\delta S(x)}{\delta x}$$
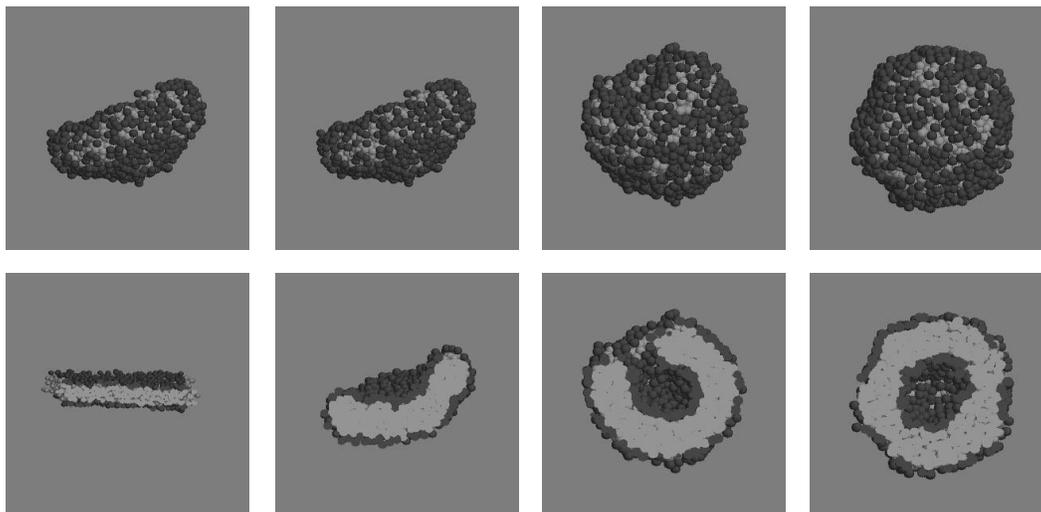
Figure 4. Formation of a vesicle in a DPD simulation [135]. The surrounding 188 000 water molecules have been removed for clarity. The lower panels show the corresponding cut through the middle of the vesicle during the formation process.

Table 1. Amplitudes of conservative interactions used in the present dissipative particle dynamics simulations for liposome formation.

	Head	Tail	Water
Head	100	160	100
Tail	160	100	160
Water	100	160	100

where x characterizes the state of the system, $L(x)$ is an antisymmetric matrix, and $M(x)$ is a symmetric and positive definite matrix. They are connected to the second and the first law of thermodynamics, respectively. The terms $E(x)$ and $S(x)$ are functionals for the total energy and entropy, respectively.

The equation above is the so-called GENERIC form, which uses the Poisson bracket and projection operator formalism to connect different levels of description. The basic ingredients of the theory are the energy and entropy functionals, which are the generators for reversible (energy) and irreversible (entropy) dynamics and the corresponding matrices.

The mathematical details are somewhat lengthy, and we refer the reader to the original articles [25–27] and the review by Español in Ref. [28], which discusses analytical approaches to coarse-graining at length.

3.3. Field Theoretical Methods

In contrast to the particle-based coarse-grained methods, the central idea of phase-field modeling is to provide a continuum description of a system using general symmetry principles, conservation laws, order parameters, and a free-energy functional. The name *phase-field* was coined by G. J. Fix [136] in 1982, but the approach itself is older [137]. This approach has several attractive features. First, it operates in meso- and macroscales, thus enabling studies of time- and length scales that would otherwise be difficult to reach. Second, it is suitable for describing out-of-equilibrium systems and interfaces between different phases that appear naturally in this approach. Third, it is possible to use projection operator techniques to study these systems using analytical theory.

Phase-field models have been used extensively in connection with problems involving nucleation and growth and spinodal decomposition (see, e.g., Ref. [138, 139]). In materials science, it has been used to study the properties of such diverse systems as directional solidification [140], morphological instabilities [141], eutectics [142], and charge-density waves [143], to mention a few examples. Surprisingly, the phase-field approach has received very little attention in soft matter problems, although it has great potential addressing questions related to macroscopic behavior. For details of phase-field models in a general context, we refer the reader to the recent article by Ala-Nissila et al. [22]. In the following text, we will describe the physical basis of this approach.

3.3.1. Landau Theory of Continuous Phase Transitions

The phase-field approach is based on Landau's theory of phase transition. Thus, the system is defined by an order parameter reflecting local order in the system. It can be, for example, a density difference between different liquids. From that it is clear that the order parameter is by definition a coarse-grained quantity. As such it does not include any information about the molecular details of the system. For the Ising model, one can perform a formal mapping to obtain the order parameter, which in that model is the magnetization of a block.

The Landau theory of continuous phase transitions is phenomenological and stresses the importance of overall general symmetry properties and analyticity over microscopic details in determining the macroscopic properties of a system; an excellent discussion is given in Ref. [110]. The validity of this approach is based on the following assumptions.

First, it is possible to define an order parameter Ψ, that characterizes the order in the system in the following way: $\Psi = 0$ in the disordered state (above the critical temperature T_c), and Ψ is small and finite in the ordered state (below T_c).

Second, it is possible to describe the system with a free energy. Close to T_c, it can be expanded in powers of the order parameter; that is,

$$F(\Psi) = \sum_{n=0}^{\infty} a_{2n} \Psi^{2n} \tag{17}$$

where F is the free energy and the expansion coefficients are phenomenological parameters that depend on temperature and the microscopic properties of the system under consideration. Therefore, at least in principle, the a_{2n} can be derived from first principles. The smallness of the order parameter guarantees that the expansion converges and can be truncated at a finite power.

Third, the expansion, Eq. (17), of the free energy must be consistent with the high temperature symmetry properties of the system under consideration. Because the Landau theory describes continuous transitions, no odd powers are allowed in the expansion of the free energy. When the system is cooled down to $T < T_c$, there is a spontaneous symmetry breakdown. This means that the ground state of the system has lower symmetry than the free energy shows: States $+\Psi_0$ and $-\Psi_0$ are equivalent, but the system must select one of them (see Fig. 5). Mathematically speaking, the Hamiltonian must commute with the symmetry group of the high-temperature phase.

Fourth, the free energy must be analytic. In addition to analyticity in Ψ, it is also required that the expansion coefficients be regular functions of the temperature. Because the second-order term is dominant, $a_2(T)$ must vary smoothly from $a_2(T) < 0$ for $T < T_c$, to $a_2(T) > 0$ when $T > T_c$, with $a_2(T = T_c) = 0$. This implies that $a_2(T) \sim (T - T_c)$. To see this, we first minimize the free energy with respect to Ψ,

$$\frac{\partial F}{\partial \Psi} = 2a_2\Psi + 4a_4\Psi^3 = 0 \tag{18}$$

The solutions are

$$\Psi_0 = 0 \qquad \text{and} \qquad \Psi_0 = \sqrt{-\frac{a_2}{2a_4}} \tag{19}$$

The first is the high-temperature solution, and the latter provides the minima for $T < T_c$. From Eq. (18), we can see that a_4 must be greater than zero, as otherwise $\Psi \to \infty$ would minimize the free energy, rendering the above construction useless. Next, we Taylor expand the expansion coefficients around T_c:

$$a(T) \approx a(T_c) + (T - T_c)\frac{\partial a(T)}{\partial T}\bigg|_{T=T_c} + \frac{1}{2!}(T - T_c)^2\frac{\partial^2 a(T)}{\partial T^2}\bigg|_{T=T_c} + \cdots \tag{20}$$

and

$$a_4(T) \approx a_4(T_c) + (T - T_c)\frac{\partial a_4(T)}{\partial T}\bigg|_{T=T_c} + \frac{1}{2!}(T - T_c)^2\frac{\partial^2 a_4(T)}{\partial T^2}\bigg|_{T=T_c} + \cdots \tag{21}$$
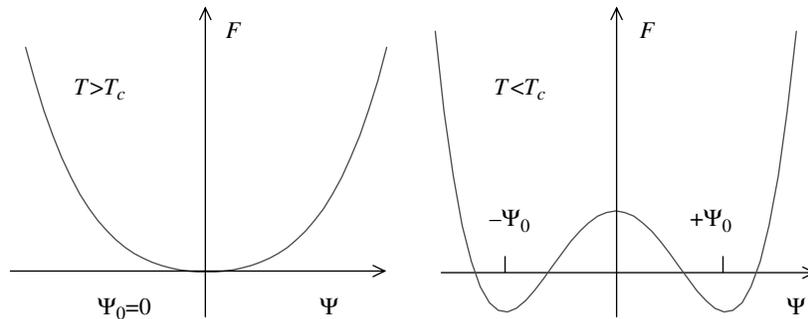
Figure 5. Schematic picture of the Landau free energy. Upon cooling below T_c the system undergoes a spontaneous symmetry breaking.

As argued above, $a_4(T) > 0$, and therefore we can approximate $a_4 = a_4(T_c) = $ const. This requirement is enough to guarantee the finiteness of the order parameter. Then, $a_2(T)$ must change its sign on cooling below T_c. Then,

$$\left.\frac{\partial a_2}{\partial T}\right|_{T=T_c} = \text{const.} > 0 \tag{22}$$

and $a_2(T) \sim (T - T_c)$.

The Landau theory is a mean-field theory because it does not take into account spatial inhomogeneities or thermal fluctuations. To account for inhomogeneities, we have to let the order parameter to become space dependent [i.e., $\Psi \equiv \Psi(\vec{x})$]. Conceptually, we should think of Ψ as a coarse-grained order parameter; that is, Ψ is defined only over a certain length scale. We must define a short wavelength (ultraviolet) cutoff for Ψ in such a way that Ψ varies smoothly in space; the order parameter cannot fluctuate on smaller length scales than the cutoff. In practice, the cutoff may often be thought of as the lattice spacing.

The response of the system to spatial inhomogeneities can be described as internal rigidity; Anderson [144] provides an in-depth discussion about the emergence and importance of rigidity in physical systems. Physically, it is clear that large differences between neighboring points are unfavorable. In other words, rigidity is a generic property of the system, and therefore there must be an energy cost associated with spatial inhomogeneities. Because the order parameter is a smooth and slowly varying function, we can take the rigidity of the system into account by making a gradient expansion and retaining only the lowest-order term compatible with the symmetry properties. The validity of the truncation of the gradient expansion depends on the smoothness and slowness of the order parameter variations.

Because the order parameter is now a local variable, the free energy in Eq. (17) becomes a functional of Ψ,

$$\mathscr{F}[\Psi] = \int d^d\vec{x}\left[F(\Psi) + \frac{1}{2}K(\nabla\Psi)^2\right] \tag{23}$$

where K is a phenomenological parameter describing the rigidity of the system and $F(\Psi)$ is the free energy defined in Eq. (17). Equation (23) is also known as the Ginzburg–Landau–Wilson free energy. Clearly, in Eq. (23), K must be positive for the free energy to be bounded from below.

3.3.2. Dynamical Models

The above description does not include time dependence. How time dependence can be added is to use linear response theory and assume relaxational dynamics (i.e., the system is dissipative and it is driven to equilibrium; see Fig. 5). The resulting models are often described as models A, B, C, and so on, depending on their symmetry properties, the dimensionality of the order parameter, and the conservation laws [137].

Being the simplest, model A serves as an example. It is a nonconserved system, and the equation of motion can be given as

$$\frac{\partial \Psi(\vec{x}, t)}{\partial t} = -\Gamma\frac{\delta\mathscr{F}[\Psi(\vec{x}, t)]}{\delta\Psi(\vec{x}, t)} + \eta(\vec{x}, t) \tag{24}$$

where we have included thermal fluctuations as Gaussian random noise, with the first and second moments defined as

$$\langle\eta(\vec{x}, t)\rangle = 0 \tag{25}$$

and

$$\langle\eta(\vec{x}, t)\,\eta(\vec{x}', t')\rangle = 2\Gamma k_B T\delta(\vec{x} - \vec{x}')\delta(t - t') \tag{26}$$

The angular brackets denote an average, and Γ is a kinetic coefficient describing the relaxation rate. An example of the evolution of a system described by model A is given in Fig. 6.
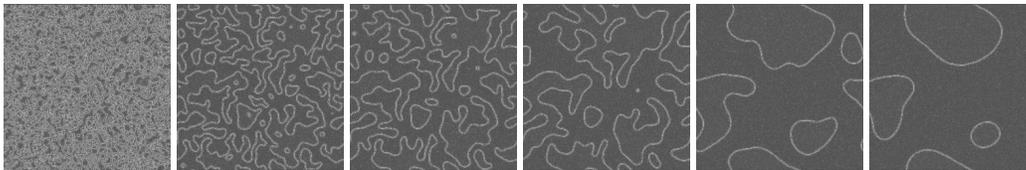
Figure 6. Phase separation in model A described by Eq. (24). The panels from left to right show the evolution of the system from random initial conditions towards equilibrium. Periodic boundary conditions are clearly visible.

Other models can be obtained by including conservation laws; for example, the dynamical equation of motion for a conserved order parameter is given by model B,

$$\frac{\partial \Psi(\vec{x}, t)}{\partial t} = \Gamma \nabla \left[\frac{\delta \mathscr{F}[\Psi(\vec{x}, t)]}{\delta \Psi(\vec{x}, t)} + \eta(\vec{x}, t) \right] \tag{27}$$

This can be derived using the continuity equation.

The system may have several order parameters that are coupled, thus leading to more complicated free energies and equations of motion; see Ala-Nissila et al. [22] and references therein.

It is rather surprising that this type of approach has, thus far, had only limited attention in soft matter and biological modeling. There are a few notable exceptions, though. Shore et al. [145, 146] coupled a phase-field model to the Navier–Stokes equation, including viscoelasticity to melt fracture in polymer extrusion experiments. The details of that model are beyond this review, but it suffices to note that the study of Shore et al. is a good example of the power of the phase-field approach, as it provides physical insight into (in this case, instabilities) both fundamentally and industrially important problems.

Other examples of the use of field theoretical models in soft matter include the OCTA software package, which in one of its parts uses self-consistent field theory from the Nagoya group [39, 40]. Self-consistent field theory has been used in the theoretical description of polymers for a long time. A very recent example of the developments in that field is the elegant study of tetrablock copolymers by Drolet and Fredrickson [147].

Another major effort to build modeling software around field-theoretical ideas is the MESODYN project of the Leiden group [24, 148]. They use the Ginzburg–Landau approach and DFT to build a systematic and computationally tractable system for polymer melts. A detailed description is again beyond this review, but with their approach it is possible to study phase separation even in three dimensional systems. This approach is closely related to model B of critical dynamics [137], as discussed above.

4. METHODOLOGICAL ISSUES IN MOLECULAR SIMULATIONS

One of the intriguing issues in science is indeed that there is the artistic side, too: The art of doing science. In the field of molecular simulations and computational sciences, this is largely related to the inventions and new ideas of seeing things done in an accurate but efficient manner. This brings us to the methodological side of doing molecular simulations. In this section, we consider a few aspects related to both atomistic and coarse-grained descriptions of molecular systems that illustrate the importance of developing the methodology.

We discuss three highly important methodological issues whose role for the reliability and accuracy of molecular simulations is particularly significant. First, especially for MD simulations of biologically relevant soft matter systems on atomic level, the treatment of long-range electrostatic interactions is a major issue. If this is not handled with care, the interpretation of simulation results may be very problematic because of the underlying artifacts caused by the mistreatment of electrostatics. Second, as all stochastic simulation methods are based on the use of noise produced by so-called pseudorandom number generators, the quality of pseudorandom number sequences is of prominent importance in all cases in which they are employed to generate the dynamics for the systems under study. Because pseudorandom

number generators produce deterministic rather than unpredictable sequences of numbers, the "randomness" of pseudorandom numbers is a very subtle issue and should never be taken for granted. Finally, we discuss an old issue that had been thought to be overcome several years ago: the artifacts resulting from integration schemes that yield the dynamics for systems governed by Newton's equations of motion. In contrast to classical MD, where this issue is well under control, in stochastic simulation techniques such as dissipative particle dynamics, the case is more subtle. Here we discuss most recent developments in this field and show how the problems can be overcome in an efficient and reliable fashion.

4.1. How to Handle Long-Range Electrostatic Interactions

Electrostatic interactions are an inherent part of all biologically relevant soft matter systems. For example, the behavior of water is to large extent driven by hydrogen bonds, in which electrostatics plays an important role. In the same spirit, the stability of proteins and DNA is largely based on hydrogen bonding. DNA, in particular, is a prominent example of the importance of electrostatics because it is probably the most charged molecule found in living systems. Other examples including the self-assembly and overcharging of DNA-lipid complexes [79, 80, 149–153] are therefore obvious and lead naturally to a wide class of biologically important systems composed of lipid molecules. The most important entities in this context are cell membranes that surround the cell. In addition to protecting it, cell membranes provide a heterogeneous, semipermeable, and fluidlike environment for numerous processes [154–158] facilitated by ion channels and other proteins embedded in or attached to membranes, thereby, for example, governing the transport of molecules and ions across a lipid membrane.

The importance of electrostatics in membrane systems can be demonstrated by a practically relevant example. Of the many techniques available for the introduction of genetic material into cells, many are based on nonviral delivery vectors composed of cationic liposomes. Understanding the properties of complexes composed of cationic membranes and DNA is therefore largely based on the understanding of the electrostatics associated with these systems. From the point of view of molecular simulations, it is hence crucial to treat long-range electrostatic interactions as accurately as possible.

Techniques available for this purpose are wide and include, for example, Ewald summation techniques [76] based on solving the Poisson equation for the electrostatic potential in a periodic system in a fashion in which all charged particles and their periodic images are accounted for systematically. The Particle–Mesh–Ewald (PME) technique [60, 159], specifically, has found increasing popularity in MD studies of soft matter systems. Other commonly used techniques include the fast multipole method [60, 160], the reaction-field technique, and truncation methods. The truncation method is particularly appealing because it leads to considerable reduction in the computational load. As a consequence, it is frequently used when computational requirements are substantial because of large system sizes or long timescales, which may be the case, for example, in studies of lipid–protein systems [161], self-assembly of lipids [162], and membrane fusion [163]. There is reason to emphasize, however, that the accuracy of electrostatics should not be sacrificed at the expense of speed.

The key point here is that the results based on MD simulations of biophysical systems depend on the scheme chosen for electrostatics. It has been shown that the truncation of electrostatic interactions may affect structural properties of systems like water [164–168], peptides [169, 170], proteins [171], and DNA [172, 173]. Here we demonstrate and discuss this issue in the context of lipid bilayers.

4.1.1. Model of a Dipalmitoylphosphatidylcholine Bilayer

We discuss this issue in a one-component lipid bilayer consisting of $N = 128$ dipalmitoylphosphatidylcholine (DPPC) molecules at 323 K (see Fig. 7); which is above the so-called main phase transition temperature T_m. The temperature T_m is characteristic to any lipid species and depends on various molecular details such as the headgroup and the length and saturation of lipid acyl chains. The main transition temperature characterizes the transition for the pure one-component lipid system between the high-temperature liquid-crystalline
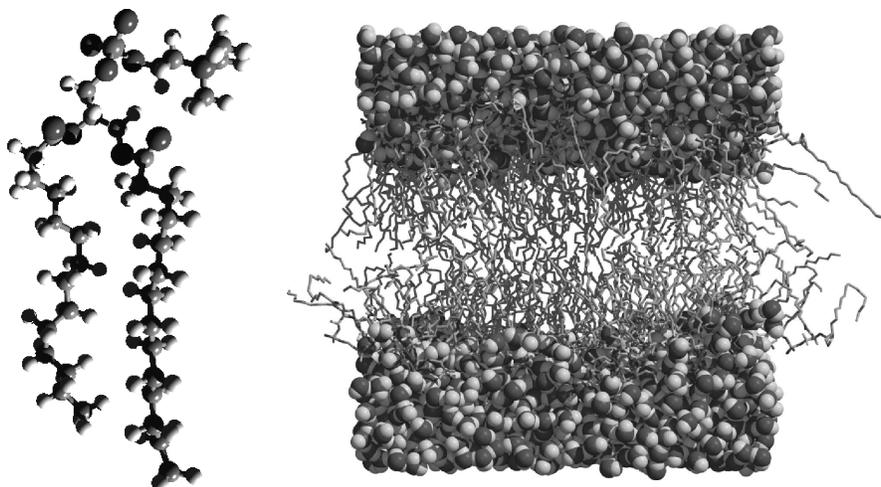
Figure 7. Representation of a saturated phospholipid molecule (left) and a lipid bilayer made of 128 dipalmitoylphosphatidylcholine and 3655 water molecules (right).

(also called liquid-disordered) phase and the low-temperature solid-ordered phase. The two phases are distinctly different in nature, as the high-temperature phase above T_m is characterized by the absence of any translational order in the plane of the membrane, as well as the absence of conformational ordering of the acyl chains. In contrast, the low-temperature phase below T_m is solidlike because of the translational order in the bilayer plane, in addition to which the acyl chains are highly ordered. In this work, in which we consider the DPPC bilayer at 323 K above $T_m = 315$ K, we are therefore in the high-temperature liquid–crystalline phase.

The bilayer is fully hydrated by 3655 water molecules, for which we used the single-point charge (SPC) model [174]. The choice of initial configuration and force fields is discussed in Ref. [175].

For electrostatic interactions, we discuss here the two most commonly used techniques. First, we consider group-based long-range electrostatic interactions handled by using an abrupt cutoff at the truncation distance r_{cut}. For r_{cut}, commonly used values in lipid membrane simulations range from 1.2 to 2.0 nm. Here we consider a number of values in the same range, with an objective to find possible trends in the limit where the truncation distance is as large as possible compared to the linear system size. To this end, we have chosen $r_{cut} = 1.8$, 2.0, and 2.5 nm.

In addition to truncation, we discuss the PME [60, 159] technique that takes the long-range interaction fully into account through Ewald summation rules. Although it has been shown that PME suffers from certain drawbacks as a result of the periodic nature of the simulation box, the PME method has been found to work well in various soft matter systems. Thus, it is used as a reference in this work.

It is worthwhile to emphasize that the present choice of interactions, time steps, thermostats and barostats, and the choice to use the twin-range scheme for long-range interactions, follows the practice made in numerous MD simulations of lipid bilayers. In particular, this approach follows the choice made for a DPPC bilayer in the pioneering works by Berger et al. [176] and Tieleman and Berendsen [177].

4.1.2. Area Per Molecule in a Lipid Bilayer

In lipid membrane systems, the area per molecule is perhaps the most central quantity. From the experimental point of view, it is one of the few quantities that can be determined rather accurately. In addition, the area per molecule is closely related to the ordering of lipid acyl chains in the hydrophobic part of the membrane, and area fluctuations in turn are related to a number of processes such as permeation through membranes, hole formation, and lateral diffusion of lipids and other molecules in the plane of the bilayer. To investigate

the influence of electrostatic interactions on the structural properties of lipid membranes, we therefore focus here on the area per lipid.

Results in Fig. 8 illustrate the time dependence of the area per DPPC molecule $A(t)$ over a timescale of 50 ns. As the equilibration of $A(t)$ takes about 10 ns, it is reasonable to discard this part of the trajectory and use only the last 40 ns for analysis. The simulations using PME then yield $\langle A \rangle = (0.655 \pm 0.010)$ nm^2. This result is consistent with recent experiments summarized by Nagle and Nagle [157], which for DPPC give $\langle A \rangle = 0.64$ nm^2. Truncation at 2.5 nm reduces the size of the bilayer by about 8% compared to the PME result, giving $\langle A \rangle = (0.604 \pm 0.009)$ nm^2. Further decrease of the cutoff distance to 2.0 nm leads to $\langle A \rangle = (0.582 \pm 0.027)$ nm^2. Finally, the smallest truncation distance of $r_{\text{cut}} = 1.8$ nm considered here yields $\langle A \rangle = (0.551 \pm 0.005)$ nm^2. Without doubt, this result deviates substantially from the PME result. Figure 8 also reveals that if the scheme is changed from truncation at 1.8 nm to PME, the behavior of $\langle A \rangle$ rapidly evolves to the limit found for PME (see the dashed curve in Fig. 8 that starts at $t = 10$ ns). Thus, the deviations between the results for different schemes are solely the result of electrostatics, and not caused by initial conditions, for example.

Other simulation studies both in the low-temperature gel phase [178] and in the high-temperature fluid phase [179] are consistent with these findings. The study by Anézo et al. [179] is exceptionally revealing, as it covers a wide range of different electrostatic schemes and demonstrates the sensitivity of the area per lipid on the scheme chosen for electrostatic interaction.

The interactions used in the above systems have been completely similar in all cases, except for the scheme chosen for the long-range electrostatic interactions, yet the deviations between different model systems have been surprisingly large. To gain a better understanding of the reasons that lead to the above differences, we next focus on the most central quantity in all condensed and soft matter systems (i.e., the radial distribution function).

4.1.3. Radial Distribution Functions

The radial distribution function (RDF), $g(r)$, describes the probability of finding a pair of atoms a distance r apart, relative to the probability expected for a completely random distribution at the same density [61]. For the ideal gas characterized by the absence of conservative interactions, the RDF is therefore equal to 1 for all distances r. Interacting systems in turn are characterized by peaks and valleys that describe regions of high and low density, respectively. Hence, the RDF is an excellent means to gauge the ordering of any system.

Figure 9 shows the intermolecular RDF's between the nitrogen (N) atoms in the DPPC headgroups in a lipid bilayer. We find that the results for the PME technique are as expected, as in that case there is a strong main peak at about 0.85 nm and essentially no structure beyond 1.5 nm. This is consistent with the liquid–crystalline phase in which the lipids are
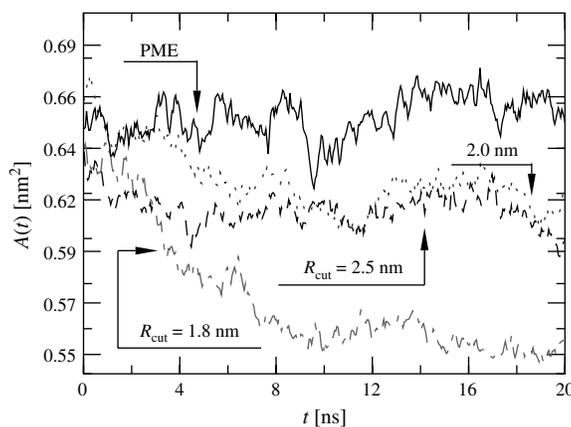
Figure 8. Temporal behavior of the area per molecule in a model system of a DPPC bilayer. Adapted from Ref. [175].
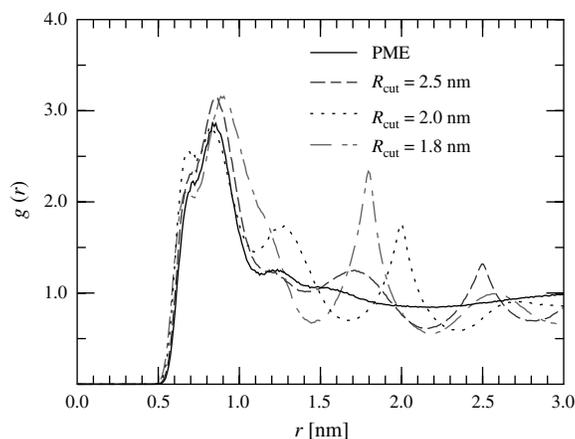
Figure 9. Radial distribution function for intermolecular N–N pairs in the headgroups of the dipalmitoylphosphatidylcholine molecules. Note the peaks exactly at the truncation distance. Adapted from Ref. [175].

in a fluid-like state, characterized by the lack of long-range ordering in the plane of the membrane. The truncation techniques, in contrast, lead to pronounced peaks exactly at the truncation distance. For the truncation distance of 1.8 nm, the RDF has a prominent peak exactly at 1.8 nm. When the truncation distance is increased, the peak moves accordingly.

The above finding gives rise to major concern as it indicates that the truncation of electrostatic interactions may change the phase behavior of the system. As a matter of fact, this is the case, as was shown by Patra et al. in Ref. [175] for the RDF between the centers of mass positions of the lipids. The artificial ordering resulting from truncation leads to unexpected peaks that are entirely unphysical, thus changing the phase behavior of the membrane from a liquid–crystalline to some semi-fluid-like phase characterized by ordering at the length scale of the truncation distance. At the same time, the PME method does not yield any structure of the same kind.

Without any doubt, the above artifacts in structural quantities are the result of the fact that the truncation method does not account for the long-range component of electrostatic interactions. Apparently, the truncation of electrostatic interactions gives rise to artificial ordering in the bilayer plane, thus changing the phase behavior of the system. Although related artifacts have been observed in a number of different soft matter systems, the artifacts seem to be more pronounced in lipid membranes. Why so? To address this question, let us first recall that the dielectric constant of water is about 80. This implies that long-range interactions in water are strongly screened. Inside the membrane, however, the dielectric constant is about 2–4, and in the vicinity of the water-membrane interface it is about 20–40. Screening of long-range interactions in a membrane is therefore considerably weaker compared with pure water, thus enhancing the magnitude of artifacts compared to water-like systems. Second, it has been found that the energetic cost of changing the area per lipid in a membrane is very small, on the order of thermal energy [180], which implies that even minor changes in electrostatics may lead to substantial changes in the area of the membrane. As numerous both structural and dynamical quantities are influenced by the (excess) area in a membrane, it is therefore understandable how they are so strongly influenced by the choice of electrostatics in lipid membrane simulations. Clearly, this is a major issue and warrants particular attention in large-scale simulations.

4.1.4. Discussion on Ways to Handle Electrostatic Interactions

The above example demonstrates that the choice of a scheme for electrostatic interactions may dramatically affect the structural properties of lipids in a DPPC bilayer. Recent studies in the same lipid bilayer system [181] have further shown that dynamical properties such as lateral and rotational diffusion are equally sensitive to the choice of electrostatics. Now, of course, we can claim that the above findings are specific to this system rather than being generic and holding true in a variety of different lipid membrane systems. On one hand,

the idea of the specific nature of our findings is true: The influence of the electrostatic scheme on the structure and dynamics depends on the force field and other practical details chosen for the model. However, on the other hand, we may also conclude that the results discussed above are by and large generic. In particular, for the truncation methods, the location of artificial peaks found exactly at the truncation distance in the radial distribution functions is the most clear example to support this view. More generally, although this discussion has centred on lipid membranes, it is likely that similar kinds of findings would be faced in many other soft matter systems such as membrane-DNA complexes and systems composed of proteins and other solutes embedded in or attached to membranes. As we prefer to be on the safe side, we conclude that the truncation of electrostatic interactions should be avoided, if possible.

Having made this conclusion, let us ask ourselves why the truncation method has been used in the first place, and why it is still rather popular in large-scale simulations. The answer lies in the computational cost associated with handling electrostatic interactions. In this model system composed of a DPPC bilayer in water, the computational cost of using PME is larger by a factor of 1.8 compared with truncation at 1.8 nm. In large-scale simulations, where the computations may take several months (or even years) of CPU time, this is a major issue. In the absence of better techniques, the truncation distance is therefore still commonly used in systems where the computational load is excessive.

Recent studies may provide a better solution to this situation, however: The reaction-field technique [61]. In the reaction-field method, electrostatic interactions are explicitly accounted for up to some truncation distance (typically about 1.0–1.5 nm), beyond which the remaining long-range contribution is described by a reaction field correction using a single dielectric constant. Whereas the origin of this method goes back to the days of Onsager in the 1930s, it is still a reasonable and relatively useful choice for treating long-range interactions in an efficient and rather accurate manner. Of course, as the reaction-field method is typically implemented using a single dielectric constant through the whole system, it may be problematic in nonisotropic cases such as water-membrane systems and other interfaces. Further work would therefore be highly welcome for developing the RF approach in this respect. Nevertheless, we are confident that the reaction-field method offers a safer choice than the methods based on an abrupt truncation of electrostatics. This conclusion is supported by recent studies by Anézo et al. [179]. Recent investigations on a DPPC bilayer system provide further support for this idea [181, 182], showing that the artifacts in RDFs can be largely avoided by using the reaction-field technique instead of truncation. Yet, the computational load of the RF approach is comparable to the truncation technique.

Despite the above discussion, we wish to conclude that approaches such as the PME technique are currently the method of choice, if the computational load is not a limiting factor.

4.2. Role of Random Numbers in Parallel Stochastic Simulations

Perhaps surprisingly, random numbers are used extensively in numerous both scientific and practical applications. Stochastic simulations in physical sciences [60, 183, 184] characterize this fact very well. There, where the aim is to model complex systems over large length and timescales, atomistic approaches such as classical molecular dynamics are therefore no longer feasible. Examples of such stochastic techniques include the MC method, Brownian dynamics simulations, and DPD, all of which use pseudorandom numbers to generate the dynamics of model systems under study. Because of the deterministic nature of pseudorandom numbers, it is obvious that if there are any significant correlations within pseudorandom number sequences, then the dynamics of these model systems will be biased and the reliability of the whole approach may become questionable. To avoid such concerns, pseudorandom number generators should be thoroughly tested before extensive use in model simulations. This is particularly true as regards stochastic simulations in a parallel environment and in large-scale studies of soft matter systems, where the number of random numbers used is typically huge.

Below, we consider these two topics side by side. We first discuss recent ideas [185] on how to test the quality of random numbers in parallel applications in terms of random walks.

As it turns out, this approach is both simple and efficient and is able to reveal correlation effects in a number of commonly used pseudorandom number generators. Second, we study the effects of the quality of random numbers on DPD simulations [11, 17–19], which are based on solving Newton's equations of motion in the presence of a stochastic force component. The results indicate a few promising generators whose performance in the present tests is rather remarkable.

4.2.1. How to Test Parallel Random Number Generators by Random Walks

Parallel computing and parallel random number generation have been active fields of research for a relatively long time. With the exception of a few interesting studies [186–188], much less attention has been devoted to design tests, which are specifically suited for gauging correlations in parallel random number generators. This is partly the result a great number of general test methods that have been developed and used during the last few decades [47]. In this general approach that is common to numerous standard tests, correlations are looked for from a very long random number sequence $\{x_i\}$. This approach works very well in serial computing, where the whole random number sequence is used on a single CPU. In parallel applications, however, the case is more complicated. Then it is more practical to consider relatively short subsequences $\{x_i\}^{(k)}$, $i = 1, \ldots, \Omega_k$, where subsequences $k = 1, \ldots, m$ do not overlap and are used on distinct CPUs during the calculation, as the task is distributed as subtasks between a number of different computing units. Obviously, the interest is now on cross-correlations between distinct random number sequences $\{x_i\}^{(1)}, \ldots, \{x_i\}^{(m)}$. To study such effects, one needs specific approaches that gauge long-range correlations between blocks of random numbers.

   We consider this problem in terms of random walks, which are a common tool in a variety of disciplines, including physics, chemistry, biology, and economics. We use random walks to characterize the quality of random numbers in parallel calculations. The key idea is to consider a number of diffusing random walkers, each of which is governed by a distinct random number sequence. Through studies of their mutual correlations we are then able to characterize and quantify possible correlations between separate pseudorandom number sequences. To this end, we calculate quantities, such as intersection probabilities between different random walks, and compare their asymptotic behavior to known theoretical limits. The difference between simulation results and theoretical predictions serves as a measure of correlations in the random number generator under study.

   In the following, we briefly describe three tests [185] that are based on this idea. They study both correlations within a single random number sequence $\{x_i\}^{(k)}$ and cross-correlations between distinct random number sequences $\{x_i\}^{(1)}, \ldots, \{x_i\}^{(m)}$. The sequences $\{x_i\}^{(k)}$ are generated by the pseudorandom number generator in question from $\{x_i\} = x_1, \ldots, x_\Omega, x_{\Omega+1}, \ldots, x_{2\Omega}, \ldots$ such that we obtain nonoverlapping sequences $\{x_i\}^{(1)} = x_1, \ldots, x_\Omega$, $\{x_i\}^{(2)} = x_{\Omega+1}, \ldots, x_{2\Omega}$, and so on. Here we consider the case where the sizes Ω_k of sequences $\{x_i\}^{(k)}$ are equal for all k. Random numbers x_i are uniformly distributed between zero and one.

4.2.2. Tests Based on Random Walks

In the height-correlation test, we consider the position y_i of a one-dimensional random walker versus the number of jumps made i. The position $y_t = \sum_{i=1}^{t} \delta y_i$ is a sum of displacements δy_i, which are random variables

$$\delta y_i = \begin{cases} +1, & \text{if } x_i \leq 1/3 \\ 0, & \text{if } 1/3 < x_i \leq 2/3 \\ -1, & \text{otherwise} \end{cases} \tag{28}$$

In this manner, we construct the paths $y_i^{(1)}$ and $y_i^{(2)}$ from the sequences $\{x_i\}^{(1)}$ and $\{x_i\}^{(2)}$, respectively. Using the initial condition $y_0^{(1)} = y_0^{(2)} = 0$, the height between the two random

walkers is defined as $h_t = y_t^{(1)} - y_t^{(2)}$. For a random process, the corresponding correlation function $H_t \equiv \langle |h_t - h_0| \rangle \sim t^\phi$ decays asymptotically as a power law with an exponent $\phi = 1/2$ [189].

The intersection test deals with two random walkers on a square lattice. The random walkers start from the origin at the same time and are allowed to jump independently on a lattice. Meanwhile, we consider the probability I_t that the two random walkers after t jumps have no intersection other than their common starting point. Note that the two random walks need not meet at the same site at the same time, but any common point in their paths is regarded as an intersection. For a random process, I_t behaves asymptotically like a power law $I_t \sim t^{-\alpha}$ with an exponent $\alpha = 5/8$ [190, 191].

In the S_N test, we consider N random walkers in one dimension and let them move simultaneously without any interaction such that, at any jump attempt, they can make a jump to the left or to the right with equal probability. After $t \gg 1$ jumps by all random walkers, the mean number of sites visited $S_{N,t}$ has an asymptotic form $S_{N,t} \sim f(N)t^\gamma$, where the scaling function $f(N) = (\ln N)^{1/2}$ and $\gamma = 1/2$ [192]. The value of γ observed serves as a measure of correlations.

Values of the parameters used in the tests discussed below are given in Table 2. There Ω is the number of jumps in a single random walk, and M is the number of independent runs in the test.

4.2.3. Results for Some Pseudorandom Number Generators

Here we discuss results of the three tests for a few carefully chosen pseudorandom number generators. The present results complement previous investigations discussed in more detail in Ref. [185].

The generators discussed in this work represent a variety of different generators often used within the physics community. R250 is an implementation of the generalized feedback shift-register (GFSR) algorithm [193] $x_n = x_{n-250} \oplus x_{n-103}$, where $\oplus$ is the bitwise exclusive OR operator. R89 is another example of GFSR generators and uses $x_n = x_{n-89} \oplus x_{n-38}$. RAN is a "minimal" linear congruential generator of Park and Miller (with multiplier 16807 and modulus $2^{31}-1$) combined with a Marsaglia shift sequence, and it has been proposed in a recent edition of *Numerical Recipes* [194]. RAN2 is based on the 32-bit combination generator first proposed by L'Ecuyer [195] and later published in *Numerical Recipes* [196], using shuffling. RANMAR [197, 198] combines a lagged Fibonacci generator with a simple arithmetic sequence and has been suggested as a good candidate when one aims towards a "universal generator" [198]. In addition, we test the Mersenne Twister [199], which has a huge period and good theoretical properties in view of recent studies by Matsumoto and Nishimura [199]. Finally, we consider the most "luxurious" version of RANLUX [200, 201], which is based on ideas of deterministic chaos. In RANLUX4, one generates 389 random numbers, delivers 24 of them, and throws the remaining 365 numbers away.

To determine the exponents ϕ, α, and γ, we considered their "running exponents" defined as

$$\phi_t \equiv \frac{\log(H_{t+\delta t}/H_t)}{\log[(t+\delta t)/t]} \tag{29}$$

where the present example is for the height correlation test. Running exponents of α and γ were determined in a similar fashion. The "time window" δt used in this work varied between 50 and 1000.

Demonstrative results for the running exponents are shown in Fig. 10. The initial behavior of some generators is clearly different and is discussed below in more detail. Nevertheless,

Table 2. Values of the parameters Ω and M used in the tests. The S_N test was carried out with $N = 2$.

	Height Correlation Test	Intersection Test	S_N Test
Ω	2000	4000	2000
M	10^7	10^8	10^8
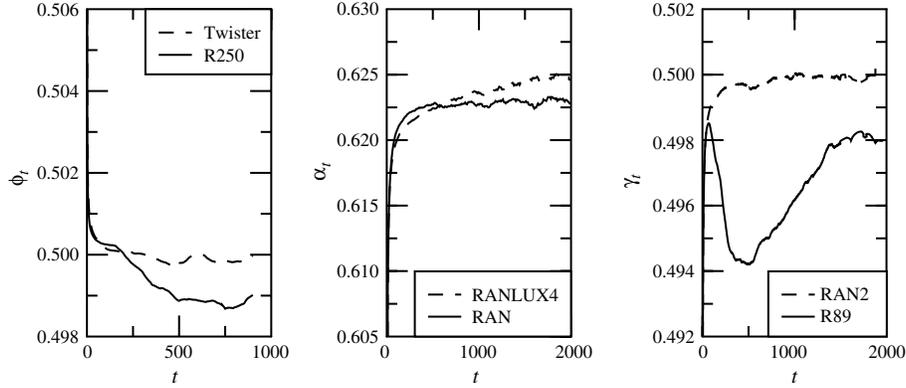
Figure 10. Demonstrative results of the running exponents (left) ϕ_t, (middle) α_t, and (right) γ_t in the three tests. The exponents shown in Table 3 have been extracted from the large-t limit. For clarity's sake, results of only a few generators are shown here.

asymptotically, we find the running exponents of all generators to converge to some limiting value at large t. This regime was therefore used to determine the exponents.

The results for the exponents ϕ, α, and γ are given in Table 3. Results of the GFSR generators R89 and R250 are not surprising, as they have recently failed in various random walk tests (see references in [185]). RANLUX4 and RANMAR, on the other hand, perform considerably better, in agreement with some previous studies [185, 198, 202–205]. Equally good performance is found for the Mersenne Twister and RAN2, whereas the results of RAN in the intersection test are not fully convincing. Then, the exponent α deviates from the theoretical limit, and this deviation seems to be systematic, as is demonstrated by Fig. 10 (middle). When the intersection test was repeated for RAN, we found $\alpha = 0.6239(4)$, which again fails the test.

The tests presented above focus on the asymptotic behavior of the corresponding correlation functions. This is because only the asymptotic behavior of the correlation functions is known theoretically. Yet correlations in pseudorandom number sequences are playing a role also at shorter scales, as is very evident from the running exponents in Fig. 10. For cases in which such short-range correlations are of interest, one can use a two-way analysis in which various pseudorandom number generators are judged against a reference generator. This approach was used in Ref. [185], which revealed further correlations in a few random number generators, including RANMAR.

At this stage, we conclude that even highly recommended and commonly used pseudorandom number generators may fail in tests that focus on their weak points. In general, one can safely say that there are no good pseudorandom number generators. All of them are bad. It is simply a question of finding those that are better and more reliable than the others. We will discuss this broad issue in more depth in Section 4.3.3. Meanwhile, to get further

Table 3. Results for the exponents of the three tests. The notation 0.4989(2) means 0.4989 ± 0.0002.

RNG	Height Correlation Test ϕ	Intersection Test α	S_N Test γ
R250	0.4989(2)	0.6265(5)	0.4984(1)
R89	0.4984(2)	0.6205(5)	0.4981(1)
RANLUX4	0.5001(1)	0.6257(5)	0.5000(1)
RANMAR	0.5000(1)	0.6250(4)	0.5001(1)
Twister	0.4999(1)	0.6242(5)	0.5000(1)
RAN	0.4999(1)	0.6224(4)	0.5000(1)
RAN2	0.5001(1)	0.6241(5)	0.5000(1)
Exact	1/2	$5/8 \approx 0.625$	1/2

Note: The exponents were extracted from the asymptotic tail of the corresponding correlation functions, and the exponents that deviate from the exact value by more than two error bars have been underlined.

insight into the use and effect of random numbers in stochastic simulations, we consider their influence on DPD simulations.

4.3. Role of Random Numbers in Dissipative Particle Dynamics Simulations

Above we have stressed the importance of developing techniques for multiscale modeling. Development of stochastic simulation techniques also plays a crucial role in this respect. This is simply because many important processes in soft matter systems take place at mesoscopic length and timescales (roughly 1–1000 nm and 1–1000 ns) that are beyond the limits of atomic-scale molecular dynamics. To overcome this problem, stochastic simulation methods such as DPD, discussed in Section 3.2, are hence needed.

The role of random numbers in DPD is crucial. This is because particles in DPD model systems are partly driven by stochastic noise, which is generated by pseudorandom numbers. Any correlations in pseudorandom number sequences can therefore lead to serious problems if they interfere with the true dynamics of the system.

To study how sensitive DPD actually is to such underlying correlations in pseudorandom number sequences, we consider a simple model fluid system described by N particles with masses m_i, coordinates $\vec{r}_i$, and velocities $\vec{v}_i$. Interparticle interactions are characterized by the pairwise conservative, dissipative, and random forces exerted on particle i by particle j, respectively, and are given by

$$\vec{F}_{ij}^C = \alpha\omega(r_{ij})\vec{e}_{ij}$$
$$\vec{F}_{ij}^D = -\gamma\omega^2(r_{ij})(\vec{v}_{ij}\cdot\vec{e}_{ij})\vec{e}_{ij} \tag{30}$$
$$\vec{F}_{ij}^R = \sigma\omega(r_{ij})\xi_{ij}\vec{e}_{ij}$$

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$, $r_{ij} = |\vec{r}_{ij}|$, $\vec{e}_{ij} = \vec{r}_{ij}/r_{ij}$, and $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$. The weight function for the different interaction terms has here been chosen to follow the same form. Random numbers appear in these equations by describing ξ_{ij}, which are symmetric random variables with zero mean and unit variance. The ξ_{ij} are expected to be uncorrelated for different pairs of particles and different times, and it is indeed random numbers whose task it is to satisfy this condition.

Remaining details are fixed by adopting a commonly made choice for the weight function

$$\omega(r_{ij}) = \begin{cases} 1 - r_{ij}/r_c & \text{for } r_{ij} < r_c \\ 0 & \text{for } r_{ij} > r_c \end{cases} \tag{31}$$

with a cutoff distance r_c [17] and $\omega^R(r_{ij}) = \omega(r_{ij})$. Then the equations of motion are given by the set of stochastic differential equations

$$d\vec{v}_i = \frac{1}{m_i}\left(\vec{F}_i^C\,dt + \vec{F}_i^D\,dt + \vec{F}_i^R\sqrt{dt}\right) \tag{32}$$

where $\vec{F}_i^R = \sum_{j\neq i}\vec{F}_{ij}^R$ is the total random force acting on particle i (with $\vec{F}_i^C$ and $\vec{F}_i^D$ defined correspondingly).

4.3.1. Model System

To maximize the role of random numbers, we study a simple model fluid system of identical particles ($m_i = m\ \forall i$) in the absence of conservative forces ($\alpha = 0$). This choice corresponds to an ideal gas, which provides us with some exact theoretical results to be compared with those of model simulations. The dynamics of the system then arise only from random noise and from a dissipative coupling between pairs of particles. The random force strength is chosen as $\sigma = 10$, and the dissipation strength is given by the fluctuation–dissipation relation $\sigma^2/\gamma = 2k_BT^*$ [18], where the desired thermal energy is chosen as $k_BT^* = 1$. The length scale is defined by $r_c = 1$, and time is given in units of $r_c\sqrt{m/k_BT}$.

In our simulations we use a three-dimensional box of size $10 \times 10 \times 10$ with periodic boundary conditions and a particle density $\rho = 4$. Equations of motion [Eq. (32)] are solved in practice using the recently suggested integration scheme known as DPD-VV [125–127],

and the time step used in this procedure was $\Delta t = 0.001$. With this choice, the average temperature during simulations $\langle k_{\mathrm{B}} T \rangle$ remains at the desired value within 1%. The results shown below are for simulations of 10^6 time steps after equilibration, which takes about 200 CPU-hours each on a typical RISC workstation.

Because the results depend on how the random number sequences are used, we discuss this issue in some detail. Consider particles $i = 1, \ldots, N$. For every pair of particles at a given moment, we need to generate an independent and symmetric random variable ξ_{ij} with zero mean and unit variance, which here is done using uniformly distributed random numbers with unit variance [11]. Then using the particle "1" as an example, one finds all other particles "j" that interact with the particle "1" (i.e., particles for which $r_{1j} \le r_c$). Suppose that the number of such particles $\{j_1, j_2, \ldots, j_K\}$ is K. Then one generates a random number sequence $\{x_i\} = x_1, \ldots, x_K$ such that x_1 determines ξ_{1j_1}, x_2 determines ξ_{1j_2}, and so forth. The same procedure is then performed for all particles interacting with the particles "2" through "$N-1$," such that each pair of particles is considered exactly once.

The results below have not been published elsewhere, which possibly explains our interest in specifying this model in great detail.

4.3.2. Results for Some Pseudorandom Number Generators

In this discussion, we consider a small sample of the generators described in Section 4.2.3. We focus on the radial distribution function $g(r)$ [206] that is one of the most central observables in studies of liquids and solid systems. For the ideal gas considered here, $g(r) \equiv 1$ for all r, and therefore any deviation from one has to be interpreted as an artifact caused by the employed simulation scheme. This approach provides a simple and effective measure of correlation effects and serves our main interest of seeing how sensitive DPD is to the underlying correlations in pseudorandom number sequences.

Results shown in Fig. 11 (left) indicate that differences between the generators are small. For all generators tested, the radial distribution function is almost flat and has only minor deviations from the expected behavior. Did we expect something else that would have allowed us to conclude that some pseudorandom number generators lead to results that are simply nonsense? Maybe. At anyrate, we are definitely glad that we did not find major artifacts.

Aside from noise effects, the minor deviations found in Fig. 11 (left) may result from three different sources. First, the expected behavior of $g(r) = 1$ is rigorously true only in the limit $\Delta t \to 0$, whereas in practice the time step is always finite. Second, how Eq. (32) is integrated also affects physical quantities such as $g(r)$. Third, the deviations may result from correlations in random number sequences.
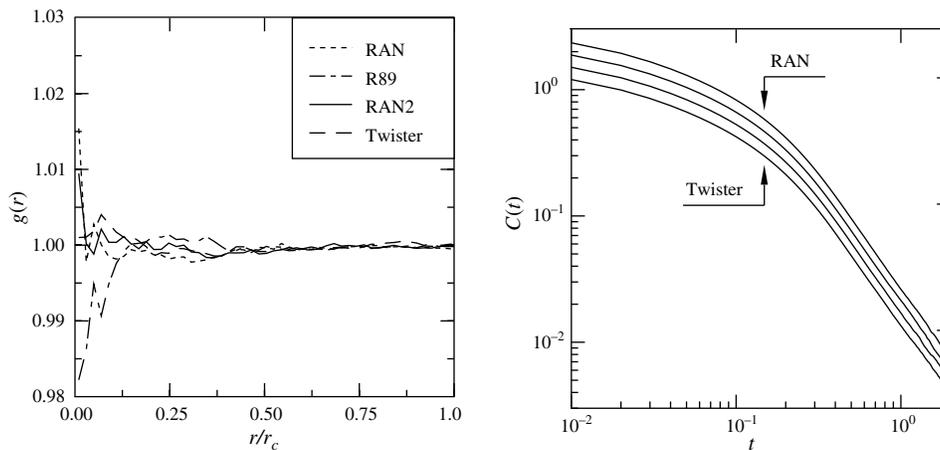
Figure 11. Left: Demonstrative results of the radial distribution function $g(r)$ for a few random number generators tested. Note that error bars are largest for small r, as the data was collected with a fixed bin size $dr = r_c/100$. Right: The decay of the velocity correlation function $C(t)$ for a few generators from top to bottom: RAN, R89, RAN2, and the Mersenne Twister. Results of R89, RAN2, and the Mersenne Twister have been shifted to clarify the presentation.

In this case, the time step Δt and the integration procedure are fixed, and therefore their effect is similar for all generators. Thus, the best we can do is to compare the four curves in Fig. 11 (right) with one another. This comparison leads to a simple conclusion that all generators yield approximately similar results within error bars. Only R89 deviates from other generators at small r, and even in this case the fluctuations are rather weak.

To complement this test, we also calculated the isothermal compressibility κ_T defined as $\kappa_T \rho k_B T = 1 + 4\pi\rho \int_0^\infty dr r^2[g(r) - 1]$, which is an example of a thermodynamic response function. Results of all generators were found to be very similar. Finally, we used the single-particle velocity correlation function $C(t) = \langle \vec{v}_i(t) \cdot \vec{v}_i(0) \rangle$ [206] to gauge possible problems with time-dependent quantities such as diffusion. These studies revealed that the results of all generators for $C(t)$ were essentially identical [see Fig. 11 (right)].

These results indicate that DPD model simulations are relatively insensitive to the underlying correlations in pseudorandom number sequences. This idea is supported by further calculations using rather poor generators (such as the GFSR generator $x_n = x_{n-31} \oplus x_{n-3}$), whose results were comparable to those discussed above. However, we cannot conclude that this finding is generic, as the test results depend on how random numbers are used within a simulation. In this case, we feel that the fluid-like nature of the system plays certain role. Namely, as was described in Section 4.3.1, the motion of the particle i is influenced by random forces acting between i and other neighboring particles nearby. Because the number of neighboring particles K is usually small and not conserved, the motion of the particle i is dictated by a short random number sequence $\{x_i\} = x_1, \ldots, x_K$, whose size K fluctuates in time. Furthermore, the random number sequences $\{x_i\}^{(1)} = x_1, \ldots, x_{K_1}$ and $\{x_i\}^{(2)} = x_{1+L}, \ldots, x_{K_2+L}$ used at consecutive moments for the same particle are separated by a sequence of size $L \sim \mathcal{O}(N\langle K \rangle)$. Thus, to conclude, it seems plausible that only strong, short-range correlations within $\{x_i\}$ influence the dynamics of simple fluid systems that have been considered in this work.

We found above that the quality of pseudorandom numbers is likely not crucial in DPD model simulations of simple fluid systems. This favorable idea should be taken with a grain of salt, however. A number of simulation studies have shown [185, 207, 208] that the underlying correlations in pseudorandom number sequences may interfere with the true dynamics of model systems. Therefore it is possible that the same problems are lurking behind the shadows and could be faced in DPD simulations, too. An example of such a situation is DPD simulations in a parallel environment, where pseudorandom number sequences can be used in various different ways (such as the leap-frog and blocking techniques, to name just two examples). This warrants particular care to be taken when pseudorandom number generators are being used in future large-scale applications. Although such tests have not been carried out yet (to our knowledge), they would be highly welcome to clarify this issue.

4.3.3. Discussion on the Future Aspects of Using Pseudorandom Numbers

The Buffon experiment [209] is probably one of the first and best known examples of the MC method. In this experiment, one throws needles of equal length at random over a plane marked with parallel and equidistant lines. By counting the number of intersections between lines and needles, one can estimate the value of π. Provided that the needles were thrown by a pseudorandom number generator, how precise and "universal" values for π would you expect?

If your answer is positive and in favor of using pseudorandom number generators without any concern, perhaps a few practical examples might change your mind. In the late 1980s, it was found [210] that the critical exponents in the three-dimensional Ising model depended on the random number generator used in the MC simulations. Later in the 1990s, Ferrenberg et al. [207] draw the same conclusion in the two-dimensional Ising model when cluster algorithms were employed for the dynamics of the system. Similar findings in other contexts are numerous and include, for example, studies of surface growth, deposition problems, and random walks [202, 208, 211].

The above examples demonstrate that in stochastic simulation techniques, there is an extra degree of freedom that should always be mentioned when the results are reported

in the literature: the pseudorandom number generator. This may sound sarcastic, but the fact is that the results of stochastic simulations are a function of the pseudorandom number generator used.

As a consequence, numerous test methods have been developed with an aim to finding (undesired) correlation effects from the pseudorandom number sequences. The standard tests that focus on general properties of random number sequences form the most significant test bench in this field [47]. They are complemented by more specific tests, often known as application-specific tests [202, 203] that are designed to mimic some particular application. For research topics such as surface growth or surface diffusion, using lattice-gas models, the Ising model is one of the most convenient starting points when one aims to develop an application-specific test for these purposes. In addition to these, theoretical tests such as the spectral test and discrepancy complement our knowledge of the properties of pseudorandom number algorithms. In all, there is a huge body of tests available. The key issue is to use them and, further, to develop new test methods for ever increasing requirements. When done carefully, one can avoid undesired artifacts resulting from the pseudorandom number generator used.

We have demonstrated above how correlations in some commonly used pseudorandom number generators may change the behavior of physical quantities such as time-dependent correlation functions. However, we have also found that some recently suggested generators such as the Mersenne Twister and RANLUX4 passed all the present tests. A pessimist could now say that all pseudorandom number generators have inherent weaknesses, and that eventually they fail anyway. Optimistically speaking, however, one can conclude that there are still many generators whose properties are reasonably good even for very challenging applications. To make sure that the situation will remain equally positive in the future as the computational power is increased, further work is definitely called for to develop more reliable pseudorandom number generators. When combined with novel test methods to challenge them, we are on the right track.

4.4. How to Integrate Equations of Motion in DPD Simulations

In the above discussion, we have noted that there is an increasing demand for useful and efficient methods to deal with complex soft matter systems in the mesoscopic regime. DPD [11, 17–19] (see Section 3.2) is a particularly appealing technique in this regard. This is mainly because the "particles" of DPD correspond to coarse-grained entities, which allows the use of large time steps (compared to typical molecular dynamics simulations) as pair potentials for the coarse-grained description are considerably softer than for the underlying atomic description. Further, the pair potentials in DPD can be chosen such that the hydrodynamic modes are accounted for in a proper fashion.

In practice, to obtain full hydrodynamics, the choice of interactions is done through a pairwise coupling of particles through random and dissipative forces. As we will note below, this makes the integration of the equations of motion a nontrivial task. The main difficulty here arises from the dissipative force, which depends explicitly on the relative velocities of the particles, whereas the velocities in turn depend on the dissipative forces. An accurate description of the dynamics therefore requires a self-consistent solution, which in turn is computationally demanding.

Here we discuss how this problem can be overcome. In addition to the self-consistent approach, we discuss other complementary methods that account for the velocity dependence of dissipative forces in some approximate manner, allowing the integration to be carried out to a sufficient degree of computational efficiency.

4.4.1. Equations of Motion in DPD

The time evolution of particles in DPD can be described by the Newton's equations of motion

$$d\vec{r}_i = \vec{v}_i \, dt$$

$$d\vec{v}_i = \frac{1}{m_i}(\vec{F}_i^C \, dt + \vec{F}_i^D \, dt + \vec{F}_i^R \sqrt{dt})$$

(33)

Here $\vec{F}_i^C = \sum_{i \neq j} \vec{F}_{ij}^C$ is the total conservative force acting on particle i, and $\vec{F}_i^D$ and $\vec{F}_i^R$ are defined correspondingly for the dissipative and random forces, respectively.

The reason why we come back to this issue is the fact that Eq. (33) is at the heart of the problem as regards DPD. It shows that the velocities of the particles depend on the dissipative forces acting on them. At the same time, the dissipative forces in turn depend on the relative velocities of the DPD particles [see Eq. (30)]. The velocities and dissipative forces of all the particles are therefore intertwined. This is precisely the reason why the self-consistent solution is needed in the first place.

Equation (33) furthermore shows that the similarity of this case to classical MD simulations is close. Nevertheless, there is reason to point out that the random force term in Eq. (33) is coupled to the factor $\sqrt{dt}$ instead of dt. This can be justified by a Wiener process, as in stochastic differential equations. Here, it suffices to notice that physically the Wiener process models intrinsic (thermal) noise in the system and provides the simplest approach to modeling Brownian motion using stochastic processes (see Refs. [11, 18]).

4.4.2. Model Using DPD

Although the above continuous-time version of DPD satisfies detailed balance and describes the canonical NVT ensemble, in practice the time increments in Eq. (33) are finite and the equations of motion must be solved by some integration procedure. To study possible artifacts caused by the choice of the integration scheme, we consider a gaslike system of N identical particles. The model discussed here is essentially similar to that presented in Section 4.3.1.

In essence, we consider an ideal gas ("ideal DPD fluid") in three dimensions, where the random force strength is chosen to be $\sigma = 3$ in units of $k_B T^*$ and the particle number density is $\rho = 4$. The weight function $\omega^R(r_{ij}) = \omega(r_{ij})$, as well as other practical details have been chosen as in Section 4.3.1. Finally, the random numbers were generated by the Mersenne Twister.

The reason for using this model is that the ideal DPD fluid model provides us with some exact theoretical results that can be compared to results from model simulations. Hence, the dynamics of the present system arise only from thermal noise and dissipative coupling between pairs of particles.

4.4.3. Integration Schemes

The integration schemes discussed below have been suggested very recently. They complement each other in the sense that the velocity dependence of the dissipative forces is accounted for in all cases, but the approaches differ substantially from each other.

The standard velocity–Verlet (VV) [212] is typically the integrator in classical MD simulations. It is time-reversible and symplectic and works well especially at large time steps [61]. Here we use VV as a starting point for integrators designed specifically for DPD simulations.

The modified VV algorithm (DPD-VV) differs from the standard VV in the sense that it accounts for the velocity dependence of dissipative forces in an approximate fashion by updating the dissipative forces for a second time at the end of each integration step. When this idea is extended further, we obtain a self-consistent variant of DPD-VV. This SC-VV algorithm [125–127] determines the velocities and dissipative forces self-consistently through functional iteration, and the convergence of the iteration process is monitored by the instantaneous temperature $k_B T$. This approach is similar in spirit to the self-consistent leap-frog scheme introduced recently by Pagonabarraga et al. [213]. Other schemes considered here include the approach by den Otter and Clarke (OC) [214], in which the parameters of a leap-frog algorithm are adjusted such that the effects resulting from the velocity dependence of dissipative forces are reduced as much as possible, and the approach (S1) by Shardlow [215], in which the integration process is factorized such that the conservative forces are calculated separately from the dissipative and random terms.

In addition to the above integrators, we also discuss the performance of Lowe's integration scheme [133] based on combining the Andersen thermostat [134] with the VV scheme in a momentum conserving DPD-like fluid. This approach differs from the "traditional" DPD

in the sense that now the dissipative and random forces are not explicit. Rather, they are described by the Andersen thermostat that thermalizes the system for pairs of particles. In practice, one first integrates Newton's equations of motion with a time step Δt, and then for all pairs of interacting particles one decides with a probability $\Gamma \Delta t$ ($0 < \Gamma \Delta t \leq 1$) whether to take a new relative velocity from a Maxwell distribution. The dynamical properties of the system can be tuned by the parameter Γ [133], which is inversely proportional to the decay time for relative velocities.

A full description of the integrators is given in Ref. [126].

4.4.4. Demonstrative Results

As discussed above, this model is characterized by the absence of conservative forces, and thus any artifacts arising from the velocity-dependent forces are expected to be pronounced in this model. To study this possibility, we first discuss the deviations of the observed kinetic temperature

$$\langle k_{\mathrm{B}} T \rangle = \frac{m}{3N - 3} \left\langle \sum_{i=1}^{N} \vec{v}_i^2 \right\rangle \tag{34}$$

from the desired temperature $k_{\mathrm{B}} T^*$. As the simulations are employed in the canonical ensemble, the temperature conservation is one of the main conditions for a reliable performance of the model. The results for $\langle k_{\mathrm{B}} T \rangle$ shown in Fig. 12 show that DPD-VV and SC-VV are reasonably good at small time steps, whereas larger time increments lead to major deviations from the desired temperature. For OC, $\langle k_{\mathrm{B}} T \rangle$ decreases monotonically with Δt for $\Delta t \leq 0.15$. For larger time steps, we find the temperature to increase rapidly. Overall, we can conclude that the deviation in the case of OC is greater than for DPD-VV but considerably smaller than in the case of SC-VV. The Shardlow S1 integrator, in turn, has excellent temperature control, and the deviations remain less than 0.5% up to $\Delta t = 0.2$. The best temperature control is found for the method by Lowe, however, yielding $\langle k_{\mathrm{B}} T \rangle = k_{\mathrm{B}} T^*$ for all time steps Δt. Although Lowe's approach depends on the dynamic variable Γ, the conclusions with regard to temperature conservation have not been found to depend on the value chosen for this parameter [126].

Next, we examine the radial distribution function $g(r)$ [216], which is one of the most central observables in studies of structural properties of liquids and solids. For the ideal gas, the radial distribution function provides an excellent test for the integrators, as then $g(r) \equiv 1$ at the continuum limit. Therefore, any deviation from unity has to be interpreted as an artifact resulting from the integration scheme employed.

Results for $g(r)$ are shown in Fig. 13. We find that the deviations from the ideal gas limit $g(r) = 1$ are pronounced for OC, indicating that even for small time steps, this integration
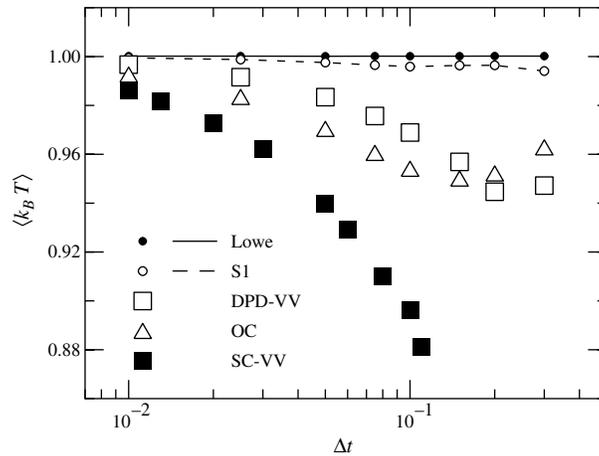
Figure 12. Results for the deviations of the observed temperature $\langle k_{\mathrm{B}} T \rangle$ from the desired temperature $k_{\mathrm{B}} T^* \equiv 1$ versus the size of the time step Δt [126]. The error in $\langle k_{\mathrm{B}} T \rangle$ is of the order of 10^{-4}. Adapted from Ref. [126].
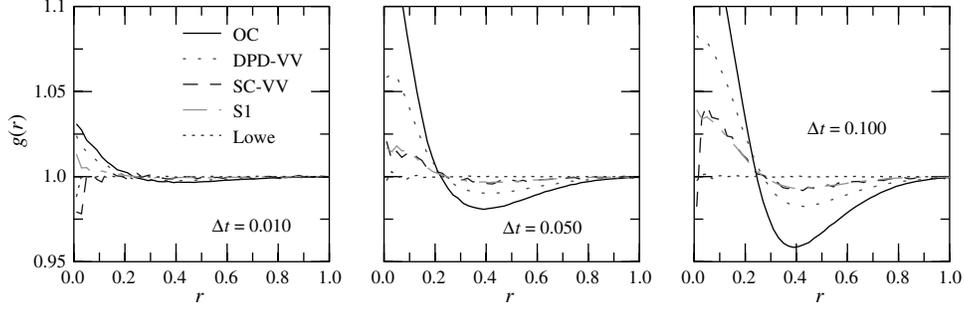
Figure 13. Radial distribution functions $g(r)$ with several values of time step Δt in model A: (a) $\Delta t = 0.01$, (b) $\Delta t = 0.05$, and (c) $\Delta t = 0.1$ [126]. The error in $g(r)$ is greatest at $r = 0.01$, where it takes the value of 0.01. Adapted from Ref. [126].

scheme gives rise to unphysical correlations. The performance of DPD-VV is considerably better, although artificial structures are yet rather pronounced, whereas SC-VV and S1 lead to a radial distribution function that is close to the theoretically predicted one. Completely structureless $g(r)$ is found only for the integrator by Lowe, however. Again, in this case, we have tested the behavior of $g(r)$ with various values of Γ, but the results remain the same. This confirms the expectation that Γ does not influence the equilibrium properties of the system.

The radial distribution function reflects equilibrium (time-independent) properties of the system. To complement the comparison of different integrators, we next consider the tracer diffusion coefficient

$$D_T = \lim_{t \to \infty} \frac{1}{6\,t} \langle [\vec{r}_i(t) - \vec{r}_i(0)]^2 \rangle \tag{35}$$

which can provide us with information of possible problems on the dynamics of the system. Here $\vec{r}_i(t)$ is the position of a tagged particle, and the mean-square displacement is then averaged over all particles in a system to get better statistics for D_T.

The results for the diffusion coefficient D_T in Fig. 14 are essentially consistent with the conclusions above. The integrator OC is not very useful in a system of this kind, since it seems to lead to substantial deviations from the expected behavior. The SC-VV and the integrator by Lowe perform much better, and the integrators S1 and DPD-VV are most stable in this case.

4.4.5. Discussion on Methodology of DPD

Above, we discussed the performance of various novel integration schemes that have been designed specifically for DPD simulations. On the basis of the ideal gas results, we can conclude that there are two approaches whose performance is above the others. Of these,
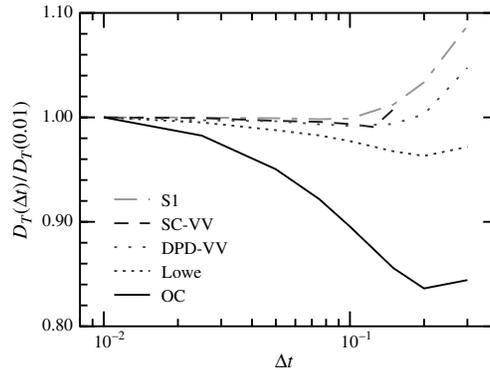
Figure 14. Results for the tracer diffusion coefficient $D_T(\Delta t)/D_T(0.01)$ versus the time step Δt [126]. The error in $D_T(\Delta t)/D_T(0.01)$ is on the order of 0.001. Adapted from Ref. [126].

Shardlow's integration scheme is based on splitting the equations of motion and can be applied to the usual DPD picture, whereas the approach by Lowe is distinctly different in nature and is related to the classical work by Andersen. Results for other, more strongly interacting models support this conclusion [126]. It is also noteworthy that both of these schemes are fast and easy to implement. Furthermore, and what is important when Lowe's method is compared to Shardlow's integration scheme, it provides an alternative and a very attractive description of dissipative particle dynamics. It is actually rather surprising that Lowe's approach has not attracted considerable attention yet. Perhaps the situation will improve shortly.

In addition to this issue, the methodology and the theoretical background of DPD have taken many significant steps very recently. From a simulation point of view, DPD has been shown to be a promising technique for studies of soft matter systems under non-equilibrium (flow) conditions [128]. On the theoretical side, a few extensions have been suggested recently [217, 218]. The Smoothed Particle Hydrodynamics framework suggested by Pep Español and his collaborators is particularly interesting as it lacks a number of limitations imposed by the traditional DPD picture [218]. We also would like to mention the recent work by Chris Lowe [133], whose alternative description for DPD is an attractive idea since it allows one to tune solvent properties in a reasonable and efficient way. Finally, as coarse graining is one of the grand challenges in the field of multiscale research, we would like to mention the recent work by Lyubartsev et al. [219]. Therein, the authors presented an approach to coarse-grain a MD description by the IMC technique such that the effective interactions found through this procedure can be applied to DPD simulations to study the large-scale properties of aqueous electrolyte solutions. This case will be discussed in more detail in Section 6.

We conclude that the development of methodology is an integral part of computational research. This is particularly true in soft matter research that deals with various fundamental questions, having a major effect on our understanding of biological systems. In particular, we hope that the case studies discussed above have highlighted some of the problems that may be hiding under the simulation protocols. Great care is therefore warranted to avoid artifacts resulting from the methods used, and more attention is called for to develop more accurate and efficient techniques to deal with complex, biologically relevant soft matter systems.

5. NANOSCIENCE OF BIOPHYSICAL SYSTEMS THROUGH ATOMIC-SCALE MOLECULAR DYNAMICS

At present, the length scales accessible to atomic-level molecular dynamics simulations are on the order of 5–50 nm. This may sound modest, but in practice it allows one to gain plenty of insight into the understanding of various biologically relevant molecules and molecular systems under hydrodynamic conditions. The above length scale further stresses the fact that MD is currently the method of choice for processes and phenomena that take place in the nanoscopic regime. In this section we therefore discuss a few examples that hopefully highlight the capability of molecular simulations, and MD in particular, in revealing the microscopic mechanisms and physical laws that govern biophysical systems on the atomic level. Because the field is enormous, we are bound to focus on one topic. We hence concentrate on the properties of lipid bilayers that are commonly adapted as basic models of cell membranes. In addition to these properties, we also, though very briefly, discuss recent computational modeling of other biomolecules related to membrane systems.

5.1. Lipid Membranes in a Nutshell

All cells are surrounded by cell membranes that are like thin elastic sheets. The thinness of the membranes is genuinely interesting, as whereas the diameter of a cell is typically 20–30 μm, the thickness of a cell membrane is just a fraction of this, being typically about 5 nm. Yet membranes are involved in essentially all processes related to the flow of nutrients and ions into and out of the cell. Membranes further provide an environment in which a

variety of biochemical processes occur, and in general the role of membranes in cells is crucial, as they function to organize biological processes by compartmentalizing them.

The main structural components of native biological membranes [220–222] are lipids (Greek: lipos, fat). As shown in Fig. 7, they have a polar headgroup and a nonpolar hyrdocarbon tail. Because the polar "water-loving" headgroups are able to form hydrogen bonds, they favor contact with water, whereas the "water-hating" nonpolar tails try to avoid water as much as possible. This "schizophrenic" nature of lipid molecules is the underlying reason why they self-assemble as bilayer-like structures, where two lipid monolayers face each other (see Fig. 7).

Although this idea of describing lipid membranes simply as lipid bilayers is very convenient, it is yet somewhat misleading because biological membranes are not like single-component lipid bilayers but, rather, are mixtures of various types of lipids that differ in a number of ways [222]. Some differences may seem minor, such as the length of hydrocarbon tails or the position of a double bond in the tails of monounsaturated lipids. Yet there are differences that make some lipids distinctly different from each other, such as the chemical composition and the net charge of the polar headgroup. In all, there are hundreds (or even thousands) of different lipid species that are found in biological membranes. In addition to this, to make molecular modeling more challenging, one should account for the fact that lipid membranes include numerous kinds of proteins either embedded in (integral proteins, ion channels) or attached peripherally to the membrane. The proteins are also related to a dynamic rubberlike network known as the cytoskeleton [222], which is attached to the inner surface of the membrane, whereas the outer leaflet of the membrane is covered by a network that in this case is made of glycocalyx carbohydrates. All together, for a physicist modeling soft matter systems, biological membranes are a mess.

Life would be much easier if biological membranes could be described by simple theoretical tools such as models related to the Ising model. Perhaps surprisingly, this approach has indeed been applied to cell membranes, and the outcome has been very successful [223–228]. Sarcastically speaking, it has been said that life can be described by the Ising model, too: Of the two states available in the traditional Ising model, one is then "alive" and the other is "dead." However, even though the Ising model and related approaches can be very useful in some cases, it is clear that it is a very coarse-grained description for complex systems such as lipid membranes. We cannot expect that approaches of that kind can provide any insight into the understanding of membrane systems over microscopic scales. In this nanoscopic regime, where atomic and molecular degrees of freedom cannot be neglected, more detailed approaches are needed to resolve the microscopic mechanisms that dictate the properties of given systems.

5.2. Molecular Dynamics of One- and Two-Component Lipid Membranes

Molecular dynamics provides a unique means to consider lipid membranes as well as other biologically relevant soft matter systems from an atomic perspective (see recent review articles in Refs. [84–88]). In particular, it allows one to gain insight into the nature of atomic-scale phenomena and related issues with a level of detail missing in any experimental technique. Nevertheless, MD also allows one to analyze membrane properties that are of experimental interest, thus providing a connection to macroscopic properties of the system.

As MD simulations of native biological membranes are not feasible, a natural starting point is to consider simpler models that include only the most essential features of cell membranes. This line of thought has lead to numerous studies of single-component lipid bilayers. Even though this approach may sound too simplified, it can yield a substantial amount of knowledge of model membrane systems, and as a matter of fact, the same idea has been employed in experiments.

First, MD simulations of single-component lipid bilayers were carried out in the 1980s and early 1990s [229, 230]. The timescales and system sizes were obviously very limited, but the early studies were already able to demonstrate the usefulness of MD simulations in this field. This is particularly true as regards the structure of lipid bilayers, which was then the key

aspect because of the relatively short timescales accessible through MD simulations (on the order of 1 ns at that time). More recently, and in particular during the last few years, better computer resources have made it possible to examine some of the dynamical aspects, too.

Another direction of research that has attracted plenty of activity is the study of two-component lipid bilayers. This is simply a result of the fact that there is a rather limited number of lipid species that are abundant in cell membranes. As an example, in human erythrocytes the most abundant lipid species are [220] cholesterol (25%), phosphatidylcholine (PC) (19%), phosphatidylethanolamine (PE) (18%), sphingomyelin (SM) (17.5%), and phosphatidylserine (PS) (8.5%). In human myelin membranes the abundant lipid species are [222] cholesterol (27%), ganglioside (26%), PC (10%), PE (20%), SM (8.5%), and PS (8.5%). Whereas PC, PE, and SM molecules are neutral but zwitterionic, PS is charged (anionic). For the discussion below, it is remarkable to note that cholesterol, PCs, and SMs together make up about 62% of the composition of human erythrocytes, and about 45% of human myelin membranes. Other highly concentrated molecules playing a major role in cell membranes are, for example, glycolipids, sugar-containing lipid molecules whose polar headgroups are attached to oligosaccharides. Thus, they play a prominent role in cell–cell interactions.

The most relevant case for all eucaryotic plasma membranes that do contain large amounts of cholesterol is a mixture of PC and cholesterol molecules. The role of cholesterol is particularly important, as cholesterol is considered as one of the main regulators of the fluid-like nature of membranes that is the key to their various dynamic properties. The importance of cholesterol is further emphasized by the recent view by Simons and Ikonen [231], who proposed the so-called raft model for biological membranes. The raft model is in part based on the (rather old) idea of lateral organization of heterogeneities in cellular membranes, and second, for an assumption that sphingolipids and cholesterol segregate in membranes to form microdomains termed rafts. Cholesterol seems to be crucial in the formation of rafts, which have been reported to be rich in sphingolipids, cholesterol, and saturated lipids. Recent works have proposed rafts to be enriched in some integral proteins as well as in GPI-anchored (glycosylphosphatidylinositol-anchored) proteins in the outer monolayer of a plasma membrane. As a consequence, rafts have been thought to act as platforms for adhesion and signaling, as they confine proteins involved in a signal transduction event into the same microdomain. Anyhow, despite many studies, unambiguous evidence for the existence of rafts is currently lacking. For a more thorough account of this topic the reader is referred to reviews in Refs. [232–235]. Nevertheless, without doubt, the role of cholesterol in lipid membranes is clear: There is a variety of reasons to consider cholesterol as one of the most important (lipid) molecules in cells.

Having said this, let us consider a PC/cholesterol bilayer system as an example to demonstrate how MD can be used to probe complex biologically relevant systems. As the space is limited, we focus on a few structural quantities that are closely related to experiments.

Figure 15 shows a snapshot of a DPPC/cholesterol bilayer after a 100-ns MD simulation [236]. To better understand the structure of the bilayer, mass density profiles along the normal direction of the membrane are shown in Fig. 16 [237]. Experimentally, information of the same kind can be obtained by, for example, x-ray diffraction measurements that yield electron density profiles of given bilayer systems [157].

The center of the bilayer is at $z = 0$; thus we find that the thickness of the membrane in this case is about 5 nm. Water penetrates rather deeply into the polar headgroup region of the bilayer, where the phosphate and choline groups of DPPC molecules (as well as the hydroxyl groups of cholesterol molecules) reside, whereas water does not favor the hydrophobic region characterized by the DPPC tails and the steroid ring structure of cholesterol. Rather, the cholesterol molecule prefers to locate itself such that the polar OH-group is in the vicinity of the water-membrane interface and the steroid structure is in close contact with the acyl chains of DPPC molecules. The density at the center of the membrane is rather small, implying that the free volume and free voids are presumably largest at this part of the system. One might then assume that the intermonolayer friction is very small. We are not aware of related experiments, however; thus, this idea should be taken as suggestive. What is definitely clear, however, is that the headgroup-water interface is the most dense part of
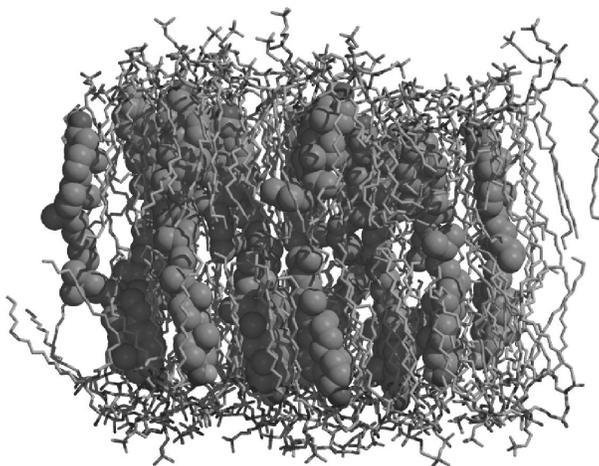
Figure 15. A snapshot of a lipid bilayer of dipalmitoyphosphatidylcholine and cholesterol molecules. The cholesterol mole fraction is 20%. For clarity, water is not shown.

the system. For one-component PC systems, the pressure across the membrane also seems to largest at the interface [180].

As the rigid steroid ring system of cholesterol is in close contact with the DPPC tails, one might expect that this has some influence on the ordering of lipid tails. This is precisely what happens as is depicted in Fig. 17. There, the order parameter S_n describes the ordering of the hydrocarbon tails in DPPC molecules following a definition based on the second order Legendre polynomial. In this case, $S_n = \frac{1}{2}\langle 3\cos^2\theta_n - 1\rangle$, where θ_n is the angle between the bilayer normal and the orientation of the vector along a C–H bond of the nth carbon atom of the hydrocarbon chain [86]. The same order parameter is commonly adapted to studies of liquid crystals. In membrane systems, S_n can be determined through nuclear magnetic resonance measurements. In this case it suffices to mention that $S_n = 0.5$ corresponds to full ordering (where acyl chains are full-trans—like a "zig-zag" conformation), whereas $S_n = 0$ corresponds to a disordered chain, in which case the hydrocarbon chain acts like a random walker. The main finding here is that for this system in the fluid-like state, above the so-called main phase transition temperature, cholesterol orders the acyl chains of
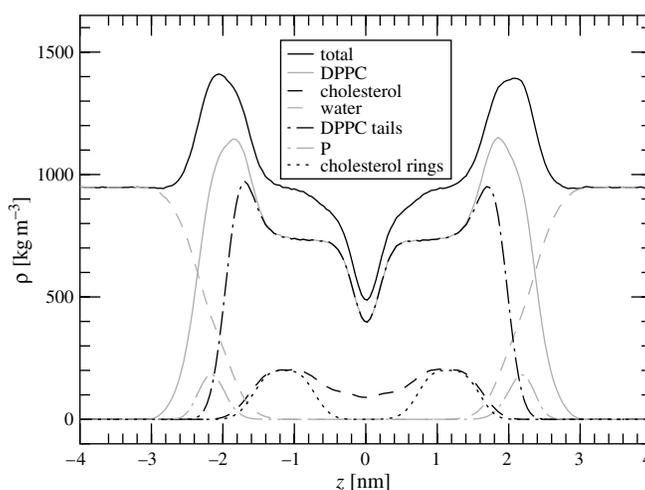
Figure 16. Mass density profiles of various structural components in a dipalmitoylphosphatidylcholine (DPPC)/cholesterol bilayer with a cholesterol mole fraction of 20% [237]. The results are shown across the bilayer in the z-direction of membrane normal. The case $z = 0$ corresponds to the center of the lipid bilayer. The different curves represent total mass density (full black), DPPC (full gray), cholesterol (dashed black), water (dashed gray), DPPC tails (dash-dotted black), phosphor atoms (P) (dash-dotted gray), and cholesterol rings (dotted black).
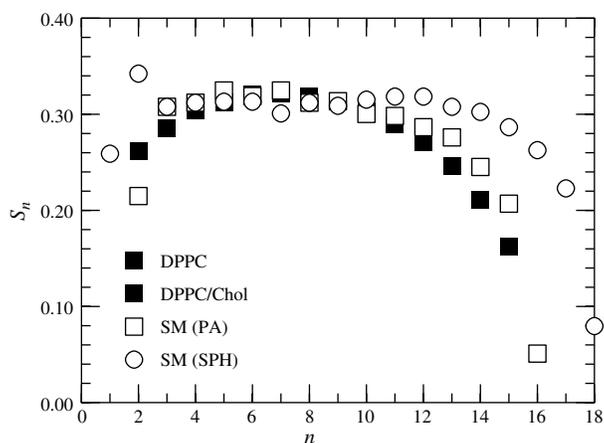
Figure 17. Order parameter S_n of the hydrocarbon chains of lipid molecules. Small n corresponds to carbons (methylene groups) close to the headgroup, whereas large n corresponds to carbons close to the center of the bilayer. In addition to the pure DPPC system [175], results are shown for a DPPC/cholesterol system at a cholesterol concentration of 20% [236] and for a pure SM bilayer (data taken from Ref. [248]). In the case of SM, the order parameters along both the sphingosine (SPH, 18:1) and palmitoyl (PA, 16:0) chains are given.

DPPC molecules. As recent MD simulations for lipid bilayer mixtures of glycerophospholipids and cholesterol have revealed, the enhancement of acyl chain ordering resulting from to an increasing cholesterol concentration is related to a number of intriguing phenomena in membranes [236–246]. For example, as the cholesterol concentration is increased, the area per molecule in the plane of the bilayer decreases, the permeation of molecules across the bilayer is affected because of changes in the free energy profile, the likelihood of gauche conformations of lipid acyl chains are affected considerably, the lateral diffusion of lipids and cholesterol molecules is reduced, and the bilayer in general becomes more rigid. These observations are in good agreement with experimental findings.

As far as other lipid membrane systems are concerned, the key issue is that MD simulations provide a versatile and essentially unique tool to analyze membrane properties in atomic detail. Basically, all possible issues (excluding electronic degrees of freedom) can be determined in full detail. In addition, as experimental techniques typically consider macroscopic aspects of the system, or the averages taken in experimental measurements are global ones, MD simulations allow one to gain a much deeper understanding of the system. In addition to global averages, one can calculate local properties, such as the influence of cholesterol, on the ordering of phospholipids that are neighbors to the given cholesterol [244]. One can also follow the trajectories of individual molecules and signaling events in time, and thus, for example, resolve the diffusion mechanisms and diffusion pathways of molecules migrating inside lipid bilayers. The work by Söderhäll and Laaksonen of the diffusion of ubiquinone inside a DMPC bilayer provides an excellent recent example of this issue [247]. Now, as basically every atomic detail is within reach, does this imply that MD is the method of choice without any limits? It does not. There are two major issues that limit the use of molecular dynamics for simulations of biomolecular systems. First, the development of force fields together with rigorous testing to confirm their validity, and second, the resources. The first of these is a general issue with which one has to learn to live—it is and it will be part of our lives in all cases in which new models are being designed and tested. The second one will get better in time as the development of algorithms and the progress in computer technology permits computer simulations of more and more challenging problems.

Now, let us come back to the idea of lipid rafts. The second major component in rafts are sphingolipids, and SM in particular. It is therefore rather surprising that there are just a few studies that have explored SM systems through MD simulations [248–251]. In essence, one has found that SM is distinctly different from other lipids such as PCs in the sense that its hydrogen-bonding capacity is substantial. Unlike PCs, for example, SM molecules are able to act both as a donor and as an acceptor, thus allowing strong intramolecular as well as intermolecular hydrogen bonding. This leads to compact SM bilayers in which the area per

molecule in the plane of the membrane is much smaller than in PC systems, which in turn is expected to play a key role in various dynamical processes such as lateral and rotational diffusion as well as permeation of molecules across SM bilayers. Further, the small area per molecule and the concominant reduction of free volume inside SM bilayers is manifested in the ordering of SM hydrocarbon tails. As shown in Fig. 17, the order parameter S_n for a pure SM bilayer is about 50% larger than that for a pure DPPC bilayer, and approximately of the same magnitude as the order parameter for a DPPC/cholesterol mixture with a cholesterol content of 20%.

As there are no published computational studies of lipid bilayers with both cholesterol and SM, essentially all relevant questions with regard to the atomic perspective of lipid raft systems remain to be posed. We are looking forward to future work that hopefully will address these issues, together with many other questions related to many-component lipid membrane systems in general.

5.3. Recent Studies of Lipid Membranes Through Molecular Dynamics

Developments in the field of MD simulations of model lipid bilayer systems have been discussed in a number of rather recent review articles [84–88]. In this work, we therefore focus on the advances made in the last few years.

Most studies in the field of model membranes have concentrated on the "high-temperature case" characterized by the fluid-like L_α phase above the main phase transition temperature. Much less attention has been paid to the low-temperature gel phase characterized by ordering in the plane of the membrane, as well as conformational ordering of the lipid acyl chains. The studies by Venable et al. [178] and Snyder et al. [252] provide an exception to the rule. Although certain observations such as tilting of lipid acyl chains and reduced area per molecule can be made in agreement with experimental data, a detailed and solid analysis of the low-temperature phase is very difficult to perform because of the dynamics, which are very slow in this phase.

Rather extensive MD investigations over the years have lead to a good understanding of bilayers composed of saturated lipids, whereas the understanding of the properties of bilayers composed of unsaturated lipids has been much more limited. Recently, though, there has been increasing interest in investigating these systems [253–255]. This is simply because unsaturated lipids, and polyunsaturated lipids in particular, are an essential component of some cellular membranes. For example, the cell membrane of the nervous system is known to contain a large fraction of polyunsaturated docosahexaenoic, a fatty acid with six double bonds. Polyunsaturated lipids are further thought to play a key role in neurological dysfunctions such as Parkinson's disease and in diseases such as atherosclerosis and cancer. Although there is a long way to go until the properties of these systems are understood in detail, MD studies have already demonstrated that bilayers comprising polyunsaturated lipids are distinctly different from bilayers composed of saturated lipids. Among others, they have shown that polyunsaturated lipids in bilayers are highly flexible and characterized by a high degree of conformational disorder. As a consequence, as these findings imply that the influence of polyunsaturated lipids on the structure within a bilayer is considerable, it has been suggested that polyunsaturated lipids are particularly relevant for the proper functioning of membrane proteins [253].

Phospholipid (Langmuir) monolayers are rather commonly used as model systems of lipid bilayers. This approach is definitely attractive, as the monolayer, being half of a membrane, is a well-defined planar system that is easier to control than model vesicles and liposomes. As a consequence, it has been adopted to a wide range of studies, including intermolecular interactions between various molecules in monolayers, drug delivery issues, and penetration of peptides into lipid monolayers [256, 257]. The actual correspondence between quantities measured for bilayers and monolayers is not often clear, however, which may render interpretation of the data. In a recent study, Kaznessis et al. [258] employed MD simulations to investigate lipid monolayers at the water/air interface and found evidence that some structural quantities can indeed be different from their counterparts in lipid bilayers. In particular,

they observed that the order parameter S_n for a DPPC monolayer was essentially constant ($S_n \approx 0.2$) for carbons 2–14. In DPPC bilayers the behavior of S_n is markedly different (see Fig. 17). Further studies would be highly welcome to clarify the monolayer versus bilayer issue in more detail.

Biological membranes at physiological pH are surrounded by an aqueous buffer containing ions such as Na^+, Ca^{2+}, and Cl^-. At the same time, about 10% of the lipids in biomembranes are charged (typically anionic), which implies that their charge has to be compensated by counterions such as Na^+. Recent MD studies [259–261] have thus aimed to clarifing the effect of salt on lipid bilayers. As Boeckmann et al. demonstrated very clearly [259], this task is computationally very demanding. This is because relaxation times associated with the diffusion of ions are very large, which in turn implies that the timescales required for equilibration and gathering sufficient statistics are substantial (on the order of 100 ns or more). Nevertheless, recent attempts have shown that salt plays a major role in determining the structural and electrostatic properties of lipid bilayers and has a major effect on the lateral diffusion of lipids in a bilayer.

In addition to the above studies for pure lipid bilayers, we would like to mention recent work on cationic lipid bilayers [151, 262] because of their relevance in gene therapy, drug delivery, and specifically on DNA–membrane interactions. Some years ago, Bandyopadhyay et al. performed an MD study of a lipid mixture of neutral (zwitterionic) dimyristoylphosphatidylcholine (DMPC) and cationic dimyristoyltrimethylammonium propane (DMTAP) in the presence of a short DNA fragment [151]. It took several years until this very elegant piece of work was complemented by a second study focusing on cationic bilayers. Gurtovenko et al. [262] studied DMPC/DMTAP lipid bilayer mixtures over a wide concentration range of DMTAP and showed that there is a strong interplay between positively charged TAP headgroups and the zwitterionic PC heads. This interplay was shown to lead to considerable reorientation of PC headgroups for an increasing DMTAP concentration, in agreement with experimental data. It is likely that these observations play a significant role in the condensation of DNA onto the membrane surface.

Biomembranes are semipermeable barriers that control the flow of ions and other molecular species into and out of the cell. As a consequence, the permeation of small molecules such as water and ions has been studied rather extensively during the last decade. Recently, more ambitious attempts have been made to study the permeation and distribution of more complex molecules such as sugars, amphiphilic drugs, and anesthetic molecules in lipid bilayers [263–267]. It is evident that the influence of the compound on the structure of lipid bilayers depends on the compound in question. Thus, there is currently no point in drawing any general conclusions. Rather, we would like to emphasize the challenge associated with the modeling of these processes; the study of the distribution of compounds in membranes is in general hindered by the limited timescale accessible through MD simulations.

The same problem is encountered in studies of enzyme–membrane and protein–membrane systems because of large relaxation times associated with the conformational changes of the molecules. Nevertheless, some successful attempts have been made recently to better understand the action and function of different kinds of enzymes and protein molecules in membranes. For example, lipases are enzymes involved in a variety of processes, such as signal transduction and the hydrolysis of some bonds in phospholipids. MD simulations have been able to shed light on the influence of these molecules on the structure of lipid bilayers [268] and on the orientation and conformations of certain enzymes at the water–membrane interface [269]. As far as integral proteins are concerned, let us here mention just a few demonstrative examples of recent activities in the field. Sansom et al. [270, 271] have investigated potassium ion channels that enable rapid movement of K^+ ions passively across a membrane, and Baudry et al. [272] have featured recent progress in the modeling of light-driven proton pumps in the purple membrane. Perhaps the most topical issue, however, is the work based on MD studies of water channels [54, 55]. This example is related to the Nobel Prize in Chemistry 2003 and highlights the remarkable fact of how experiments and computer simulation studies can indeed complement each other.

5.4. From Atomistic to Coarse-Grained Molecular Descriptions of Lipid Membranes

In this chapter, we have in many places faced a fact that biological systems are characterized by a variety of different length and timescales. For the shortest ones, there are indeed techniques based on quantum mechanics and atomic-scale MD that describe systems in full atomic detail. As a consequence, they are obvious techniques for studies of microscopic properties. However, these approaches are not feasible for cases in which the length scales are larger than $\sim$10 nm and the timescales are beyond $\sim$100–1000 ns.

What we can do under these circumstances is to replace the atomic-level description of the system with a less detailed one. (See discussion in Section 3.) In practice, one still employs MD simulations for a system of interacting particles, using precisely the same framework as in atomic-level MD. The only difference is the molecular description and the choice of interactions. Although the above discussion has been based on an atomic view, in coarse-grained MD the molecular description does not account for atomic details. Rather, the particles describe clusters of atoms rather than individual ones.

One example of this approach is the recent work by Marrink et al. [273]. They used the Gromacs simulation package (first developed for atomic MD) to study membranes comprising of lipid molecules. The particles in this description did not represent atoms, however. Rather, they represented clusters of methyl groups or clusters of atoms in the polar head-group, thus reducing the computational complexity of the model significantly. Goetz et al. [274, 275], Shelley et al. [276–278], and Kranenburg et al. [120, 121] have recently pursued ideas of a similar nature. Another approach is to calculate material properties from atomic-scale MD simulations and to then use them to parameterize mesoscale and macroscale models. Ayton et al. have recently applied this idea to study elastic properties of membrane systems beyond the atomic regime, including both mesoscopic [279] and macroscopic levels [280–283].

Overall, the key word in these simplified descriptions is coarse graining, which takes us higher in the hierarchy of molecular simulations and allows one to design and study mesoscopic, and possibly even macroscopic, models for biological membranes and other soft matter systems.

6. SIMULATIONS OF SOFT MATTER SYSTEMS OVER MESOSCOPIC SCALES: COARSE-GRAINING OF IONIC SOLUTIONS

Thus far we have discussed molecular level modeling in detail and has reviewed some of the recent developments in multiscale modeling in soft matter and biological systems. Further, we have discussed some practical aspects such as the importance of high-quality random numbers in numerical simulations and the use of MD simulations for microscopic studies of biophysical systems. Here, we take one step further and discuss how soft matter systems can be coarse grained and modeled over scales far larger than those accessible through atomic-level MD. More specifically, we use one particular coarse-graining method, namely the IMC method (see Section 3.1) to combine micro- and mesoscale approaches to demonstrate coarse-graining in practice. We discuss the merits and problems of this method and put it into a more general context. The method we have chosen is applied to an ionic solution. Despite its apparent simplicity, this system, consisting of water and NaCl only, is nontrivial from both a practical and a theoretical point of view.

6.1. Coarse-Graining Strategy

The approach we present here is conceptually simple and systematic, combining classical molecular dynamics (MD) and a momentum conserving DPD thermostat (i.e., we obtain the conservative potentials through a coarse-graining procedure from MD simulations and employ DPD as a thermostat in a manner discussed in Section 3.2. Figure 18 shows an illustration of the original and coarse-grained systems.
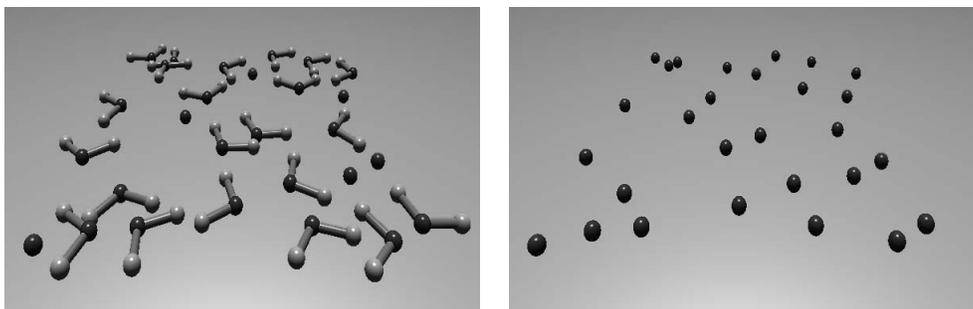
Figure 18. The coarse-grained and the underlying molecular system. Left: aqueous NaCl with atomic detail. Right: the coarse-grained model in which waters have been replaced with just one particle.

This method has several advantages, as will be discussed below, and it consists of three independent parts: first, MD simulations are carried out to obtain radial distribution functions $g(r)$ between different atoms, molecules, or molecular groups; second, the IMC procedure (see Section 3.1 for details) is applied to invert the radial distribution functions to obtain effective interaction potentials $V^{\text{eff}}(r)$ between the new interaction sites; and finally, the effective potentials are used within the DPD algorithm (see Section 3.2 for details) to study the long time- and length scale properties with full hydrodynamics for the new coarse-grained system.

One of the main advantages of this approach is that it allows easy tuning of the level of coarse-graining while preserving the essential molecular information. It also provides a well-defined and systematic coupling between the microscopic and coarse-grained systems in the sense that the pair correlation properties of the two descriptions are consistent. Although we have not checked that, this indicates that the coarse-graining procedure proposed here also preserves the essential information of phase behavior. Furthermore, this method speeds up simulation times considerably, thus making it possible to examine large-scale properties of complex molecular systems with full hydrodynamics. Finally, it allows one to study the limits of coarse-graining (i.e., how coarse-graining depends on physically relevant variables such as the density, temperature, pressure, and salt concentration).

6.2. Obtaining the Interaction Potentials

Let us first validate the method against the underlying microscopic MD simulations. To be self-consistent, the coarse-grained potentials should produce the same pair correlation functions. It is important to note that coarse-graining of even simple liquids is not a simple feat, especially if one wants to reproduce both static and dynamic properties. This is where the approach and results presented here are particularly valuable.

For the underlying microscopic MD simulations, we used the flexible SPC water model [284] and the so-called Smith–Dang parameters for Na$^+$ and Cl$^-$ ions [285]; that is, the Lennard–Jones parameters for the sodium ions were $\sigma = 2.35$ Å and $\varepsilon = 0.544$ kJ/mol, and for chloride $\sigma = 4.4$ Å and $\varepsilon = 0.419$ kJ/mol. The temperature was set to 300 K, the simulations were done in the NVT ensemble, and we used a salt concentration of 0.87 M. Electrostatic interactions were computed using the Ewald summation. The full set of simulation parameters are given in [109, 286]. As the central quantity, the radial distribution functions between all different pairs of particles were computed from the MD simulations. They were used as an input in the IMC procedure (described in Section 3.1). The new coarse-grained interaction potentials are shown in Fig. 19. The potentials are not as "soft" as the DPD potentials in Fig. 3 but have extra features and a relatively hard core. That simply reflects the degree of coarse-graining and the fact that this method preserves the identity of the coarse-grained entities.

As a little, digression a short comment regarding the model should be said. Whereas there are several well established and extensively studied models for water (see, e.g., Ref. [287] and references therein), flexible SPC being one of them, that is not the case for NaCl. The chosen Smith–Dang parameterization is hand optimized for NaCl and reproduces some of the
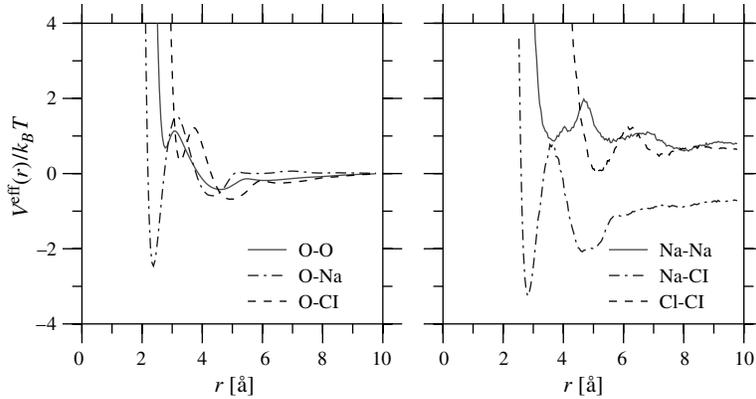
Figure 19. The effective interaction $V^{\text{eff}}(r)$ versus the radial distance r for all pairs of particles as obtained from the radial distribution functions of MD simulations at 0.87 M salt concentration. "O" refers to the oxygen atom in a water molecule.

properties well. In a recent study [288], we compared a number of force fields for simulations of NaCl in aqueous solutions and found that there are some important differences between them. We are pointing this out because one should always validate the coarse-grained results against the particular model used to obtain the new potentials.

A comparison of the effective interactions to the potentials of mean force revealed [288] that the overall structures of the potentials of mean force were roughly similar to those of the effective interactions—as expected—but there were significant differences in the form and amplitudes of some of the peaks. Furthermore, it is important to notice that the long-range tails of $V^{\text{eff}}(r)$ include a contribution from Coulombic forces, whereas the tails of the potentials of mean force do not (i.e., they decay rapidly to 0 for small r, resembling the behavior typical for the commonly used Derjaguin–Landau–Verwey–Overbeek [DLVO] theory for colloids; see, e.g., Ref. [289]). From the comparison with the results in Ref. [288], we conclude that the potential of mean force serves as a reasonable approximation of $V^{\text{eff}}(r)$ but is distinctly different quantitatively. This difference becomes more pronounced for charged particles for which even the qualitative behavior may be different.

At this point, we would like to stress the following six features:

1. All DPD calculations that will be discussed below were carried out with the $V^{\text{eff}}(r)$ obtained from the MD simulations at 0.87 M salt concentration. As was noticed previously [286], the effective potentials are rather weakly affected by a change of salt concentration (i.e., it is not necessary to recompute the effective interactions for all different salt concentrations).

2. We used a cutoff $r_c = 9.6$ Å in all the cases. This was done to enhance the efficiency of the model, and it did not give rise to artifacts in, for example, $g(r)$, as explicit tests confirmed. If larger cutoffs are desired, one can either compute $V^{\text{eff}}(r)$ over a larger interparticle distance or approximate the long-distance part by a Coulombic tail [109].

3. The electrostatic interactions that were explicitly accounted for in the MD simulation are now implicitly included in the effective interactions. This is reflected in the shape of the potentials.

4. DPD simulations were carried out with coarse-grained water molecules, each of them represented by a single spherical particle. The dynamics resulting from hydrogen bonds are implicitly described by the effective interactions.

5. The $V^{\text{eff}}(r)$ have a relatively soft core; that is, the potentials shown in Fig. 19 were extrapolated to a finite value as $r \to 0$.

6. The weight functions were chosen as to be of the standard form $\omega^R(r) = [1 - r/r_c]$ (see Section 3.2). The dissipation strength γ [see Eq. (13)] was determined such that the decay rate of the single-particle velocity correlation function at early times was approximately identical in MD and DPD simulations. That yielded $\gamma = 0.72$, which was used in simulations. Note that γ is the only fitting parameter in this model.

Whereas the first five items are straightforward, the last item is the critical step that connects the DPD thermostat to the MD simulations through the γ parameter. It is important to understand that there is no unique way to fix the strength of the dissipative force γ. Here, we adjusted the rate of the dynamics by finding an optimal γ in the sense that the initial decay rate of the single-particle velocity correlation function,

$$\phi(t) \equiv \langle \vec{v}_i(t) \cdot \vec{v}_i(0) \rangle \tag{36}$$

(separately for Na^+, Cl^-, and H_2O), was approximately similar in MD and DPD calculations at 0.87 M. This way, γ does not fix the long-time decay of $\phi(t)$s, and hence does not fix the diffusion coefficients. This scaling is in agreement with the pair collision theory [290]. The single-particle velocity autocorrelation functions are shown in Fig. 20.

6.3. Comparison Between Molecular Dynamics and Coarse-Grained Simulations

To check the consistency of our coarse-grained approach with the MD simulations, we first compared the radial distribution functions $g(r)$ of the two approaches. As can be seen in Fig. 21, the results for the $g(r)$s are in excellent quantitative agreement. In addition to this, the coordination numbers [219] were found to be in perfect agreement with those measured in the MD simulations. This proves that the method is self-consistent, at least in terms of static properties.

Even when static properties are well reproduced, there is no reason for the dynamics of the coarse-graing system to resemble the true dynamics of the system [32]. To test the dynamical properties of the system, we measured the tracer diffusion coefficient D using the Einstein relation for mean-square displacements,

$$D = \lim_{t \to \infty} \frac{1}{6\,t} \langle [\vec{r}_i(t) - \vec{r}_i(0)]^2 \rangle \tag{37}$$

which can provide us with information of possible problems on the dynamics of the system. Here $\vec{r}_i(t)$ is the position of a tagged particle at time t. To study dynamics even further, we have also computed the shear viscosity by a Green–Kubo relation [219]. Dynamical properties are a different matter, and there is no reason to expect, *a priori*, for them to match. This is particularly so in systems in which water and charges, or polarizable particles, are present, as hydrogen bonds are directional.

The results for tracer diffusion are presented in Fig. 22. It should be noted that in all of the coarse-grained DPD, we used the potentials obtained at 0.87 M salt concentration and shown in Fig. 19. Thus, this also tests the applicability of the potentials.

From Fig. 22, it is clear that within the statistical error the qualitative behavior of the diffusion coefficients versus the salt concentration is virtually identical in MD and DPD
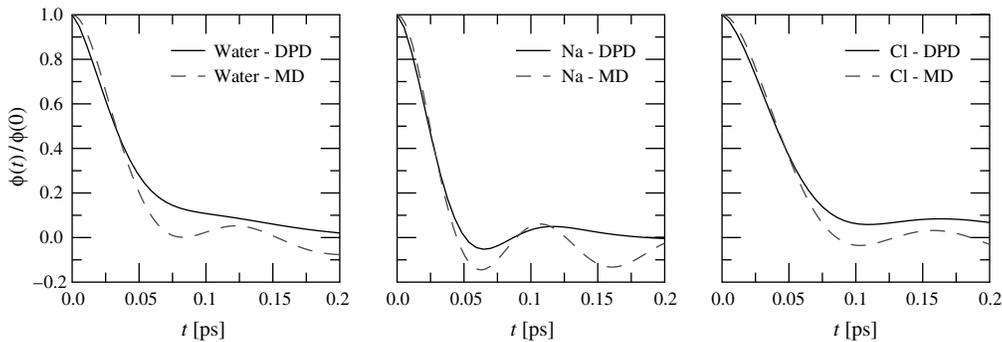
Figure 20. The decay of the single-particle velocity autocorrelation function $\phi(t)$ at early times. Results shown here are for water, Na^+, and Cl^- ions. As the data illustrates, the early-time decay of $\phi(t)$ is essentially identical between molecular density and dissipative particle dynamics (DPD) simulations for $\gamma = 0.72$ used in the DPD simulations. Adapted from Ref. [219].
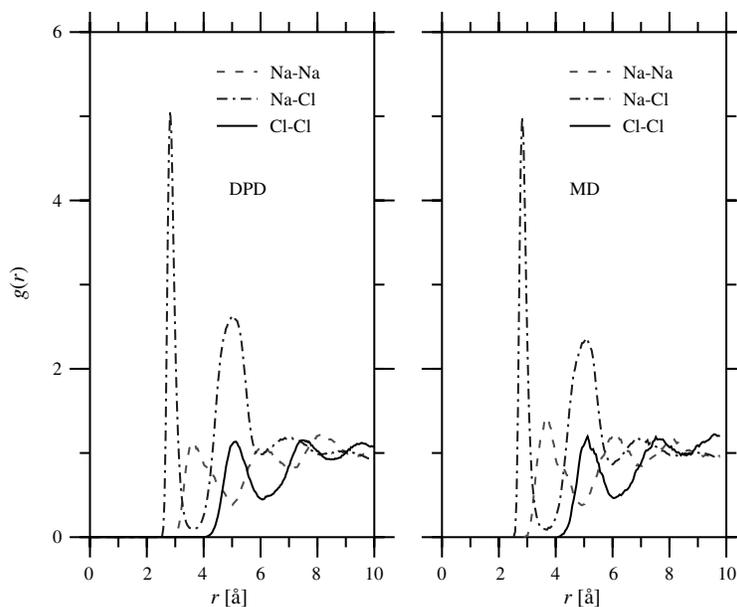
Figure 21. The radial distribution functions obtained from the coarse-grained simulation using the DPD thermostat with the effective potentials (left), and the $g(r)$ obtained from the microscopic MD simulations (right). The agreement between the two approaches is remarkably good.

simulations. Even at lower molarities, the deviations are small and the respective tracer diffusion coefficients for Na^+, Cl^-, and H_2O have the same overall behavior in both cases. The increasing deviations at low molarities are the result of the growing screening length. Furthermore, the diffusion coefficients are in good agreement with Langevin simulations (using the generalized Langevin equation formalism) and experimental results [291]. The agreement with experiments is mainly the result of MD simulations. The true test for our DPD-based approach is the comparison against the MD simulations that were used to obtain the effective potentials. As the comparison clearly shows, the coarse-grained method performs very well indeed.
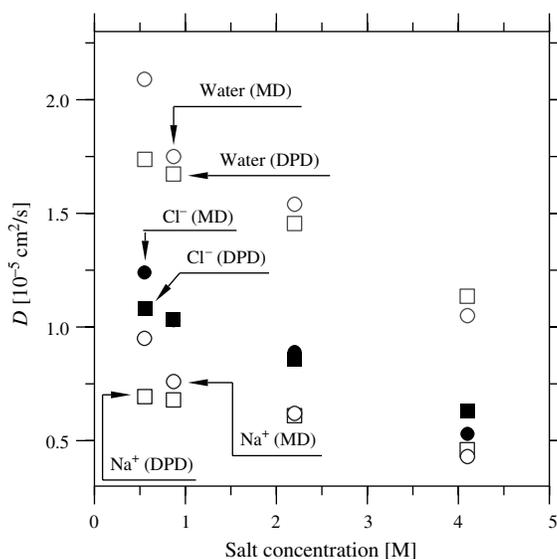
Figure 22. Comparison of the tracer diffusion coefficients obtained from molecular dynamics and dissipative particle dynamics (DPD) simulations. The DPD simulations are based on effective interactions extracted from the MD studies at 0.87 M. These effective interactions were used in all consequent DPD simulations with a varying salt concentration. DPD data has been scaled to allow comparison. Error bars are about the size of the symbols.

Our initial studies for the temperature dependence of D through DPD simulations further support the consistency of this approach, as we found that the effective diffusion barriers, obtained through an Arrhenius analysis of D around 300 K at 1 M, of H_2O and Cl^- were essentially identical, whereas the barrier of Na^+ was about 15% larger, in agreement with MD simulations. In our previous study [219], we also studied the behavior of the shear viscosity coefficient. It was found that it increases monotonously for an increasing salt concentration, in agreement with experiments [292]. The shear viscosity coefficient was not determined from MD simulations, as it would have been too time consuming.

6.4. Computational Aspects

In addition to the above physical aspects, there are important computational gains as well. The coarse-graining procedure, even at this very modest level, led to a speed-up of the order of 20, as the CPU times for MD and DPD simulations were almost identical but the particle numbers studied by MD and DPD were about 300 and 7000, respectively (at the same densities and simulation times of about 2 ns). As a result of the coarse-graining procedure, the fast internal and orientational degrees of freedom related to water have been integrated out leading to the speed-up.

We expect that in more complex systems where the molecular description can be coarse-grained more than here, leading to a molecular description with a smaller number of particles and softer interparticle interactions, the speedup will be significantly larger. A further important aspect is that once the MD simulation for obtaining the potentials has been done, the stored equilibrium configurations can be reused to obtain new effective potentials; that is, between different groupings of molecules. This provides a straightforward approach to study the effect of the level of coarse-graining on the properties of the system. In the case of lipid molecules, for example, such studies may reveal the essential interaction centers (such as molecular groups with a finite charge) and the key features of molecular structure that are crucial for the behavior of the system. Obviously, these features should then be accounted for in the coarse-grained model. Further, it is noteworthy that this methodology can be applied to a wide range of scales starting from *ab initio* simulations.

6.5. Discussion and Relation to Other Methods

In this section we have discussed in detail the linking of classical molecular dynamics simulations to a higher-level description via the IMC method [109]. We chose to have only a very modest level of coarse-graining to be able to better validate our results against the microscopic MD simulations.

The agreement between the present coarse-grained approach with the MD results and experiments clearly demonstrates that the method itself has a sound physical foundation as well as potential for further applications. In particular, despite the coarse-graining of both the particles and the interactions, the results show that the method retains the characteristics of different molecules, or molecular groups, and the essential aspects of interparticle interactions. Furthermore, although the level of coarse-graining used here has been quite modest, the results do demonstrate that the method can reduce the computational load significantly as compared to detailed molecular models.

Here, the computational gain was a factor of 20. This may sound modest but one should keep in mind that, unlike the MD simulations, the mesoscopic simulations were performed using an unoptimized simulation code and that the factor of 20 alone would be able to push us to the microsecond scale to enable studies of, for example, dynamic properties. Thus, this approach allows one to study soft matter systems over mesoscopic scales beyond the atomistic regime. It is also noteworthy that as DPD satisfies momentum conservation, this approach allows studies of a given system with full hydrodynamics that is essential in various soft matter processes. Thus, we feel that the approach presented here provides a good starting point and a platform for further studies of macromolecular systems and of the general question of bridging different time and length scales.

An important and general question regards the limitations of coarse-graining, and how much does it depend on physically relevant variables such as density, temperature, pressure,

and salt concentration. Although we lack general knowledge, the results presented here indicate that it is not necessary to obtain the effective potentials for each condition studied. This idea is supported by further ongoing studies of lipid systems (DMPC) and various solutions (LiCl, $CaCl_2$), in which case we have studied the effects of varying density and temperature. The results from these simulations are in favor of our conclusions about the generality of this approach. One of the main issues in regard to charge carrying systems is to check that the coarse-grained systems actually obey the Stillinger–Lovett sum rules [89] and thus charge conservation. Work is in progress to address that question.

We are currently applying this method to more complex biomolecular systems of amphiphilic lipid molecules in explicit water, with a stronger focus on the coarse-graining. Macromolecular systems pose new conceptual problems, and lipids, because of their amphiphilic nature, are particularly challenging. For this reason, at the moment no systematic method for coarse-graining lipids exists. We are currently extending the hierarchical approach presented here to lipids as well.

Finally, we would like to note that the method discussed here bears a very close resemblance to the approach by Müller–Plathe and coworkers in the context of polymeric systems. Elegant reviews on coarse-graining of polymeric systems are provided in Refs. [29, 32]. As far as the coarse-graining of lipid systems is concerned, the work of Shelley et al. presents the current state-of-the-art concerning this issue [276].

7. CONCLUDING REMARKS

Liquid crystals, paints, polymers, clay, foam, complex fluids, and surfactants share many things in common. Most important, these materials are distinctly different from liquids and solid condensed matter systems for which reason an idea has emerged that they constitute a new class of materials called *soft matter*. What is common to soft matter systems is that they are all characterized by weak interactions of the order of thermal energy. Because $\langle k_B T \rangle \approx 26$ meV, the energy scale associated with soft matter is well below typical interaction energies in condensed matter systems, where they are on the order of 1 eV. Second, what is peculiar to soft matter is a multitude of fascinating phenomena, many of which deal with biological issues and are therefore related to living matter and life sciences. This point is truly challenging because biological systems in particular are characterized by extremely wide length and timescales, ranging from femtoseconds to several days and from nanometers to several meters.

Without doubt, our understanding of soft matter and biological systems is largely based on experiments. Nevertheless, it is a plain fact that experiments cannot live on their own. First of all, the resolution of experimental approaches in time and space is often limited, which implies that there are always issues that cannot be probed by experiments. The interpretation of experimental results is further always based on some model, thus stressing the importance of theoretical work that can provide one with an understanding of the laws of nature that govern these systems. Analytical theory, however, is also unable to live on its own. In addition to the models whose underlying assumptions can be questioned, the analytical approach is troubled by the fact that analytical calculations are often based on approximations whose validity may be difficult to assess. To resolve these problems, and concurrently to bridge theory and experiments more closely together, it has become common to carry out computer simulations based on well-defined theoretical models.

The models used in computer simulations can be quantum mechanical, and thus be able to address questions where electronic degrees of freedom are important. However, if electronic degrees of freedom are not crucial, a somewhat simpler approach using classical molecular dynamics in atomic level is perhaps more appropriate. For phenomena that take place over much larger scales, even more coarse-grained approaches are needed, leading us to mesoscale and macroscale simulation approaches. Thus, the fact that we are dealing with a variety of different time- and length scales in biological systems implies that one needs to develop and employ multiscale modeling techniques, an approach in which different kinds of computer simulation methods are bridged together through coarse-graining approaches.

In this review, our aim has been to present an overview of the essential underlying ideas related to molecular modeling, molecular dynamics simulations in atomic detail, and mesoscopic simulation methods probing systems beyond the atomic regime. In addition, we have discussed coarse-graining techniques whose aim is to reduce the number of degrees of freedom and consequently to yield less detailed models for studies of large-scale properties. Although this overview is focused on some techniques that we have found most appropriate for the present case, we hope that it has been able to highlight the challenges and future prospects related to the field.

We consider it a fact that molecular modeling in general is a unique tool that can provide a great deal of insight into both the microscopic and the large-scale properties of biologically relevant soft matter systems. The numerous applications in biosciences as well as in technologically related soft matter fields are some of the reasons that justify and motivate this work. Yet the challenge related to understanding such extremely rich problems, many of which are of fundamental nature, provides another and even a more important reason.

ACKNOWLEDGMENTS

REFERENCES

1. M. Daoud and C. E. Williams, Eds., "Soft Matter Physics." Springer, Berlin, 1999.
2. H. Flyvjberg, J. Hertz, M. H. Jensen, O. G. Mouritsen, and K. Sneppen, Eds., "Physics of Biological Systems: From Molecules to Species." Springer, Berlin, 1997.
3. M. E. Cates and M. R. Evans, Eds., "Soft and Fragile Matter: Nonequilibrium Dynamics, Metastability and Flow." Institute of Physics Publishing, Bristol, 2000.
4. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. "Molecular Biology of the Cell," 3rd edn. Garland, New York, 1994.
5. C. Branden and J. Tooze, "Introduction to Protein Structure," 2nd edn., Garland, New York, 1999.
6. R. L. C. Akkermans and W. J. Briels, *J. Chem. Phys.* 113, 6409 (2000).
7. P. Español and F. Vázquez, *Phil. Trans. R. Soc. Lond. A* 360, 1 (2002).
8. E. G. Flekkøy, P. V. Coveney, and G. De Fabritiis, *Phys. Rev. E* 62, 2140 (2000).
9. B. M. Forrest and U. W. Suter, *J. Chem. Phys.* 102, 7256 (1995).
10. R. D. Groot and K. L. Rabone, *Biophys. J.* 81, 725 (2001).
11. R. D. Groot and P. B. Warren, *J. Chem. Phys.* 107, 4423 (1997).
12. A. Malevanets and R. Kapral, *J. Chem. Phys.* 110, 8605 (1999).
13. A. Malevanets and R. Kapral, *J. Chem. Phys.* 112, 7260 (2000).
14. M. Murat and K. Kremer, *J. Chem. Phys.* 108, 4340 (1998).
15. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, *J. Phys. Chem. B* 105, 4464 (2001).
16. W. Tschöp, K. Kremer, J. Batoulis, T. Bürger, and O. Hahn, *Acta Polymer.* 49, 61 (1998).
17. P. B. Warren, *Curr. Opin. Colloid Interface Sci.* 3, 620 (1998).
18. P. Español and P. Warren, *Europhys. Lett.* 30, 191 (1995).
19. P. J. Hoogerbrugge and J. M. V. A. Koelman, *Europhys. Lett.* 19, 155 (1992).
20. J. M. V. A. Koelman and P. J. Hoogerbrugge, *Europhys. Lett.* 21, 363 (1993).
21. E. G. Flekkøy and P. V. Coveney, *Phys. Rev. Lett.* 83, 1775 (1999).
22. T. Ala-Nissila, S. Majaniemi, and K. Elder, in "Novel Methods in Soft Matter Simulations." (M. Karttunen, I. Vattulainen, and A. Lukkarinen, Eds.). Springer, Berlin, 2004 (unpublished).
23. T. Kawakatsu, K. Kawasaki, M. Furusaka, H. Okabayashi, and T. Kanaya, *J. Phys. Condens. Matter* 6, 6385 (1994).
24. N. M. Maurits, Mathematical Modeling of Complex Systems, PhD thesis, Groningen University, 1998.
25. M. Grmela and H. C. Öttinger, *Phys. Rev. E* 56, 6620 (1997).
26. M. Grmela and H. C. Öttinger, *Phys. Rev. E* 56, 6633 (1997).
27. H. C. Öttinger, *Phys. Rev. E* 57, 1416 (1998).

28. M. Karttunen, I. Vattulainen, and A. Lukkarinen, Eds., "Novel Methods in Soft Matter Simulations." Springer, Berlin, 2004 (unpublished).
29. J. Baschnagl, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Müller-Plathe, M. Murat, W. Pauli, S. Santos, U. W. Suter, and V. Tries, *Adv. Polym. Sci.* 152, 41 (2000).
30. W. J. Briels and R. L. C. Akkermans, *Mol. Sim.* 28, 145 (2002).
31. F. Müller-Plathe, *Chem. Phys. Chem.* 3, 754 (2002).
32. F. Müller-Plathe, *Soft Materials* 1, 1 (2003).
33. J. C. Shelley and M. Y. Shelley, *Curr. Opin. Colloid Interface Sci.* 5, 101 (2000).
34. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, P. B. Moore, and M. L. Klein, *J. Phys. Chem. B* 10, 9785 (2001).
35. S. J. Marrink and A. E. Mark, *J. Am. Chem. Soc.* 125, 11144 (2003).
36. M. Kröger, *Phys. Rep.* 390, 453 (2004).
37. P. Nielaba, M. Mareschal, and G. Ciccotti, "Bridging Time Scales: Molecular Simulations for the Next Decade." Springer, Berlin, 2002.
38. ESF Network Simu—Challenges in Molecular Simulations; http://simu.ulb.ac.be.
39. M. Doi, *J. Comp. Appl. Math.* 149, 13 (2002).
40. M. Doi, *Pure Appl. Chem.* 75, 1395 (2003).
41. H. H. Goldstine and A. Goldstine, *Math. Tables Other Aids Comput.* 2, 97 (1946).
42. N. Metropolis, *Los Alamos Sci.* 125 (1987).
43. N. Metropolis and S. Ulam, *J. Am. Stat. Assoc.* 44, 335 (1949).
44. N. Metropolis, A. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* 21, 1087 (1953).
45. N. C. Metropolis, G. Reitwiesner, and J. von Neumann, *Math. Tables Other Aids Comput.* 4, 109 (1950).
46. T. E. Hull and A. R. Dobell, *SIAM Rev.* 4, 230 (1962).
47. D. E. Knuth, "The Art of Computer Programming: Seminumerical Algorithms." 3rd edn., Vol. 2. Addison-Wesley, Reading, MA, 1998.
48. B. J. Alder and T. E. Wainright, *J. Chem. Phys.* 27, 1208 (1957).
49. B. J. Alder and T. E. Wainright, *J. Chem. Phys.* 31, 459 (1959).
50. A. Rahman, *Phys. Rev.* 136A, 405 (1964).
51. J. A. Barker and R. O. Watts, *Chem. Phys. Lett.* 3, 144 (1969).
52. A. Rahman and F. H. Stillinger, *J. Chem. Phys.* 55, 3336 (1971).
53. J. A. McCammon, B. R. Gelin, and M. Karplus, *Nature* 267, 585 (1977).
54. E. Tajkhorshid, P. Nollert, M. Ø. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, *Science* 296, 525 (2002).
55. M. Ø. Jensen, E. Tajkhorshid, and K. Schulten, *Biophys. J.* 85, 2884 (2003).
56. D. Mac Kernan, Interview of Michele Parrinello, Simu News.; available at simu.ulb.ac.be, pp. 5–17.
57. T. Schlick, "Molecular Modeling and Simulation." Springer, New York, 2002.
58. H. W. Kroto, Nobel lecture in December 7, 1996; http://www.nobel.se.
59. R. J. Baxter, "Exactly Solved Models in Statistical Mechanics." Academic Press, London, 1982.
60. D. Frenkel and B. Smit, "Understanding Molecular Simulation: From Algorithms to Applications," 2nd edn. Academic Press, San Diego, 2002.
61. M. P. Allen and D. J. Tildesley, "Computer Simulation of Liquids." Clarendon Press, Oxford, 1993.
62. A. R. Leach, "Molecular Modelling: Principles and Applications," 2nd edn. Prentice Hall, Harlow, 2001.
63. R. G. Parr and W. Yang, *Ann. Rev. Phys. Chem.* 46, 701 (1995).
64. R. Car and M. Parrinello, *Phys. Rev. Lett.* 55, 2471 (1985).
65. M. Eichinger, P. Tavan, J. Hutter, and P. Parrinello, *J. Chem. Phys.* 110, 10452 (1999).
66. W. Andreoni, A. Curioni, and T. Mordasini, *IBM J. Res. Dev.* 45, 397 (2001).
67. G. J. Martyna and M. E. Tuckerman, in "Bridging Time Scales: Molecular Simulations for the Next Decade" (P. Nielaba, M. Mareschal, and G. Ciccotti, Eds.), pp. 381–412. Springer, Berlin, 2002.
68. R. Kapral and G. Ciccotti, in "Bridging Time Scales: Molecular Simulations for the Next Decade" (P. Nielaba, M. Mareschal, and G. Ciccotti, Eds.), pp. 445–472. Springer, Berlin, 2002.
69. F. Jensen, "Introduction to Computational Chemistry." Wiley, Chichester, 1999.
70. J. M. Thijssen, "Computational Physics." Cambridge University Press, Cambridge, 1999.
71. R. Bonneau and D. Baker, *Annu. Rev. Biophys. Biomol. Struct.* 30, 173 (2001).
72. P. Carloni, U. Rothlisberger, and M. Parrinello, *Acc. Chem. Res.* 35, 455 (2002).
73. U. Rothlisberger, M. Sprik, and J. Hutter, in "Bridging Time Scales: Molecular Simulations for the Next Decade" (P. Nielaba, M. Mareschal, and G. Ciccotti, Eds.), pp. 413–444. Springer, Berlin, 2002.
74. S. Y. Reddy, F. Leclerc, and M. Karplus, *Biophys. J.* 84, 1421 (2003).
75. E. Lindahl, B. Hess, and D. van der Spoel, *J. Mol. Model.* 7, 306 (2001).
76. C. Sagui and T. A. Darden, *Annu. Rev. Biophys. Biomol. Struct.* 28, 155 (1999).
77. G. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, *Mol. Phys.* 87, 1117 (1996).
78. B. J. Berne, G. Ciccotti, and D. F. Coker, Eds., "Classical and Quantum Dynamics in Condensed Phase Simulations." World Scientific, Singapore, 1998.
79. E. Giudice and R. Lavery, *Acc. Chem. Res.* 35, 350 (2002).
80. V. Makarov and B. M. Pettitt, *Acc. Chem. Res.* 35, 376 (2002).
81. B. Isralewitz, M. Gao, and K. Schulten, *Curr. Opin. Struct. Biol.* 11, 224 (2001).
82. C. L. Brooks III, *Acc. Chem. Res.* 35, 447 (2002).
83. V. Daggett, *Acc. Chem. Res.* 35, 422 (2002).

84. L. Saiz, S. Bandyopadhyay, and M. L. Klein, *Bioscience Rep.* 22, 151 (2002).
85. L. Saiz and M. L. Klein, *Acc. Chem. Res.* 35, 482 (2002).
86. D. P. Tieleman, S. J. Marrink, and H. J. C. Berendsen, *Biochim. Biophys. Acta* 1331, 235 (1997).
87. S. E. Feller, *Curr. Opin. Colloid Interface Sci.* 5, 217 (2000).
88. H. L. Scott, *Curr. Opin. Struct. Biol.* 12, 495 (2002).
89. J. P. Hansen and I. R. McDonald, "Theory of Simple Liquids." Academic Press, London, 1986.
90. P. G. Bolhuis, A. A. Louis, J.-P. Hansen, and E. J. Meijer, *J. Chem. Phys.* 114, 4296 (2001).
91. A. A. Louis, P. G. Bolhuis, J.-P. Hansen, and E. J. Meijer, *Phys. Rev. Lett.* 82, 2522 (2000).
92. R. L. Henderson, *Phys. Lett.* 49A, 197 (1974).
93. P. Hohenberg and W. Kohn, *Phys. Rev. B* 136, 864 (1964).
94. R. Faller, H. Schmitz, O. Biermann, and F. Müller-Plathe, *J. Comp. Chem.* 20, 1009 (1999).
95. S. Girard and F. Müller-Plathe, *Mol. Phys.* 101, 779 (2003).
96. D. Reith, H. Meyer, and F. Müller-Plathe, *Macromolecules* 34, 2235 (2001).
97. E. Falck, O. Punkkinen, I. Vattulainen, and T. Ala-Nissila, *Phys. Rev. E* 68, 050102(R) (2003).
98. E. Falck, J. Lahtinen, I. Vattulainen, and T. Ala-Nissila, submitted, 2003.
99. S. H. Lee and R. Kapral, *Physica A* 298, 56 (2001).
100. A. Malevanets and J. M. Yeomans, *Europhys. Lett.* 52, 231 (2000).
101. N. Kikuchi, A. Gent, and J. M. Yeomans, *Eur. Phys. J. E* 9, 63 (2002).
102. T. Ihle and D. M. Kroll, *Phys. Rev. E* 63, 020201 (2001).
103. T. Ihle and D. M. Kroll, *Phys. Rev. E* 67, 066705 (2003).
104. T. Ihle and D. M. Kroll, *Phys. Rev. E* 67, 066706 (2003).
105. E. Tüzel, M. Strauss, T. Ihle, and D. M. Kroll, *Phys. Rev. E* 68, 036701 (2003).
106. A. Lamura and G. Gompper, *Eur. Phys. J. E* 9, 477 (2002).
107. N. Kikuchi, C. M. Pooley, J. F. Ryder, and J. M. Yeomans, *J. Chem. Phys.* 119, 6388 (2003).
108. I. Pagonabarraga and D. Frenkel, *J. Chem. Phys.* 105, 5015 (2001).
109. A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* 52, 3730 (1995).
110. N. Goldenfeld, "Lectures on Phase Transitions and the Renormalization Group." Addison-Wesley, Reading, MA, 1992.
111. W. Dzwinel and D. A. Yuen, *J. Coll. Interf. Sci.* 225, 179 (2000).
112. P. Español, *Europhys. Lett.* 39, 605 (1997).
113. P. Español, *Europhys. Lett.* 40, 631 (1997).
114. P. Español, M. Serrano, and I. Zuñiga, *Int. J. Mod. Phys. C* 8, 899 (1997).
115. R. D. Groot and T. J. Madden, *J. Chem. Phys.* 108, 8713 (1998).
116. R. D. Groot, T. J. Madden, and D. J. Tildesley, *J. Chem. Phys.* 110, 9739 (1999).
117. S. Jury, P. Bladon, M. Cates, S. Krishnaa, M. Hagen, N. Ruddock, and P. Warren, *Phys. Chem. Chem. Phys.* 1, 2051 (1999).
118. S. Yamamoto, Y. Maruyama, and S. Hyodo, *J. Chem. Phys.* 116, 5842 (2002).
119. S. Yamamoto and S. Hyodo, *J. Chem. Phys.* 118, 7937 (2003).
120. M. Kranenburg, M. Venturoli, and B. Smit, *Phys. Rev. E* 67, 060901(R) (2003).
121. M. Kranenburg, M. Venturoli, and B. Smit, *J. Phys. Chem. B* 107, 11491 (2003).
122. M. G. Burke, R. Woscholski, and S. N. Yaliraki, *Proc. Natl. Acad. Sci.* 100, 13928 (2003).
123. P. B. Warren, *Phys. Rev. E* 68, 066702 (2003).
124. J. A. Elliot and A. H. Windle, *J. Chem. Phys.* 113, 10367 (2000).
125. G. Besold, I. Vattulainen, M. Karttunen, and J. M. Polson, *Phys. Rev. E* 62, R7611 (2000).
126. P. Nikunen, M. Karttunen, and I. Vattulainen, *Comput. Phys. Commun.* 153, 407 (2003).
127. I. Vattulainen, M. Karttunen, G. Besold, and J. M. Polson, *J. Chem. Phys.* 116, 3967 (2002).
128. T. Soddemann, B. Dunweg, and K. Kremer, *Phys. Rev. E* 68, 046702 (2003).
129. R. D. Groot, *J. Chem. Phys.* 118, 11265 (2003).
130. R. D. Groot, *J. Chem. Phys.* 119, 10454 (2003).
131. C. Sagui and T. Darden, *J. Chem. Phys.* 115, 6578 (2001).
132. C. Sagui, L. G. Pedersen, and T. Darden, *J. Chem. Phys.* 120, 73 (2004).
133. C. P. Lowe, *Europhys. Lett.* 47, 145 (1999).
134. H. C. Andersen, *J. Chem. Phys.* 72, 2384 (1980).
135. P. Nikunen, I. Vattulainen, and M. Karttunen, (2004) (unpublished).
136. G. J. Fix, in "The Mathematics of Finite Elements and Applications IV" (J. R. Whiteman, Ed.), p. 265. Academic Press, London, 1982.
137. P. C. Hohenberg and B. I. Halperin, *Rev. Mod. Phys.* 49, 435 (1977).
138. J. D. Gunton and M. Droz, "Introduction to the Theory of Metastable and Unstable States." Springer, Berlin, 1983.
139. J. D. Gunton, M. San Miguel, and P. S. Sahni, in "Phase Transitions and Critical Phenomena" (C. Domb and M. S. Green, Eds.), Vol. 8, p. 267. Academic Press, London, 1983.
140. N. Provatas, Q. Wang, M. Haataja, and M. Grant, *Phys. Rev. Lett.* 91, 155502 (2003).
141. J. Müller and M. Grant, *Phys. Rev. Lett.* 82, 1736 (1999).
142. F. Drolet, K. R. Elder, M. Grant, and J. M. Kosterlitz, *Phys. Rev. E* 61, 6705 (2000).
143. M. Karttunen, M. Haataja, K. R. Elder, and M. Grant, *Phys. Rev. Lett.* 83, 3518 (1999).
144. P. W. Anderson, "Basic Notions of Condensed Matter Physics." Addison-Wesley, Redwood City, 1984.
145. J. D. Shore, D. Ronis, L. Piché, and M. Grant, *Phys. Rev. Lett.* 77, 655 (1996).

146. J. D. Shore, D. Ronis, L. Piché, and M. Grant, *Phys. Rev. E* 55, 2976 (1997).
147. F. Drolet and G. H. Fredrickson, *Phys. Rev. Lett.* 83, 4317 (1999).
148. J. G. E. M. Fraaije, B. A. C. van Vlimmeren, N. M. Maurits, M. Postma, O. A. Evers, C. Hoffmann, P. Altevogt, and G. Goldbeck-Wood, *J. Chem. Phys.* 106, 4260 (1997).
149. B. M. Ladanyi and M. S. Skaf, *Annu. Rev. Phys. Chem.* 44, 335 (1993).
150. I. Koltover, T. Salditt, J. O. Rädler, and C. R. Safinya, *Science* 281, 78 (1998).
151. S. Bandyopadhyay, M. Tarek, and M. L. Klein, *J. Phys. Chem. B* 103, 10075 (1999).
152. A. Y. Grosberg, T. T. Nguyen, and B. I. Shklovskii, *Rev. Mod. Phys.* 74, 329 (2002).
153. M. Patra, M. Patriarca, and M. Karttunen, *Phys. Rev. E* 67, 031402 (2003).
154. M. Bloom, E. Evans, and O. G. Mouritsen, *Q. Rev. Biophys.* 24, 293 (1991).
155. R. Lipowsky and E. Sackmann, Eds., "Structure and Dynamics of Membranes: From Cells to Vesicles." Elsevier, Amsterdam, 1995.
156. K. M. Merz, Jr. and B. Roux, Eds., "Biological Membranes: A Molecular Perspective from Computation and Experiment." Birkhäuser, Boston, 1996.
157. J. F. Nagle and S. Tristram-Nagle, *Biochim. Biophys. Acta.* 1469, 159 (2000).
158. J. Katsaras and T. Gutberlet, Eds., "Lipid Bilayers: Structure and Interactions." Springer, Berlin, 2001.
159. U. Essman, L. Perera, M. L. Berkowitz, H. L. T. Darden, and L. G. Pedersen, *J. Chem. Phys.* 103, 8577 (1995).
160. L. Greengard and V. Rokhlin, *J. Comput. Phys.* 73, 325 (1987).
161. R. Zangi, M. L. de Vocht, G. T. Robillard, and A. E. Mark, *Biophys. J.* 83, 112 (2002).
162. S. J. Marrink, E. Lindahl, O. Edholm, and A. E. Mark, *J. Am. Chem. Soc.* 123, 8638 (2001).
163. S. J. Marrink and D. P. Tieleman, *Biophys. J.* 83, 2386 (2002).
164. H. E. Alper and R. M. Levy, *J. Chem. Phys.* 91, 1242 (1989).
165. S. E. Feller, R. W. Pastor, A. Rojnuckarin, A. Bogusz, and B. R. Brooks, *J. Phys. Chem.* 100, 17011 (1996).
166. P. Mark and L. Nilsson, *J. Comput. Chem.* 23, 1211 (2002).
167. H. E. Alper, D. Bassolino, and T. R. Stouch, *J. Chem. Phys.* 98, 9798 (1993).
168. H. E. Alper, D. Bassolino-Klimas, and T. R. Stouch, *J. Chem. Phys.* 99, 5547 (1993).
169. P. E. Smith and B. M. Pettitt, *J. Chem. Phys.* 95, 8430 (1991).
170. H. Schreiber and O. Steinhauser, *Biochemistry* 31, 5856 (1992).
171. D. M. York, T. A. Darden, and L. G. Pedersen, *J. Chem. Phys.* 99, 8345 (1993).
172. D. M. York, W. Yang, H. Lee, T. Darden, and L. G. Pedersen, *J. Am. Chem. Soc.* 117, 5001 (1995).
173. J. Norberg and L. Nilsson, *Biophys. J.* 79, 1537 (2000).
174. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in "Intermolecular Forces" (B. Pullman, Ed.), pp. 331–342. Reidel, Dordrecht, 1981.
175. M. Patra, M. Karttunen, M. T. Hyvönen, E. Falck, P. Lindqvist, and I. Vattulainen, *Biophys. J.* 84, 3636 (2003).
176. O. Berger, O. Edholm, and F. Jahnig, *Biophys. J.* 72, 2002 (1997).
177. D. P. Tieleman and H. J. C. Berendsen, *J. Chem. Phys.* 105, 4871 (1996).
178. R. M. Venable, B. R. Brooks, and R. W. Pastor, *J. Chem. Phys.* 112, 4822 (2000).
179. C. Anézo, A. H. de Vries, H.-D. Höltje, D. P. Tieleman, and S. J. Marrink, *J. Phys. Chem. B* 107, 9424 (2003).
180. E. Lindahl and O. Edholm, *J. Chem. Phys.* 113, 3882 (2000).
181. M. Patra, M. Karttunen, M. T. Hyvönen, E. Falck, and I. Vattulainen, *J. Phys. Chem. B* (2004) (unpublished).
182. M. Patra, M. Karttunen, M. T. Hyvönen, E. Falck, and I. Vattulainen, (2004) (unpublished).
183. J. E. Gentle, "Random Number Generation and Monte Carlo Methods," 2nd edn. Springer, Berlin, 2003.
184. K. Binder and D. W. Heermann, "Monte Carlo Simulation in Statistical Physics," 4th edn. Springer, New York, 2002.
185. I. Vattulainen, *Phys. Rev. E* 59, 7200 (1999).
186. M. Mascagni, *Parallel Comput.* 24, 923 (1998).
187. K. Entacher, *ACM Trans. Model. Comput. Simul.* 8, 61 (1998).
188. A. Srinivasan, M. Mascagni, and D. Ceperley, *Parallel Comput.* 29, 69 (2003).
189. J. Krug, *Adv. Phys.* 46, 139 (1997).
190. B. Duplantier and K.-H. Kwon, *Phys. Rev. Lett.* 61, 2514 (1988).
191. K. Burdzy and G. F. Lawler, *Ann. Prob.* 18, 981 (1990).
192. H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss, *Phys. Rev. A* 45, 7128 (1992).
193. T. G. Lewis and W. H. Payne, *J. Assoc. Comput. Mach.* 20, 456 (1973).
194. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, "Numerical Recipes in Fortran 90, The Art of Parallel Scientific Computing," 2nd edn., pp. 1141–1142. Cambridge University Press, Cambridge, 1996.
195. P. L'Ecuyer, *Commun. ACM* 31, 742 (1988).
196. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical Recipes in Fortran, The Art of Scientific Computing," 2nd edn., pp. 271–273. Cambridge University Press, Cambridge, 1992.
197. F. James, *Comput. Phys. Commun.* 60, 329 (1990).
198. G. Marsaglia, A. Zaman, and W.-W. Tsang, *Stat. Prob. Lett.* 8, 35 (1990).
199. M. Matsumoto and T. Nishimura, *ACM Trans. Mod. Comp. Simul.* 8, 3 (1998).
200. M. Lüscher, *Comput. Phys. Commun.* 79, 100 (1994).
201. F. James, *Comput. Phys. Commun.* 79, 111 (1994).
202. I. Vattulainen, T. Ala-Nissila, and K. Kankaala, *Phys. Rev. Lett.* 73, 2513 (1994).
203. I. Vattulainen, T. Ala-Nissila, and K. Kankaala, *Phys. Rev. E* 52, 3205 (1995).
204. I. Vattulainen, K. Kankaala, J. Saarinen, and T. Ala-Nissila, *Comput. Phys. Commun.* 86, 209 (1995).
205. L. N. Shchur and P. Butera, *Int. J. Mod. Phys. C* 9, 607 (1998).

206. N. H. March and M. P. Tosi, "Atomic Dynamics in Liquids." Dover, New York, 1976.
207. A. M. Ferrenberg, D. P. Landau, and Y. J. Wong, *Phys. Rev. Lett.* 69, 3382 (1992).
208. R. M. D'Souza, Y. Bar-Yam, and M. Kardar, *Phys. Rev. E* 57, 5044 (1998).
209. J. M. Hammersley and D. C. Handscomb, "Monte Carlo Methods." Methuen & Co., London, 1964.
210. G. Parisi and F. Rapuano, *Phys. Lett.* 157B, 301 (1985).
211. F. J. Resende and B. V. Costa, *Phys. Rev. E* 58, 5183 (1998).
212. L. Verlet, *Phys. Rev.* 159, 98 (1967).
213. I. Pagonabarraga, M. H. J. Hagen, and D. Frenkel, *Europhys. Lett.* 42, 377 (1998).
214. W. K. den Otter and J. H. R. Clarke, *Europhys. Lett.* 53, 426 (2001).
215. T. Shardlow, *SIAM J. Sci. Comp.* 24, 1267 (2003).
216. J. P. Boon and S. Yip, "Molecular Hydrodynamics." Dover, New York, 1980.
217. C. J. Cotter and S. Reich, *Europhys. Lett.* 64, 723 (2003).
218. P. Español, in "Novel Methods in Soft Matter Simulations" (M. Karttunen, I. Vattulainen, and A. Lukkarinen, Eds.). Springer, Berlin, 2004 (unpublished).
219. A. P. Lyubartsev, M. Karttunen, I. Vattulainen, and A. Laaksonen, *Soft Materials* 1, 121 (2003).
220. D. Voet and J. G. Voet, "Biochemistry," 2nd edn. Wiley, New York, 1995.
221. G. Cevc and D. Marsh, "Phospholipid Bilayers: Physical Principles and Models." John Wiley & Sons, New York, 1987.
222. R. B. Gennis, "Biomembranes: Molecular Structure and Function." Springer, New York, 1989.
223. M. Nielsen, L. Miao, J. H. Ipsen, O. G. Mouritsen, and M. J. Zuckermann, *Phys. Rev. E* 54, 6889 (1996).
224. M. Nielsen, L. Miao, J. H. Ipsen, M. J. Zuckermann, and O. G. Mouritsen, *Phys. Rev. E* 59, 5790 (1999).
225. M. Nielsen, L. M. J. H. Ipsen, O. G. Mouritsen, M. J. Zuckermann, J. Thewalt, and M. Bloom, *Europhys. Lett.* 52, 368 (2000).
226. J. M. Polson, I. Vattulainen, H. Zhu, and M. J. Zuckermann, *Eur. Phys. J. E* 5, 485 (2001).
227. L. Miao, M. Nielsen, J. Thewalt, J. H. Ipsen, M. Bloom, M. J. Zuckermann, and O. G. Mouritsen, *Biophys. J.* 82, 1429 (2002).
228. M. J. Zuckermann, M. Bloom, J. H. Ipsen, L. Miao, O. G. Mouritsen, M. Nielsen, J. Polson, J. Thewalt, I. Vattulainen, and H. Zhu, *Methods Enzym.* (2004) (unpublished).
229. P. van der Ploeg and H. J. C. Berendsen, *J. Chem. Phys.* 76, 3271 (1982).
230. R. M. Venable, Y. Zhang, B. J. Hardy, and R. W. Pastor, *Science* 262, 223 (1993).
231. K. Simons and E. Ikonen, *Nature* 387, 569 (1997).
232. D. A. Brown and E. London, *Annu. Rev. Cell. Dev. Biol.* 14, 111 (1998).
233. M. Edidin, *Annu. Rev. Biophys. Biomol. Struct.* 32, 257 (2003).
234. J. R. Silvius, *Biochim. Biophys. Acta* 1610, 174 (2003).
235. M. Dykstra, A. Cherukuri, H. W. Sohn, S.-J. Tzeng, and S. K. Pierce, *Annu. Rev. Immunol.* 21, 457 (2003).
236. E. Falck, M. Patra, M. Karttunen, M. T. Hyvönen, and I. Vattulainen (2004) (unpublished).
237. E. Falck, M. Patra, M. Karttunen, M. T. Hyvönen, and I. Vattulainen, (2004) (unpublished).
238. A. M. Smondyrev and M. L. Berkowitz, *Biophys. J.* 80, 1649 (2001).
239. S. W. Chiu, E. Jakobsson, and H. L. Scott, *J. Chem. Phys.* 114, 5435 (2001).
240. S. W. Chiu, E. Jakobsson, R. J. Mashl, and H. L. Scott, *Biophys. J.* 83, 1842 (2002).
241. M. Pasenkiewicz-Gierula, T. Róg, K. Kitamura, and A. Kusumi, *Biophys. J.* 78, 1376 (2000).
242. T. Róg and M. Pasenkiewicz-Gierula, *Biophys. J.* 81, 2190 (2001).
243. C. Hofsäß, E. Lindahl, and O. Edholm, *Biophys. J.* 84, 2192 (2003).
244. P. Jedlovszky and M. Mezei, *J. Phys. Chem. B* 107, 5311 (2003).
245. P. Jedlovszky and M. Mezei, *J. Phys. Chem. B* 107, 5322 (2003).
246. P. Jedlovszky and M. Mezei, *J. Phys. Chem. B* 108, 465 (2004).
247. J. A. Söderhäll and A. Laaksonen, *J. Phys. Chem. B* 105, 9308 (2001).
248. M. T. Hyvönen and P. T. Kovanen, *J. Phys. Chem. B* 107, 9102 (2003).
249. E. Mombelli, R. Morris, W. Taylor, and F. Fraternali, *Biophys. J.* 84, 1507 (2003).
250. S. W. Chiu, S. Vasudevan, E. Jakobsson, R. J. Mashl, and H. L. Scott, *Biophys. J.* 85, 3624 (2003).
251. P. Niemelä, M. T. Hyvönen, and I. Vattulainen (unpublished).
252. R. G. Snyder, K. Tu, M. L. Klein, R. Mendelssohn, H. L. Strauss, and W. Sun, *J. Phys. Chem. B* 106, 6273 (2002).
253. T. Huber, K. Rajamoorthi, V. F. Kurze, K. Beyer, and M. F. Brown, *J. Am. Chem. Soc.* 124, 298 (2002).
254. L. Saiz and M. L. Klein, *Biophys. J.* 81, 204 (2001).
255. S. E. Feller, K. Gawrisch, and A. D. MacKerell, Jr., *J. Am. Chem. Soc.* 124, 318 (2002).
256. H. Möhwald, in "Structure and Dynamics of Membranes: From Cells to Vesicles" (R. Lipowsky and E. Sackmann, Eds.), pp. 161–211. Elsevier, Amsterdam, 1995.
257. R. Maget-Dana, *Biochim. Biophys. Acta* 1462, 109 (1999).
258. Y. N. Kaznessis, S. Kim, and R. G. Larson, *Biophys. J.* 82, 1731 (2002).
259. R. Boeckmann, A. Hac, T. Heimburg, and H. Grubmuller, *Biophys. J.* 85, 1647 (2003).
260. S. A. Pandit, D. Bostick, and M. L. Berkowitz, *Biophys. J.* 84, 3743 (2003).
261. S. A. Pandit and M. L. Berkowitz, *Biophys. J.* 82, 1818 (2002).
262. A. Gurtovenko, M. Patra, M. Karttunen, and I. Vattulainen, *Biophys. J.* (2004) (unpublished).
263. A. K. Sum and J. J. de Pablo, *Biophys. J.* 85, 3636 (2003).
264. A. K. Sum, R. Faller, and J. J. de Pablo, *Biophys. J.* 85, 2830 (2003).

265. M. Pasenkiewicz-Gierula, T. Róg, J. Grochowski, P. Serda, R. Czarnecki, T. Librowski, and S. Lochyński, *Biophys. J.* 85, 1248 (2003).
266. J. Ulander and A. D. J. Haymet, *Biophys. J.* 85, 3475 (2003).
267. P. Mukhopadhyay, H. J. Vogel, and D. P. Tieleman, *Biophys. J.* 86, 337 (2004).
268. M. T. Hyvönen, K. Öörni, P. T. Kovanen, and M. Ala-Korpela, *Biophys. J.* 80, 565 (2001).
269. M. Ø. Jensen, T. R. Jensen, K. Kjaer, T. Bjørnholm, O. G. Mouritsen, and G. H. Peters, *Biophys. J.* 83, 98 (2001).
270. I. H. Shrivastava and M. S. P. Sansom, *Biophys. J.* 78, 557 (2000).
271. C. E. Capener and M. S. P. Sansom, *J. Phys. Chem. B* 106, 4543 (2002).
272. J. Baudry, E. Tajkhorshid, F. Molnar, J. Phillips, and K. Schulten, *J. Phys. Chem. B* 105, 905 (2001).
273. S. J. Marrink, A. H. de Vries, and A. E. Mark, *J. Phys. Chem. B* 108, 750 (2004).
274. R. Goetz and R. Lipowsky, *J. Chem. Phys.* 108, 7397 (1998).
275. R. Goetz, G. Gompper, and R. Lipowsky, *Phys. Rev. Lett.* 82, 221 (1999).
276. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, *J. Phys. Chem. B* 105, 4464 (2001).
277. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, *J. Phys. Chem. B* 105, 9785 (2001).
278. S. O. Nielsen, C. F. Lopez, P. B. Moore, J. C. Shelley, and M. L. Klein, *J. Phys. Chem. B* 107, 13911 (2003).
279. G. Ayton and G. A. Voth, *Biophys. J.* 83, 3357 (2002).
280. G. Ayton, S. G. Bardenhagen, P. McMurtry, D. Sulsky, and G. A. Voth, *IBM J. Res. Dev.* 45, 417 (2001).
281. G. Ayton, S. G. Bardenhagen, P. McMurtry, D. Sulsky, and G. A. Voth, *J. Chem. Phys.* 114, 6913 (2001).
282. G G. Ayton, A. M. Smondyrev, S. G. Bardenhagen, P. McMurtry, and G. A. Voth, *Biophys. J.* 82, 1226 (2002).
283. G. Ayton, A. M. Smondyrev, S. G. Bardenhagen, P. McMurtry, and G. A. Voth, *Biophys. J.* 83, 1026 (2002).
284. K. Toukan and A. Rahman. *Phys. Rev. B* 31, 2643 (1985).
285. D. E. Smith and L. X. Dang, *J. Chem. Phys.* 100, 3757 (1994).
286. A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* 55, 5689 (1997).
287. P. Mark and L. Nilsson, *J. Phys. Chem. A* 105, 9954 (2001).
288. M. Patra and M. Karttunen, *J. Comput. Chem.* 25, 678 (2004).
289. D. F. Evans and H. Wennerström, "The Colloidal Domain: Where Physics, Chemistry, Biology, and Technology Meet," 2nd edn. Wiley, New York, 1999.
290. A. J. Masters and P. B. Warren, *Europhys. Lett.* 48, 1 (1997).
291. M. Canales and G. Sesé, *J. Chem. Phys.* 109, 6004 (1998).
292. D. R. Lide, Ed., "CRC Handbook of Chemistry and Physics," 82nd edn. CRC Press, Boca Raton, 2001.