Computational Modeling of Realistic Cell Membranes

Siewert J. Marrink,*† Valentina Corradi,+ Paulo C.T. Souza,† Helgi I. Ingolfsson,§ D. Peter Tieleman,+ and Mark S.P. Sansom∥

†Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands
‡Centre for Molecular Simulation and Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
§Biosciences and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, United States
∥Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

ABSTRACT: Cell membranes contain a large variety of lipid types and are crowded with proteins, endowing them with the plasticity needed to fulfill their key roles in cell functioning. The compositional complexity of cellular membranes gives rise to a heterogeneous lateral organization, which is still poorly understood. Computational models, in particular molecular dynamics simulations and related techniques, have provided important insight into the organizational principles of cell membranes over the past decades. Now, we are witnessing a transition from simulations of simpler membrane models to multicomponent systems, culminating in realistic models of an increasing variety of cell types and organelles. Here, we review the state of the art in the field of realistic membrane simulations and discuss the current limitations and challenges ahead.

CONTENTS
1. Introduction 6185
2. Computational Tools 6186
  2.1. All-Atom Models 6187
    2.1.1. Challenge of Atomistic Force Fields 6187
    2.1.2. CHARMM 6187
    2.1.3. AMBER 6187
    2.1.4. Slipids 6188
    2.1.5. GROMOS 6188
    2.1.6. Polarizable Models 6188
    2.1.7. Limitations/Developments of AA Models 6188
    2.1.8. Setup Tools 6189
  2.2. CG Models 6189
    2.2.1. Top Down versus Bottom Up 6190
    2.2.2. Martini Model 6190
    2.2.3. SDK Model 6190
    2.2.4. ELBA Model 6191
    2.2.5. SIRAH Force Field 6191
    2.2.6. Solvent-Free Models 6191
    2.2.7. Limitations/Developments of CG Models 6191
    2.2.8. High-Throughput Tools 6192
  2.3. Supra-CG Models 6192
    2.3.1. Supra CGing Approaches 6192
    2.3.2. Few-Bead Lipids 6193
    2.3.3. Reduced Protein Models 6193
    2.3.4. Meso Models 6193
3. Increasing Complexity 6193
  3.1. Multicomponent Membranes 6194
    3.1.1. Lipid Domains 6194
    3.1.2. Protein−Lipid Binding Sites 6195
    3.1.3. Lipid-Mediated Protein Oligomerization 6197
    3.1.4. Membrane Curvature Generation and Sensing 6199
  3.2. Realistic Cell Membranes 6201
    3.2.1. Plasma Membranes 6201
    3.2.2. Organelle Membranes 6203
    3.2.3. Bacterial Membranes 6203
    3.2.4. Skin Models 6204
    3.2.5. Complications of Complexity 6204
  3.3. Toward Full Cell Models 6205
    3.3.1. Viral Envelopes 6205
    3.3.2. Large-Scale Membrane Organization 6205
    3.3.3. Membrane Remodeling 6206
    3.3.4. In Silico in Vivo 6207
4. Outlook 6207
Author Information 6207
Acknowledgments 6208

Special Issue: Biomembrane Structure, Dynamics, and Reactions

Received: July 23, 2018
Published: January 9, 2019

© 2019 American Chemical Society

DOI: 10.1021/acs.chemrev.8b00460
Chem. Rev. 2019, 119, 6184–6226

This is an open access article published under a Creative Commons Non-Commercial No Derivative Works (CC-BY-NC-ND) Attribution license, which permits copying and redistribution of the article, and creation of adaptations, all for non-commercial purposes.
1. INTRODUCTION

Membranes are essential components of every cell, providing the cell’s identity as well as defining a large variety of internal compartments. Typical cell membranes may contain hundreds of different lipids, asymmetrically distributed between the two bilayer leaflets and are crowded with proteins covering an estimated membrane area as large as 30%. The compositional heterogeneity of cellular membranes is now well recognized, leading to a nonuniform lateral distribution of the components. Together, lipids and proteins form distinct nanodomains with important implications for many cellular processes such as membrane fusion, protein trafficking, and signal transduction. Lipids move proteins, and proteins move lipids in a fascinating protein—lipid interplay.

Experimental techniques are getting more and more sophisticated to reveal lateral membrane organization and the principles driving it. Experimental advances include improved methods for single-particle tracking, fluorescence correlation spectroscopy, super-resolved imaging, scattering, solid-state NMR, and mass spectrometry, as well as methods to prepare asymmetric model membranes and real cell membrane extracts. However, the detailed membrane organization proves difficult to probe at the molecular level, despite progress in experimental techniques that can directly probe living cells. Computer simulations, in principle, can provide this detail. Techniques such as molecular dynamics (MD) are capable of describing the interactions between all the components in the system at atomic resolution, acting like a "computational microscope." Given enough computer power, the behavior of a system can be followed in time long enough to observe the process of interest.

The first MD simulations of surfactants and lipids appeared in the 1980s, shortly after the first published protein simulations at a time when there were only a handful of super computers available for academic research. Complexity in lipid and surfactant systems rapidly increased from simplified ordered decanoate bilayers tethered harmonically to the average position of all headgroup particles to a smectic liquid crystal made of decanol, decanoate, water, and sodium ions, a micelle, and a liquid crystalline DPPC bilayer. In the early 1990s several groups published simulation papers on phospholipids with explicit water, including the infamous Berger lipid model that, although parametrized on erroneous data, became one of the leading lipid force fields until quite recently. These early papers already targeted a set of diverse problems, including lipid bilayer structure, transport of small molecules through bilayers, effect of cholesterol, the hydration force between bilayers, and interactions with membrane-active peptides, all of which continue to be studied. The first simulations of complete membrane proteins in a lipid environment studied gramicidin A, bacteriorhodopsin, OmpF porin, and phospholipase A. An early example of protein-induced bilayer perturbation is found in the work of Tieleman et al. Simulations of membrane proteins have since grown immensely in importance and are now widely used. Comprehensive reviews of these pioneering studies are available in the literature.

As computer power grew and became more universally available, lively technical discussions appeared in the literature. Significant matters of debate included the use of cutoffs, appropriate boundary conditions for membrane simulations, as well as concerns with sampling and questions related to linking experiment and simulation. The latter two are not specific to membrane systems and, not surprisingly, continue to be major topics of both concern and continued research. In addition, during the first decade of the new millennium, we witnessed a growing range of applications of simulations involving collective lipid motion. Key pioneering examples include accessing bilayer undulatory modes, spontaneous self-assembly of lipids into a bilayer, pore formation by antimicrobial peptides or electrical fields lipid flip-flop, collective lipid flows, domain formation, membrane fusion, and many more. For an in-depth discussion on these developments, now more than 10 years ago, we refer the reader to a number of earlier reviews.

If we express the scope of a simulation as a combination of system size and simulation length, there has always been a large (maybe even up to 2−3 orders of magnitude) difference between a "typical" simulation and the largest ones in the literature. A typical scope in the early 1990s would be a bilayer model of 72−128 lipids (or 4000−15000 atoms) and simulation times of the order of a hundred picoseconds. For comparison, at the moment, early 2018, a typical simulation study might involve a combination of dozens of simulations on the order of microseconds, where a simulation system might contain 150000 atoms, an increase of at least 5 orders of magnitude. At these time and length scales, many interesting biochemical and biophysical questions can be addressed by simulations on relatively commonly available computer resources. Leadership-category machines allow access to 2−3 orders of magnitude more elaborate studies and coarse-grained models describe similar systems at a computational cost that is 2−3 orders of magnitude lower than a corresponding atomistic model. This massive increase in accessible scope, which now includes a large number of applications, has led to an explosive growth in the use of simulations to study membranes, as well as to the use of simulations in general.

Thanks to the ongoing increase in computer power, sparked by the efficient use of GPUs, together with the development of accurate atomistic and coarse-grain (CG) models and the community-based development of tools to automate setup and analysis of membrane simulations, we are now witnessing a transition from simulations of simplified, model membranes toward multicomponent realistic membranes. This transition is essential to unravel protein−lipid interplay in the crowded and complex environment of real cell membranes, where experimental detection is difficult and theoretical models fall short. In this review, we focus on this transition, which is becoming apparent during the past five years (Figure 1). We restrict ourselves to particle-based simulation methods, mostly MD, and to simulation studies addressing the lateral and spatial organizational principles of membranes. For a discussion of related topics, not covered in the current review, we refer the reader to a number of other recent reviews, for example, on membrane proteins functioning and activity, binding of membrane active peptides, nanoparticle uptake, drug-membrane interactions, ionic-liquids and membranes, pore formation, lipid flip-flop, and lipid nanodisks.

The rest of this review is organized as follows. We first give an overview of the tools comprising the computational microscope, organized by the level of resolution obtained: from all-atom models via CG models to supra-CG models.
Then we provide a comprehensive overview of the current state of the art in modeling membrane systems of increasing complexity, with sections on multicomponent systems, realistic cell membranes, and the current avenues toward full cell models. A short outlook section concludes this review.

2. COMPUTATIONAL TOOLS

At the heart of the computational “microscope” lies the simulation algorithm, for which MD is most widely used. MD simulations, in their most basic form, involve numerically solving classical equations of motion for a set of particles over a given time period. The resulting time series, called trajectory, can subsequently be visualized and analyzed in detail. MD simulation algorithms, as well as related algorithms such as Brownian Dynamics, Langevin Dynamics, and Dissipative Particle Dynamics (DPD) have been implemented in a number of simulation software packages; the most widely used in the field of membrane modeling include AMBER, CHARMM, NAMD, OpenMM, LAMMPS, ESPResSo, and GROMACS, as well as the special purpose machine ANTON with the Desmond software. A major limitation of simulations is the limited amount of sampling that can be performed, even when using the largest super computers available today. To more efficiently explore phase space, various enhanced sampling and biasing methods are available, with replica exchange MD (REMD), metadynamics, milestoning, and umbrella sampling (US) among the most popular methods in the field of biomembranes. Noteworthy are recent attempts to adopt these methods specifically in the field of membrane simulations.

Central to the success of an MD simulation is the quality of the force field (FF) (i.e., the set of parameters dictating how the particles interact). In biomolecular simulation in general, there is a variety of FFs, although they fall in a handful of families that continue to be developed and are broadly similar in terms of their potential function and main approximations.

An important distinction between the FFs is the level of resolution considered (Figure 2). Traditionally, full atomistic detail is the highest level of resolution for classical MD simulations (i.e., when quantum degrees of freedom or electronic polarizability are not considered explicitly). However, to increase the spatiotemporal range of simulations, lower resolution level FFs have been developed. These range from CG models that still contain chemical detail to supra-CG models, in which lipids and proteins are represented only qualitatively by few-bead models, and solvent is considered implicitly. Further reduction in resolution is achieved by integrating out also the lipid particles by mean-field approaches.

Figure 1. Growth of complexity of membrane models. From the pioneering stage 30 years ago, basic properties of one and two component membranes were explored around the millennium. From then on, complexity of simulated membrane systems was gradually increased, culminating in the current era of more and more realistic membrane models. POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DPPC, 1,2-dipalmitoyl-sn-glycero-3-phosphocholine; POPE, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine; DOPC, 1,2-dioleoyl-sn-glycero-3-phosphocholine; Chol, cholesterol; CLs, cardiolipins; PPPE, 1-palmitoyl-2-palmitoleoyl-phosphatidylethanolamine; PVPG, 1-palmitoyl-2-vacenoyl-phosphatidylglycerol; PVCL2, 1,10-palmitoyl-2,20-vacenoyl cardiolipin; LpsS, E. coli R1 lipopolysaccharide core with repeating units of O6-antigen. From left to right: Reprinted with permission from ref 20. Copyright 1988 AIP Publishing. Adapted from ref 26. Copyright 1993 American Chemical Society. Adapted from ref 42. Copyright 2001 American Chemical Society. Adapted with permission from ref 38. Copyright 2004 American Society for Biochemistry and Molecular Biology. Adapted from ref 311. Copyright 2014 American Chemical Society. Adapted from ref 382. Copyright 2013 American Chemical Society. Adapted from ref 593. Copyright 2014 American Chemical Society. Adapted with permission from ref 643. Copyright 2016 Elsevier.

Figure 2. Different resolutions in particle-based simulation models of lipid membranes. At the all-atom (AA) level, all atoms are considered explicitly. Upon coarse-graining, small groups of atoms and associated hydrogens are represented by coarse-grain (CG) beads. Moving down in resolution to the supra-CG level, lipids and proteins are represented only qualitatively by few-bead models, and solvent is considered implicitly. Further reduction in resolution is achieved by integrating out also the lipid particles by mean-field approaches.
models that are more generic in nature and can form a bridge to the continuum level of description. Below, we discuss the current state of the FFs in each of these categories in detail, restricting ourselves to the most popular FFs in lipid membrane simulations.

### 2.1. All-Atom Models

Generally speaking, detailed atomistic lipid parameters have been developed with the same philosophy as protein FFs and in practice in most cases are related to or part of a small number of widely used more general FFs. Although there are many FFs for lipids, and many modifications have been proposed for specific cases, there is only a few of FFs that aims to be general enough for complex membrane simulations. In the current literature, these can be divided in four families that are still being developed: CHARMM, AMBER, Lipids, and GROMOS. Given the staggering variety of lipid types, developing and testing consistent parameter sets poses significant challenges. Below we describe some of these challenges, followed by a brief description of the most widely used atomistic FFs, setup tools to build complex membrane models, and limitations of atomistic simulations. For an in-depth discussion and comparison of current atomistic FFs, see, for instance, refs 94−97.

#### 2.1.1. Challenge of Atomistic Force Fields

First, the properties of lipid bilayers are determined by the sum of a large number of interactions, some of which are weak but add up to significant contributions. An example is the strong effect of pressure on the structure of lipid bilayers, but pressure has significant contributions from long-range Lennard-Jones interactions. This makes lipid simulations quite sensitive to small variations in parameters, in particular standard schemes used to mitigate cutoff errors routinely used in molecular dynamics simulations and the related treatment of electrostatic interactions.

Second, it has only recently become practical to routinely carry out simulations on a time scale of hundreds of nanoseconds, which is required to get equilibrated properties on a bilayer of ca. 250 lipids of one type of lipid. Thus, any change in parameters requires a large amount of computer time to investigate. For binary mixtures in liquid crystalline phases or their cholesterol-containing analogues (liquid disordered), equilibration times increase to microseconds and much more in the presence of ordered domains. A related problem is that periodic boundary conditions effect the properties of lipids in simulations. Some of the first simulations of bilayers used 32−100 lipids per leaflet, but this amounts to 5−10 lipids in each of the x and y dimension and an artificially constrained length scale compared to the characteristic length scale of lipid interactions in experimental systems.

Third, biological membranes contain a large number of lipid components, which are made of a combination of a limited number of different head groups, linkages, and a limited number of different tails. In principle these components should be transferable in FFs, but this requires an additional, large, amount of testing. For mixtures, the number of possible combinations explode. In practice, these components are not reliably transferable and might be considered a reasonable initial model.

Fourth, detailed experimental structural data, primarily from neutron and X-ray scattering and from NMR, have been available for a growing number of lipids, starting with phosphatidylcholine (PC) lipids, but is insufficient to validate models of all biologically interesting lipids. Force field development and detailed experiments these days often go hand-in-hand, as simulations augment the interpretation of experimental results and in some cases drive experiments to parametrize new lipids and more complex systems. Recent reviews on comparing atomistic simulations and experiments include refs 98 and 99. In simulations, PC lipids have generally been the easiest to model, but the resulting parameters have not reliably transferred to other lipid types. More recently, a wider range of model lipids has been studied experimentally, primarily by scattering, including phosphatidylserine (PS), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and phosphatidylcholine lipids (PC) lipids,100,101 the structure of polyunsaturated lipids,102 and elements of cholesterol.103 These studies provide essential detail for the validation of simulations, but still only span a small subset of all lipids, and have been subject to several reinterpretations, while key elements like sphingomyelins have received less attention. They have also been largely limited to single-component systems, whereas more detailed experimental structural data on mixtures would be very useful for the development of simulation parameters.

Next to scattering, a second major experimental technique is deuterium NMR, which measures the average orientation of C−D bonds in deuterated lipids and can measure dynamics on relevant simulation time scales.104 Since bond orientations and detailed dynamics can be directly calculated from simulations, they are powerful validation tools.105 By selectively labeling one component in lipid mixtures, details on mixtures can also be obtained. A second major application of deuterium NMR has been the measurement of phase diagrams for simple mixtures.106 Since deuterium, unlike fluorescent probes, barely changes the chemistry of lipids, this is very important data. It remains challenging to calculate phase diagrams for computer models, but this has become feasible for CG simulations (see below) and will soon be more feasible for atomistic simulations.

#### 2.1.2. CHARMM

The most elaborate effort has been put in CHARMM36, an updated lipid FF consistent with the most recent version of the more general CHARMM FF for biomolecular simulation, which includes protein, nucleic acid, and small molecule parameters.107−109 This work is based on extensive parametrization for tails, headgroup components, and specific lipids, and has additional advantages in the large set of parametrized and tested lipids as well as the powerful setup tool CHARMM-GUI (see below).110,111 The CHARMM lipid FF was initially developed for PC lipids but has been massively extended. It includes most common lipids used in biophysical experiments, the main families of lipids found in higher organisms, bacterial lipids specific to extremophiles including ring-containing and branched lipids and hopanoids, a library of LPS from the outer membrane of Gram-negative bacteria, and yeast lipids including sterols. The main repository for lipid parameters is CHARMM-GUI, as no comprehensive review or paper describing the current CHARMM lipidome is available, although individual components have been described in more detail.112−115 The present issue has a detailed review by Leonard et al. with a comprehensive description as of 2018.116 CHARMM lipid parameters are typically used with the CHARMM protein FF, which is implemented in most of the widely used MD programs.

#### 2.1.3. AMBER

AMBER is a widely used FF for proteins, nucleic acids, and small (druglike) molecules, similar to
CHARMM. Several groups have attempted to develop AMBER lipid parameters for use with the rest of the AMBER FF, initially based on GAFF, the generalized AMBER FF. This was tested on a limited set of lipids and has not been widely used. The most recent published AMBER-based parameters set is Lipid14. Lipid14 appeared in 2014 and has not been widely used yet either. It initially had parameters for six different PC lipids with either saturated or monounsaturated chains. Lipid14 also has updated cholesterol parameters. A Lipid17 version with an expanded library is under development and available for testing at the time of writing of this review but has not been formally published yet. Compatible parameters for LPS are also available for AMBER. A major advantage of AMBER parameters for simulating complex membranes is the advanced state of the rest of the FF, but a significant amount of development is required to make the FF easily applicable to a variety of lipids and lipid mixtures.

2.1.4. Slipids. Another promising set of FF parameters has been developed by Jämbeck et al., called Slipids (for Stockholm Lipids). These have been parametrized to be consistent with AMBER, although this consistency is primarily based on the same charge derivation method as AMBER uses, and the standard for Lennard-Jones parameters is derived from CHARMM. The initial paper described DLPC, DMPC, and DPPC, which has been expanded to include monounsaturated PC and PE lipids, as well as sphingomyelin, PG, PS, and cholesterol, and most recently a set of polyunsaturated PC lipids. The protocol for parametrization is sufficiently well-defined that there is a clear path for adding new lipids. This set has not been used as widely as the CHARMM lipids and is still relatively new but so far appears a viable choice that has been used both with AMBER and CHARMM protein FFs. A recent paper derived parameters for a large set of steroids to be consistent with Slipids, which are currently not available for other force fields.

2.1.5. GROMOS. The GROMOS parameter set is based on the united-atom FF GROMOS 54A7. Mark and colleagues developed parameters for a number of lipid types that are consistent with GROMOS 54A7. Computationally, these have an advantage because in most software implementation united-atom lipids are substantially more efficient than all-atom lipids, in contrast to protein FFs where the extra hydrogens have much less impact. As for other FFs, the first lipids to be parametrized were saturated and monounsaturated PC lipids. In addition, parameters for bacterial lipids with branched fatty acids in their lipid chains, with cyclo-propane moieties, LPS, and for hopanoids and sterols are available. The parametrization is consistent in approach and atom types with GROMOS 54A7, which enables lipid–protein simulations, but the number of different lipids that is available and has been tested for this FF is rather limited.

2.1.6. Polarizable Models. Although the further improvement of standard atomistic FFs has arguably been the most important recent development, together with increased time scales accessible with newer computers and GPUs, in the slightly longer term recent work on polarizable lipid FFs may become very important. In standard atomistic FFs, we assume that the details of electronic motion are averaged out. The main consequence of this is that the partial charge of atoms cannot respond to the environment, although this is an important effect in some cases. Classical FFs that address this are called polarizable or nonadditive FFs, essentially with charges that will respond to the environment. Such FFs were routinely forecast as the next step even more than 30 years ago, but in practice their cost and the effort required to develop consistent FFs has made progress slow. In the past few years, two different approaches have been applied to membrane simulations, while a third, more detailed and expensive method has been used in other biomolecular systems but not yet on membranes to our knowledge. In the Drude oscillator model, small charges on springs attached to the nucleus (the standard atomic atom) are able to move around in response to the local electric field, thus changing the charge distribution. In the FlexQ method, charges equilibrate locally. Both methods have been applied to model systems, including PC lipids, peptides, and nucleic acids. Simulations of mixed polarizable/standard systems have also been used, as in principle the most polarizable atoms could be treated as polarizable. Examples are systems with the lipid chains as polarizable or simulations with a permeating molecule as polarizable. A third model, AMOEBA, is considerably more complicated but is now used in biomolecular simulation and would be interesting to test in membranes.

At the current state of the art, it is clear that there are viable polarizable models for membranes. They have been tested on relatively limited cases so far, primarily PC lipids. Probably the most striking difference between standard atomistic and polarizable models is a large difference in the dipole potential across the water/lipid interface. Unfortunately, this property is not easy to measure or interpret. Other properties appear less critical, and it remains to be seen in more detail where the strengths and weaknesses of these more complicated models lie.

2.1.7. Limitations/Developments of AA Models. Lipid FFs do not divide readily into neat categories, but broadly speaking, there are recognizable families in addition to a large number of more ad-hoc modifications with generally more limited reach. Such modifications allow optimizations for a specific purpose, but in the context of complex membranes, they do not generalize sufficiently to be useful. For complex membranes, a consistent set of lipid parameters, including all relevant types for the problem at hand, which may include sterols or unusual bacterial, mitochondrial, or endosomal lipids, and a consistent set of protein parameters is essential. We argue that this requirement is currently not met by any set of parameters, although CHARMM comes closest.

An additional complexity is the reliance of all FFs on very specific cutoff values for Lennard-Jones interactions and corresponding shift functions to deal with cutoff artifacts. One consequence of this is that it is not trivial to exactly match the results of simulations with the CHARMM FF in NAMD, AMBER, or GROMACS when attempting to match the original parametrization conditions in the CHARMM simulation software. Anecdotal results have been dramatically different as lipids undergo phase transitions to the gel phase at the wrong temperatures, although recent updates to simulation algorithms in different software packages offer significant improvements, tested in, for example, ref 146. One thorough solution for this would be to reparametrize entire FFs to not use cut-offs at all, which has become more realistic in recent years with the development of efficient lattice sum methods. Unfortunately, it is hard to see where the resources for the effort would come from to reparametrize the most widely used, and most complex, FFs. This is an effort that would have a
wide impact on the field, making lipid force fields more transferrable and, therefore, ought to be funded. An interesting initiative uses a form of crowdsourcing to collect validation data on a variety of lipids in an open science format. The project identified a number of issues with the headgroups and glycerol backbones of PC lipids and provides an important database of simulation data.147 A more technical consideration is that changes in algorithms, often coupled to changes in computer hardware that favor one type of optimization over another, do affect simulation results.148 This will continue to be a concern and require simple test systems for regression testing as actual research systems become increasingly complex.

In addition, there are intrinsic limitations in the use of finite systems with periodic boundaries. This has been documented for the calculation of electrostatic properties, but more recently it was shown by Camley et al. that the diffusion coefficients of membrane-embedded objects have a nontrivial dependence on both the box shape and box size, and in particular show a strong dependence on the normal direction to the membrane.149 This is perhaps counterintuitive, but the water layer surrounding the membrane couples hydrodynamically to the membrane and diffusion coefficients do not converge with increasing size of the membrane patch. Subsequent large-scale simulations confirmed this behavior, and analytical expressions to correct for these artifacts have recently been introduced.150−152 Such considerations become increasingly important as simulations model larger and increasingly complex systems and begin to overlap with direct measurements of diffusion of membrane proteins by spectroscopic methods.

One additional use of deuterium NMR that could be expanded is the measurement of order parameters of a “reporter” lipid like DMPC or POPC, which are readily available in deuterated form, as a function of concentration in mixtures. More generally, deuterium NMR has not been widely applied to mixtures, except for investigations involving cholesterol, and it is challenging to obtain funding for this, but this would be important data to validate simulations of lipid mixtures.

In addition to lipids, sterols play an important biological role and require careful parametrization. Lipid−protein interactions introduce additional complexities. A lack of useful experimental data to validate simulations is a limiting factor in model improvement in many cases. Finally, improved parameters for ions, in particular their tendency to adsorb to the membrane/water interface, remains an ongoing and important area of research.153−156

2.1.8. Setup Tools. Historically, great effort was spent on creating starting structures for simulations that were as close as possible to equilibrium, because limited simulation time scales (nanoseconds) compared to phospholipid diffusion and other motions (tens of nanoseconds or more) meant that poor starting structures completely biased the simulation results.157−161 As computers became faster, starting structures for relatively simple systems became less problematic, as even starting from random mixtures in solution resulted in equilibrated bilayers.42,162 However, for complex membranes of the type described here, or even basic mixtures or membrane proteins in basic mixtures, we are now in a situation again that it takes microseconds or much longer to equilibrate starting structures, a key prerequisite for useful simulations. A second problem is that finding errors in initial structures is almost impossible in very large simulations, which puts stringent demands on useful setup methods. This will continue to be an area of development for the foreseeable future. Here we will discuss some widely used tools.

Perhaps the most widely used tool is CHARMM-GUI, a graphical interface developed by Im and co-workers to set up a broad range of biomolecular simulations, for most of the major molecular dynamics packages. One of its uses is the conversion of CHARMM FFs to input formats that can be used in GROMACS, NAMD, OpenMM, and other software.163 For membranes, it can build structures based on a desired composition using an extensive library of lipids, including bacterial lipids, a large library of lipopolysaccharides for outer membranes from Gram-negative bacteria, and a library of yeast-specific lipids. One major problem with these systems is the slow equilibration time. A related tool has recently been developed by de Fabritis and co-workers, coined HTMD (High Throughput MD).164 HTMD offers a platform for preparation of MD simulations in general, including membrane/protein systems. Starting from PDB structures, the platform assists in building the system using well-known force fields, and in applying standardized protocols for running the simulations.

Two other methods try to use simpler model descriptions to initially equilibrate a system, after which the systems are converted to atomistic detail. The insane (INSert membrANE) method uses the Martini FF and command-line tools to create arbitrary membranes at the coarse-grained level, which can be equilibrated and then converted to atomistic simulations.165,166 This is a potentially powerful approach, but there is no guarantee at the moment that Martini and atomistic FFs (or indeed different atomistic FFs) give the same equilibrium distribution of lipids in a mixture, insane is specific to GROMACS,167 and backmapping of very complex systems from Martini to atomistic is not always straightforward.

A second way of speeding up the equilibration of membrane simulations has been put forward by the Tajkhorshid group, called the Highly Mobile Membrane Mimetic (HMMM) approach.168 In this approach, the aim is to speed up lipid diffusion as it is often found to be the rate-limiting factor in membrane dynamics. Increased lipid mobility is achieved by separating the lipid heads from the tails; in fact, the HMMM bilayer consists of two monolayers of very short tail lipids with a bulk organic (or imaginary, as it does not have to actually exist as chemical) solvent in between to represent the membrane interior. The performance of the model was tested by comparing side chain free energy profiles between HMMM and full lipid representations, showing very good agreement in the interfacial part but less accuracy in the membrane interior.169 So far, the model has been mainly applied to study binding of peripheral proteins and has been shown to be an efficient tool to predict their membrane bound state.170

2.2. CG Models

The large time and length scales over which cellular processes operate has spurred the development of a large number of CG lipid FFs, following the pioneering work of Smit et al.171 and Goetz and Lipowsky172 in the 90s. Today, CG lipid models span all the way from a generic, supra-CG level of resolution to near-atomistic models. Here we focus on models that retain chemical specificity and are therefore able to distinguish specific lipid types. These kinds of models usually group 3−6 heavy atoms per CG bead, reducing a typical lipid to around
8–14 beads. Below we discuss the overall parametrization strategy for CG models (top down versus bottom up) and describe recent progress in some of the more popular CG lipid models used for cell membranes, namely the Martini, Shinoda/Klein (SDK), the SIRAH, and ELBA FFs, as well as a number of solvent-free models. The growing number of tools to automate the simulation workflow and the limitations inherent to CGing are also discussed. For a broader overview, we direct the reader to a number of other reviews on CG membrane simulations.  

2.2.1. Top Down versus Bottom Up. Parameterization of CG models may follow either a bottom-up strategy (also denoted structure-based coarse-graining) or a top-down strategy (thermodynamic-based coarse-graining). In the bottom-up approach, effective CG interactions are extracted from reference data, such as atomistically detailed simulations or structural databanks, aiming at a faithful reproduction of the structural features of the reference data. In the top-down approach, the focus lies on reproducing experimental data, especially thermodynamic properties such as density, heat of vaporization, and partitioning data. Both approaches have their own advantages and disadvantages. Focusing on reproducing structural details often leads to highly accurate CG models; however, the accuracy is usually limited to the state point at which the parameters were derived. Besides, the resulting CG potentials typically contain detailed features that limit the integration time step and are not always straightforward to interpret from a physicochemical point of view. Relying on thermodynamic data comes at the price of limited structural accuracy but with the benefit of reproducing global partitioning of the CG molecules over a wider range of state points. In practice, many CG FFs use a combination of these two approaches to maximize accuracy on the one hand and transferability on the other. Note that, inherent to the nature of coarse graining, it is impossible to obtain fully transferable models nor to represent all features of the underlying compound at the same time (the “representability problem”). There is no unique method to construct CG potentials from higher resolution data. A full representation of higher-order correlations requires multibody potentials, which are impractical and computationally expensive, thereby defeating the purpose of coarse graining. Even when the pair correlations are well-described, other system properties such as the pressure or energy cannot be matched at the same time unless higher-order terms are included in the force field. The art of coarse graining is in the compromise of assessing which level of detail needs to be included. The best choice of CG model, in the end, will depend on the application at hand. For in depth reviews on this topic, see, for example, Brini et al., Ingólfsson et al., and Noid.

2.2.2. Martini Model. The Martini FF, developed jointly in the laboratories of Marrink and Tieleman, is currently the most widely applied CG FF for biomembranes. The philosophy behind Martini is to present an extendable CG model based on simple modular building blocks, using few parameters and standard interaction potentials to maximize applicability and transferability. Martini uses an approximate 4:1 mapping and combines top-down and bottom-up parametrization strategies. Due to the modularity of Martini, a large set of different lipid types have been parametrized, covering all common lipid heads that can be straightforwardly combined with tails of varying length and degree of saturation. More specialized lipids, such as glycolipids, PEGylated lipids, cardiolipins, tetraether lipids, lipopolysaccharides (LPS), and a variety of sterols and sterol-like compounds (cholesterol, ergosterol, hopanoids) are available as well, enabling simulation of complex membranes with realistic lipid compositions (see section 3.2). The Martini model is implemented in a number of major simulation packages, including GROMACS NAMD, LAMMPS, as well as in the Materials Science Suite. In addition to lipids, Martini has been extended to the most important classes of biomolecules (proteins, carbohydrates, nucleotides), as well as a large variety of polymers and nanoparticles. This variety makes the Martini model ideally suited to study a wide range of membrane-related processes, including interaction with non-biological particles such as polymer-induced formation of nanodisks or penetration of gold particles. For processes for which long-range electrostatic interactions are deemed important, polarizable water and ion models have been developed. A major limitation of the Martini FF is the inability to model protein folding events. The use of isotropic interaction potentials cannot capture the directionality of hydrogen-bonding patterns that underlie protein conformational stability. Instead, an elastic network is used to constrain proteins, as well as nucleotides, to a reference (e.g., X-ray) structure. A recently introduced combination of Martini with Go models allows sampling also of unfolded protein states and is a promising method to further extend the range of applications. Another limitation, that also affects all-atom FFs, is the stickiness of larger biomolecules including proteins. Although this problem can be alleviated by ad-hoc approaches, for example, by downsampling protein–protein interactions or increasing protein hydration strength, the origin of the problem appears to reside in the different CG mapping densities of these biomolecules compared to the surrounding solvent. In the forthcoming new version of the model (Martini 3.0), these interactions have been balanced more carefully, resolving this issue. More background on Martini is provided in a perspective paper by the main developers and on the Martini webportal.  

2.2.3. SDK Model. Klein and co-workers are among the pioneers in developing CG lipid models. Their model is based on a 3:1 mapping and therefore somewhat more detailed than the Martini model. Besides, the model uses softer interaction potentials, allowing for a better reproduction of heats of vaporization and surface tensions. The latest version of the model, the SDK FF (Shinoda, Devane, Klein) also combines bottom-up and top-down parametrization and has resulted in improved transferability. Applications of the SDK model include studies of the phase behavior of lipid monolayers, vesicle fusion, and membrane partitioning of fullerenes (reviewed in Shinoda et al.). Recently the model has been extended to include triglycerides, allowing the study of formation of lipid droplets. A drawback of the SDK model is that only a limited number of lipid parameters are available currently, and no compatible protein model has been developed. Furthermore, the SDK model is only implemented in the LAMMPS software package, and no active development site is maintained. A recent extension of the SDK model, called the SPICA (Surface Property fitting Coarse graining) force field, includes improved parameters for cholesterol and different lipid types allowing realistic simulations of domain formation.
2.2.4. ELBA Model. The ELBA (electrostatics-based) CG lipid FF developed by Orsi and co-workers,\(^\text{219}\) focuses on modeling lipid–water interactions and capturing important electrostatic contributions. The model uses a 3:1 mapping but represents each water molecule individually using soft sticky dipole potentials and incorporates electrostatics in the CG lipid beads as point charges or point dipoles. A few lipid types have been parametrized by matching lipid properties, such as volume and area per lipid, average segmental tail order parameter, spontaneous curvature, and dipole potential. Most recently, an ELBA model for cholesterol has been developed that matches experimental phase behavior for binary DPPC/cholesterol mixtures.\(^\text{220}\) Applications of the ELBA FF have thus far been focused on permeation of drugs and other compounds across bilayers but only using some standard lipid types. Compared to Martini, the major advantage of the ELBA models lies in the more accurate description of the electrostatic interactions. As with the SDK model, however, only few lipid types have been parametrized, and the model is only available within LAMMPS. More information is available on the Web site http://www.orsi.sems.qmul.ac.uk/elba/.

2.2.5. SIRAH Force Field. SIRAH (South-American initiative for a rapid and accurate Hamiltonian) is a top-down CG FF developed by Pantano and co-workers to model proteins and DNA.\(^\text{221,222}\) The SIRAH model has a similar mapping as the Martini model and also treats solvent and ions explicitly. Interestingly, the SIRAH FF has recently been extended to include lipids.\(^\text{223}\) So far, only parameters for DMPC lipids have been published, but the ability to model lipids opens the way to a broad range of applications involving cell membranes in the future. The FF is available for both GROMACS and AMBER. An important aspect of SIRAH is that it allows sampling of conformational changes of proteins, due to a higher resolution of the peptide backbone. More details of the FF can be found at the Web site http://www.sirahff.com/.

2.2.6. Solvent-Free Models. A number of other models should be mentioned, in particular, recent attempts to parametrize solvent-free lipid models that retain chemical detail. Implicit solvent models considerably reduce computational cost but do need to incorporate the excluded solvent interactions into the effective potentials between the CG beads. In the pioneering work of the Voth group,\(^\text{224,225}\) a bottom-up strategy based on force matching between CG and AA systems is used to derive detailed solvent-free models for a number of different lipid mixtures. Hills and co-workers used this strategy also for development of a solvent-free protein model, CgProt,\(^\text{226}\) which was recently combined with a lipid FF parametrized using the same strategy.\(^\text{227}\) Lyubartsev and co-workers\(^\text{226}\) used another bottom-up strategy, the Newton inversion method, to capture the fine details of the AA lipid models into CG potentials. Wang and Deserno\(^\text{228}\) and Sodt and Head-Gordon\(^\text{229}\) followed a more pragmatic top-down approach, adding long-range attractive interactions in the lipid tails to mimic the hydrophobic effect, tuned to fit experimental data. The model of Wang and Deserno has also been successfully combined with a CG protein model and coined the PLUM model.\(^\text{231,232}\) Curtis and Hall\(^\text{233}\) in their LIME (lipid intermediate resolution model) FF, use hard-sphere and square-well potentials in order to use discontinuous molecular dynamics and gain even greater speedup. An implicit solvent version of the Martini FF has also been developed by the Marrink group, coined Dry Martini,\(^\text{234}\) using a rescaled interaction matrix that accounts for the hydrophobic and solvation effects. The Dry Martini model can also be combined with stochastic rotational dynamics to incorporate hydrodynamics (denoted STRD Martini).\(^\text{235}\) Wan, Gao, and Fang developed a DPD model based on Martini type mapping that can be used for both lipids and peptides.\(^\text{236}\) In a recent extension of the popular CG protein model PRIMO, developed by Feig and co-workers, an implicit membrane environment has been added to study membrane protein folding and aggregation.\(^\text{237}\)

2.2.7. Limitations/Developments of CG Models. As discussed above, parametrization and validation of CG models relies either on experimental data (top-down) or higher resolution data (bottom-up). Experimental data on suitable reference systems, however, is not always available or not easy to interpret. For instance, dimerization free energies of TM peptides in model lipid membranes form a perfect test system to validate CG simulations. The free energy of this process can be easily obtained from CG simulations with the help of advanced sampling and biasing techniques. In principle, this allows comparing to the same quantity derived from association constants measured using FRET assays. However, the bound and unbound states are ill-defined, hampering a straightforward comparison. Relying on all-atom reference simulations, on the other hand, is also problematic, for two reasons. First, sampling issues at the all-atom level prevent careful validation of most processes involving protein–lipid or protein–protein binding. Second, shortcomings of the all-atom models are inherited by the CG models. In this regard, it is helpful to calibrate CG models not on a single reference FF but to use multiple ones in the absence of clearly validated targets.

Naturally, limitations of CG models arise from the reduced level of resolution. As discussed above, most CG models face limitations in the extent to which protein structural transitions can be captured, owing to the absence of directional hydrogen bonds or alternative potentials that introduce directionality. One avenue to improve the accuracy of CG models is through multiscale, combining the sampling speed of CG models with the accuracy of atomistic models. This can be achieved in a static way, in which part of the system is modeled at high resolution and surrounded by a CG environment or in a dynamic way in which molecules can change their resolution on the fly. Despite the progress in multiscale method development, applications of such methods to lipid membranes have been very limited. In a proof of principle application,\(^\text{238}\) a multiscale method was used to simulate an atomistic protein channel in a CG Martini bilayer. Proper coupling of the electrostatic interactions between the two levels of resolution, however, remained problematic due to the poor short-range screening behavior of the CG solvent. To achieve a quantitatively more accurate method, cross optimization of the interactions between CG and the atomistic FF is probably necessary as has been attempted in the PACE FF in which Martini lipids are combined with a near-atomistic protein model.\(^\text{239}\) The ELBA FF has also been used in a multiscale setup, in particular to study permeation of AA drugs across CG membranes.\(^\text{240}\) The level of detail retained in the ELBA model is high enough that the AA-CG cross interactions can be based on standard combination rules. Multiscale simulations with the SIRAH FF have also been reported\(^\text{241}\) but not (yet) involving lipid membranes. In an implicit membrane environment, the PRIMO FF can be combined with CHARMM.\(^\text{242}\)
At the moment, more powerful are so-called serial multiscale schemes that are used to reconstruct all-atom detail from a given CG configuration ("backmapping"). Most commonly applied backmapping tools for lipid systems include fragment-based approaches,243,244 simulated annealing,244 and usage of geometrical rules.166,246,247 There is also a promising new multiscale tool GADDLE maps which is based on a Monte Carlo sampling algorithm.248 Typically, backmapping is used either to validate specific interactions observed in CG simulations or to focus on some atomic details of the system of interest. Note, however, that the amount of sampling that can be performed at the atomicistic level is usually limited. Therefore, finding that a CG configuration is also stable at the atomicistic level, albeit encouraging, is not a proof of the validity of the CG model. The opposite, for example, observing that the CG configuration is unstable at the all-atom level, may however point to a limitation of the CG model.

Milano and co-workers have developed an interesting hybrid particle-field scheme, combining molecular dynamics with self-consistent field theory (the hybrid MD-SCF).249,250 The main difference of the hybrid MD-SCF method in relation to other CG approaches is that the calculation of the nonbonded interactions between the CG particles is replaced by an evaluation of an external potential on the local density. With this scheme, the hybrid MD-SCF method allows the usage of mapping and bonded parameters commonly used in other CG approaches in combination with an efficient parallelization for the calculation of interaction forces, obtained via an average density field.251 Lipid applications are still limited, which includes simulations of phospholipids in bilayer and non-lamellar phases, with lipids mapping and bonded parameters based in the Martini scheme.252,253 More recently, a flexible CG model for protein has been introduced, allowing studies of conformational changes, even in a lipid environment.254 The hybrid SCF-MD is available in a dedicated software package called OCCAM. More details of the method are available at the Web site http://www.occam/md.org/.

2.2.8. High-Throughput Tools. One of the advantages of CG models is that they provide easy access to high-throughput applications. Hundreds or thousands of simulations can be performed, systematically exploring, for example, lipid membrane composition or protein mutant libraries. A nice example is the membrane protein database MemProtMD, developed by Sansom and co-workers: based on self-assembly simulations, configurations of all classes of membrane proteins embedded in a natural lipid environment are provided.255,256 To facilitate high-throughput applications, many new and improved methods have been developed to help set up initial simulation configurations. A key example is the CHARMM-GUI framework (see also discussion above), which currently supports also the CG Martini FF.257,258 A drawback of CHARMM-GUI is that it is not command-line-based and therefore cannot be integrated into automated workflows. An example of a command-line-based tool is Moltemplate (http://www.moltemplate.org/), a generic molecular builder for LAMMPS, with support for the CG models Martini and SDK. Another command-line based tool called insane is a popular membrane-building tool associated with the Martini FF and allows for on the fly generation of new lipid templates.165 A number of programs have also been developed that automatically setup and run CG simulations for high-throughput screening of protein–protein interactions, such as Sidekick259 and Docking Assay For Transmembrane compounds (DAFT).260 To further automate the simulation workflow, current efforts are also being directed toward automated CG topology builders.261–264 Here, one of the main challenges is to automate the mapping of the underlying atomistic structure to the CG representation, a nontrivial problem. The power of such a tool is illustrated in a recent paper from Bereau and co-workers,265 who established linear relations between bulk membrane partitioning and the potential of mean force covering more than 40000 drug compounds.

2.3. Supra-CG Models

A longer-term aim of simulation of complex biological membranes is to enable us to relate molecular structures of their lipids and protein components to cellular phenotypes. This requires us to be able to compare the behavior of membrane simulations more directly to experiments at the cellular level, for example, via various super-resolution imaging modalities. The CG models described above all have a similar level of granularity, whereby each CG particle corresponds to 3–4 heavy (i.e., not hydrogen) atoms, such that, for example, a phospholipid molecule is represented by 10–15 CG beads. The advantage of this level of granularity is that it allows retention of chemical specificity of, for example, lipid headgroups in their interactions with proteins. The disadvantage is that it restricts practical applications to systems of ~2 M particles (i.e., ~8 M heavy atoms), equivalent to a length scale of <100 nm, on time scales up to the millisecond range. We need to move beyond these limitations in order to address dynamic events in membrane cell biology. For example, at the lower scale of cell membrane events, a clathrin-coated vesicle has a diameter of 100 nm and is formed by budding on a time scale of ∼10 s.266 Here we discuss current approaches to simulate such large-scale collective phenomena, requiring a further reduction in resolution denoted supra-CGing. For other reviews in this field, see, for example, refs 267 and 268.

2.3.1. Supra-CGing Approaches. In order to address events on these larger scales, supra-CGing approaches are needed. A number of approaches may be adopted in order to reach the desired meso and micro scales. At a simple level, one can employ CG models with fewer particles, for example, just a few particles per lipid molecule (e.g., the model by Ayton and Voth267) or even a few particles to represent a protein molecule or domain (e.g., models by Zhang et al.270,271). Alternatively, one may both reduce the number of particles and use modified interactions that smoothen the energy surface (as in DPD models, e.g., Venturoli et al.272). A more radical level of simplification (to reach even larger scales) may be to integrate out lipids (and water) altogether, such that proteins are represented as particles interacting in a continuum membrane environment. For all of these approaches, parameterization is a challenge, especially if one wishes to retain a degree of chemical specificity in these higher-level models, which is essential if they are to be used to address genuinely biological questions. Voth and co-workers have developed a theoretical framework for obtaining and interpreting such supra-CG models.

2.3.2. Few-Bead Lipids. A number of groups have explored CG models in which only a small number of particles are used to represent each lipid molecule.273,274 For example, Voth and colleagues have developed a framework for "aggressive" CGing of lipids in which, for example, two or three particles can represent each lipid molecule in a (solvent
free) model. This can be used to simulate, for example, 200 nm diameter lipid vesicles containing ~500000 lipids. A related model has also been developed for charged lipids to capture the electrostatic interactions of their headgroups in a “broad brush” fashion which has been used to model both mixed lipid vesicles and (peripheral) protein/charged lipid bilayer interactions. A key feature of these models is to combine analytical potentials (e.g., Gay-Berne models) to describe the generic anisotropic behavior of the lipids with more detailed force-matched potentials that provide an element of chemical specificity.

A similar level of granularity to that in “standard” CG representations is employed in DPD models, which smooth the energy surface for interactions between lipid molecules. The advantage of the soft potential employed is to enhance diffusion, although it may result in, for example, unphysical lipid overlaps. DPD models have been used, for example, to examine the effects of cholesterol on lipid bilayer structure. Comparable models have also been applied to examine mechanisms of fusion between lipid bilayer and vesicles. A supra CG model for lipids based on soft interactions has also been developed by Laradji and co-workers and applied to study biomembrane behavior, but parametrization will be challenging if biologically realistic specificity is to be preserved in such models.  

A similar level of granularity to that in “standard” CG representations is employed in DPD models, which smooth the energy surface for interactions between lipid molecules. The advantage of the soft potential employed is to enhance diffusion, although it may result in, for example, unphysical lipid overlaps. DPD models have been used, for example, to examine the effects of cholesterol on lipid bilayer structure. Comparable models have also been applied to examine mechanisms of fusion between lipid bilayer and vesicles. A supra CG model for lipids based on soft interactions has also been developed by Laradji and co-workers and applied to study biomembrane behavior, but parametrization will be challenging if biologically realistic specificity is to be preserved in such models.

A promising route to supraCG models of membrane proteins that retain a degree of specificity, in terms of the irregular shapes and dynamics of those proteins, is provided by the work of, for example, Voth and colleagues in which protein domains are represented by a small number of particles. The supraCG mapping in these ED-CG models is achieved by matching the dynamics of the CG model to a more detailed essential dynamics (ED) description derived either from atomistic simulations combined with PCA or by an elastic network model (ENM) of the protein. These models have been used in, for example, studies of membrane remodelling (see below).

An even coarser level of granularity, in terms of representation of membrane proteins, has been explored in a DPD study of the organization of membrane protein complexes in simple models of photosynthetic membranes. In this study, the protein complexes were represented by a model in which a protein (or protein oligomer) is represented by a single particle, combined with a two particle per lipid molecule model. The protein particles were parametrized phenomenologically on the basis of experimental (electron microscopy) data for their supercomplexes. Such a model allows large-scale (500 nm) organization of membranes to be explored, although the parametrization does rely on appropriate experimental data being available.

2.3.3. Reduced Protein Models. There are two broad approaches to the representation of proteins in supraCG simulations. One approach relies on idealizing/simplifying the representation of proteins in a fashion which (it is hoped) will retain the essence (but not the specific chemical details) of protein/protein and protein/lipid interactions. The other relies on simplified models of proteins comprising just a few particles, the interactions between which are parametrized on the basis of more detailed CG and/or atomistic simulations. The former approach has been used extensively within a DPD framework to study protein/lipid and protein/protein interactions in simplified models of biomembranes. The latter approach has been applied to much larger and more complex biological membrane systems in order to capture their emergent behavior on a meso scale.

In a series of DPD simulation studies, Smit et al. have modeled membrane proteins as, for example, rod-like structures with hydrophobic cores and polar caps and have used these to exploit protein/membrane interactions and also the free energy landscapes of protein–protein interactions within membranes. Protein/protein interaction potentials of mean force (PMFs) computed from DPD simulations have been used to develop larger scale 2D models in which proteins are treated as disks interacting through those PMFs. This provides an interesting route to capturing protein/protein interactions in very large-scale simulations. Weiss and colleagues have also used comparable DPD models of membrane proteins to explore, in a generalized fashion, the influence of membrane protein structure on, for example, diffusion. DPD simulations in which membrane binding proteins were represented as highly simplified Janus-like particles have been used to propose models of large-scale dynamic events such as membrane vesiculation. Again, this provides an interesting supraCG route to large-scale biomembrane behavior, but parametrization will be challenging if biologically realistic specificity and complexity is to be preserved in such models.

3. INCREASING COMPLEXITY

The complexity of cellular membrane is really staggering. There exists more than a thousand different lipid types that are found in biological membranes, with, in some cases, hundreds present in the same membrane. Embedded in this complex
lipid mixture is a plethora of membrane proteins, either transmembrane or peripherally bound. On top of this, many cell membranes are highly curved, and interact with components of the surrounding medium such as the cytoskeleton or neighboring organelles or cells. And all of this happens under, constantly changing, nonequilibrium conditions. To capture this complexity, it is evident that we need to move beyond modeling highly simplified model membrane systems containing one or two lipid components only and being surrounded by excess aqueous solvent, notwithstanding the continued usefulness of studying simplified systems.

Below, we describe our current ability to increase the level of complexity, making use of the improvements of lipid FFs as discussed in the preceding section. We first describe simulation studies that thoroughly explore the basic behavior of multicomponent lipid and lipid protein mixtures (section 3.1), followed by the ongoing efforts to model specific membranes in realistic detail (section 3.2), on our way to full cell models (section 3.3).

3.1. Multicomponent Membranes

In this section, we provide an overview of the growing body of simulation studies that consider multicomponent membranes to understand the organizational principles of membranes at a fundamental level, including formation and structure of lipid domains, binding of specific lipids to membrane proteins, membrane-mediated protein–protein interactions, and lipid-or protein-induced membrane curvature.

3.1.1. Lipid Domains. Given the diversity of lipid types in cell membranes, a nonuniform distribution of lipids in the lateral plane of the membrane is rather likely. In fact, the heterogeneous nature of the cell membrane is underlying the raft concept, which, in its current form, states that specific lipids together with proteins can cluster into nanodomains. These nanodomains may be transient and too small to be detected by experimental means but may also grow into more stable functional platforms when needed. The propensity to form distinct phases is already found in model membranes composed of ternary mixtures of saturated lipids, unsaturated lipids, and cholesterol, capable of forming coexisting liquid-ordered (Lo) and liquid-disordered (Ld) domains. In fact, ternary and quaternary mixtures display a rich behavior of domain formation processes ranging from critical fluctuations, modulated phases, all the way to macroscopic phase separation. Interestingly, extracts from real cells show similar phase behavior, pointing to the possible biological relevance of this fundamental aspect of multicomponent lipid membranes.

Although mean field theories describe these phenomena in a qualitative way, MD simulations prove essential in providing the molecular details of both the structural and kinetic aspects of lipid nanodomains. Important insight into the structure of the Lo phase has been obtained by recent all-atom models of the groups of Vattulainen and Lyman, revealing the presence of substructures within these domains. Still, with all-atom models, spontaneous segregation into coexisting Lo/Ld domains is proving difficult to observe, likely hampered by the slow kinetics of phase separation. Only the onset of the process has thus far been captured with atomistic models.

Here, CG models have proven very valuable. An important breakthrough was reported by Risselada and Marrink, who simulated the spontaneous formation of Lo and Ld domains in ternary mixtures of saturated and unsaturated lipids together with cholesterol based on the Martini model. Follow up studies have further explored the properties of these domains as a function of lipid composition, including the effect of hybrid lipids (lipids with one saturated and one unsaturated tail) that act as lineactants (i.e., decrease the line tension between the domains). Making use of high-throughput simulation strategies, complete lipid phase diagrams can nowadays be established, for example, for binary lipid/cholesterol systems as a function of temperature, as well as ternary and even quaternary mixtures. Figure 3 provides an example of a ternary phase diagram from Carpenter et al. based on the Martini model, in comparison to the experimental phase diagram. In the study of Ackerman and Feigenson, concentrations of DPPC and cholesterol are fixed, whereas the nanodomain-inducing lipid 16:0,18:2-PC (PUPC) is incrementally replaced by the macrodomain-inducing lipid 18:2,18:2-PC (DUPC). Extensive

Figure 3. Example of a multicomponent membrane phase diagram. The ternary lipid mixture dioleoyl phosphocholine (DOPC), DPPC, and cholesterol exhibits a range of interesting phase behavior. Experimental phase diagram is shown on the left, and a simulated diagram using further optimized Martini parameters for DOPC and DPPC on the right. Inserts show snapshots of four of the different simulations illustrating the phase separation. The lipids are colored red, blue, and green for DOPC, DPPC, and cholesterol, respectively. Adapted from ref 335. Copyright 2018 American Chemical Society.
simulations of this four-component system reveal that lipid demixing increases as the amount of DUPC increases, in agreement with the experimental phase diagram. Furthermore, domain size and interleaflet alignment change sharply over a narrow range of replacement of PUPC by DUPC, indicating that intraleaflet and interleaflet behaviors are coupled. It turns out to be not always trivial to assign phases in these mixtures, due to the challenge of identifying a phase based on local physical properties, in addition to challenges due to finite size and hysteresis effects.328,331,337

An interesting and ongoing topic of discussion is the extent to which the domains in opposing leaflets are registered (i.e., occupy the same lateral position as a result of an interleaflet coupling mechanism). In principle, domains can be both registered or antiregistered, or anything in between. Theoretically, driving forces that govern the extent of registration include minimization of the line tension at the domain boundaries, minimization of the interleaflet surface tension, release of curvature frustration, electrostatic coupling, lipid or cholesterol tail interdigitation, and flip-flopping of additives or cholesterol.338–340 Currently, simulation studies in support of each of these mechanisms can be found,48,341−349 probably pointing to a subtle interplay of all of these effects occurring in realistic membranes.

To add another layer of complexity, the organization of the nanodomains can be tuned by a number of additional factors, as evidenced by simulation studies on the effect of stress340 and immobilization,351,352 as well as the addition of other additives such as hydrophobic compounds,353,354 sugars,355,356 or ions.357 However, predictions from simulations are not always in agreement with experiments,358 pointing to differences between experimental and computational time and length scales (notably, the contribution of the domain boundaries is much more dominant in simulation studies359 and/or deficiencies in the nature of the FFs, and/or challenges in the interpretation of experimental data.

Due to the different nature of the Lo and Ld domains, the former being enriched in cholesterol and saturated lipids and more densely packed, one naturally expects a nonequal distribution of membrane proteins between these phases. Experiments confirm this expectation, showing a common preference for proteins to reside in the Ld domains, unless specific lipid anchors, commonly post-translationally attached to membrane proteins in vivo, are present.360 A number of recent MD studies have addressed the driving forces underlying this sorting process. A pioneering study was performed by Schäfer et al.361 revealing the molecular mechanism behind the generic preference for peptide and proteins to reside in Ld domains. In accordance with the authors, inclusions disturb the tight packing of saturated lipids and cholesterol in the Lo domain, providing an enthalpic driving force for sorting into Ld domains. In a subsequent work, De Jong et al.362 demonstrated that lipid anchors can indeed provide a counter-force to steer proteins toward the more ordered Lo domains. The importance of lipid anchors in dictating sorting behavior is also clearly demonstrated in the work of Gorfe and co-workers,165−167,359 as well as others.362,366 One highlight is the study showing that different Ras variants (H-Ras, N-Ras, K-Ras) have different propensities to segregate into Lo or Ld domains, driven by the opposite preference of palmitoyl and farnesyl anchors for ordered and disordered membrane domains.363 Localization of Ras clusters at the domain boundaries may further lead to a reduction in line tension and destabilization of the domains. Parton et al.367 show that influenza hemagglutinin, a TM protein containing a number of palmitoyl anchors, also resides in proximity of Lo regions.

Apart from lipid anchors, simulation studies have revealed a number of other mechanisms that affect sorting. Restriction of tilt, for instance as a consequence of protein anchoring to the cytoskeleton, is important as it prevents release of lipid mismatch through sorting.368 Likewise, fixed membrane curvature may lead to sorting of TM peptides.369 In addition, specific lipids can mediate the sorting behavior. A striking example is the sorting of peptides into Lo domains under the influence of gangliosides.362,370 Another example is the observation of cholesterol mediated sorting in a joint experimental-computational framework.371 Here it is shown that cholesterol can constrain the structural adaptations at the peptide-lipid interface under mismatch, resulting in a sorting potential.

For more in depth discussion on the topic of lipid domain simulations, please consult recent reviews from Bennett and Tieleman372 and Rög and Vattulainen,373 as well as the comprehensive review of Hof and co-workers covering both experimental and computational studies on membrane nano-domains.374

### 3.1.2. Protein−Lipid Binding Sites

Identifying lipid binding sites on membrane proteins is a rapidly growing area. Experimentally more and more lipid binding sites are being discovered,375 thanks to an increasing number of techniques.

Traditionally, X-ray techniques may reveal tightly bound, cocrystallized lipids, but advanced mutagenesis studies or chemical cross-linking techniques are used to probe also weaker bound lipids that are washed away under the harsh crystallization conditions. New techniques such as the use of lipid nanodisks to isolate membrane proteins with their native lipid environment,376 as well as mass-spectrometry (MS),377 hold a lot of promise to further this development. On the basis of the strength of binding, two classes of lipid binding sites can be differentiated, namely specific and nonspecific binding sites. The former involves tightly bound lipids that occupy specific sites inside or at the protein surface. The latter refers to lipids only showing a weak protein affinity, occupying the annular shell around the protein. The ability of membrane proteins to recruit and bind specific lipid types is of functional importance. For instance, lipid binding may dictate the sorting behavior of proteins between different membrane domains and may facilitate protein insertion by lowering the cost of hydrophobic mismatch;378 bound lipids may either protect proteins against aggregation (locking mechanism) or bridge proteins together into functional supercomplexes (bridging mechanism); specific binding sites may be involved in enzymatic reactions (e.g., donating protons or electrons) or more generally provide structural stability and stabilize specific protein conformations.

Computational studies are entering this field at a rapid pace. Both all-atom and CG simulations have proven useful to look at both specific and nonspecific lipid binding. In a typical setup, a membrane protein is embedded in a bilayer composed of two or three components, including the putative binding lipids. Binding sites are then identified by constructing density maps (“heat” maps) and some user defined density threshold. In principle, MD simulations also allow for quantification of the strength of lipid binding, through computation of a PMF.379,401,85,380 In AA simulations, strongly bound lipids can usually be distinguished from weak or nonbinding lipids, but
equilibration remains problematic. In particular, lipid exchange rates may become prohibitively large. Even nonbinding lipids can occupy a given site for hundreds of nanoseconds. Currently, the best strategy is to use multiple runs from different starting configurations to get a handle on the reproducibility of the results. For instance, Rogaski and Klauda\textsuperscript{181} generated five different orientations of a peripheral membrane protein to study its binding to a lipid bilayer and found consistent interaction modes requiring the presence of anionic lipids. Alternatively, CG models can be used. At CG resolution, reversible lipid–protein binding events can be observed on time scales of 10–100s of microseconds. In a state-of-the-art example, Arnarez et al.\textsuperscript{382} identified six cardiolipin (CL) binding sites on the respiratory chain complex cytochrome bc\textsubscript{1}. To provide a fully atomistic view of these binding sites, the CG configurations were backmapped to AA resolution, a procedure also frequently used by the group of Sansom.\textsuperscript{383,384} AA simulations can also serve to refine X-ray data, as shown by Aponte-Santamaria et al.\textsuperscript{385,386} in the case of cocrystallized DMPC lipids around aquaporin-0.

The full power of simulation models in this area is best demonstrated by a fast-growing number of studies that can reproduce experimental binding sites. An impressive example is the specific binding of C18-SM to a binding pocket in the TM domain of the COPII machinery protein.\textsuperscript{407} In this joint computational-experimental study, MD simulations reveal a close interaction between C18-SM and the transmembrane domain, as suggested by mutagenesis data. Interestingly, the interaction is found to be very specific, depending on both the headgroup and the backbone of the sphingolipid, as well as on a signature sequence of the protein. The verification of experimental cholesterol binding sites of GPCRs is another hot topic. Successful validations have now been made in case of the human A(2A) adenosine receptor,\textsuperscript{388–390} the \(\beta\)-adrenergic receptor,\textsuperscript{391,392} and rhodopsin,\textsuperscript{393} recently reviewed by Sengupta et al.\textsuperscript{394} (Figure 4). Simulations of sterol binding to the known sterol binding sites of the oxysterol binding protein osh4\textsuperscript{395} and the mitochondrial voltage gated anion channel VDAC1,\textsuperscript{396} cholesterol binding to the family of monoamine transporters,\textsuperscript{397} as well as a study of binding of cholesteryl esters to their binding pocket in the cholesteryl ester transfer protein\textsuperscript{398} are also worth mentioning. An increasing number of examples exist also for anionic lipids, which are key regulators of membrane protein behavior. For instance, simulations of the Sansom group identify phosphatidylserine-triphosphate (PIP\textsubscript{3}) binding sites on the pleckstrin homology domain\textsuperscript{399} and PIP\textsubscript{2} binding sites on the inwardly rectifying potassium (Kir) channels\textsuperscript{400,401} in line with X-ray and mutagenesis data. Other examples include binding sites for CL on the respiratory chain complexes cytochrome c oxidase and cytochrome bc\textsubscript{1},\textsuperscript{382,401} reproducing sites known from earlier structural studies and buried into protein cavities, as well as validation of CL binding sites on the mitochondrial ADP/ATP carrier\textsuperscript{402,403} and reproduction of known DPPC binding sites of potassium channels KcsA and chimeric KcsA-Kv1.3.\textsuperscript{404}

Given the ability of current simulation studies to reproduce known lipid binding sites, the prediction of novel binding sites becomes interesting. A number of studies on respiratory chain complexes reveal hitherto unknown CL binding sites on the membrane-exposed surfaces of these proteins.\textsuperscript{382,401,405–408} Surface-bound CLs could play an important role in, for example, proton uptake or by providing structural integrity of the complexes and supercomplexes. CL binding sites are also predicted for a number of other bacterial proteins.\textsuperscript{409–412} Likewise, simulation studies on GPCRs are pointing at novel cholesterol binding sites, for instance, in case of the serotonin(1A) receptor,\textsuperscript{413} the A(2A) adenosine receptor,\textsuperscript{414} and the Smoothened receptor,\textsuperscript{415} and on the importance of PIP\textsubscript{2} in regulating conformational states.\textsuperscript{415} Cholesterol binding sites are also found on the Kir2.2 channel, depending on the open/close state of the channel.\textsuperscript{416} In a study on nicotinic acetylcholine receptors (nAChR), cholesterol competes with PUFA containing lipids to occupy binding sites.\textsuperscript{417} The list of predicted lipid binding sites keeps growing. Other examples are the discovery of a PE binding site, stabilized by a lipid-mediated salt bridge, on lactose permease,\textsuperscript{418} and of numerous PG binding sites on the ammonium transporter AmtB.\textsuperscript{419} Binding sites for PI lipids were found in a combined experimental and computational study of the eukaryotic purine symporter UapA.\textsuperscript{420} Their presence at the dimer interface suggests a role in structural stability of the complex. A similar study revealed PIP\textsubscript{2} binding to mammalian two pore channels, forming a cross-link between two parts of the channel and enabling their coordinated movement during channel gating.\textsuperscript{421} The power of combining modeling and experimental studies is further illustrated by the discovery of ceramide binding sites on one of the TM helices of Late Endosomal Protein LAPTM4B.\textsuperscript{422} Although quite often overlooked, the absence of lipid binding sites is also useful information that can be extracted from simulations. For instance, in a multiscale study, Stansfeld et al. report no specific DPPC binding sites for various members of the aquaporin family.\textsuperscript{423}

Noteworthy are also an increasing number of studies that reveal pathways for protein-mediated lipid flip-flop, a mechanism that has been hypothesized already a while ago but not been observed directly (Figure S). Khelashvili et al. identified the existence of a series of weak lipid binding spots along the surface of Opsin, allowing flip-flopping of POPC lipids (Figure 5A). Spontaneous penetration of lipid head groups into the membrane interior along the aque duct nhTMEM16, a fungal scramblase, was observed during a simulation of Tjakhroshid and co-workers\textsuperscript{224} (also noted earlier\textsuperscript{244}), resulting in a continuous file of lipids connecting the outer and inner leaflets (Figure 5C).
Complexes is observed in the study of Huber et al.428 Here, spontaneous lipid scrambling to occur on a submicrosecond (i.e., lipid lined) pore. MD simulations show indeed nonspecific flippase activity of the SecYEG complex.427 (C) Overlay of POPC configurations bound to a fungal scramblase, defining a flip-flop pathway. In all snapshots, the lipid phosphate groups are represented by red spheres. The tails are colored from yellow to green to visualize the flip-flop pathway.

Figure 5. Protein-induced lipid flip-flop. (A) Overlay of POPC lipids bound to Opsin at intermediate stages of the flip-flop pathway.423 (B) CG trajectory of a DOPG lipid during a (partial) flip-flop mediated by the SecYEG complex.427 (C) Overlay of POPC configurations bound to a fungal scramblase, defining a flip-flop pathway.424 In all snapshots, the lipid phosphate groups are represented by red spheres. The tails are colored from yellow to green to visualize the flip-flop pathway.

POPC lipid from the inner leaflet to the outer leaflet of the membrane was captured during the simulation. Additional full translocations of lipids, including the charged lipid POPs, took place when a transmembrane voltage was applied. Further details of the translocation pathway of nHTMEM16 were revealed in a similar study by Lee et al.425 and in a study on the TMEM16K member of the family.426 In another study, Koch et al.427 discovered a potential flippase activity of the SecYEG channel. They found the presence of a single binding site for DOPG lipids that can be reached from both membrane leaflets and thus allows spontaneous flip-flop of an anionic lipid (Figure 5B). Transfer of lipids between different protein complexes is observed in the study of Huber et al.428 Here, extensive simulations were used to show a downhill pathway of transfer of lipid A from bacterial outer membrane models through CD14 to the terminal TLR4/MD-2 complex. Aksimentiev and co-workers429 reported a synthetic DNA nanopore specialized in flipping lipids, by stabilizing a toroidal (i.e., lipid lined) pore. MD simulations show indeed spontaneous lipid scrambling to occur on a submicrosecond time scale.

Furthermore, MD simulations are frequently used to study nonspecific lipid binding. Key papers in this area are from Vattulainen and co-workers430,431 in which they show strong correlations between the lateral diffusion of membrane proteins and a shell of 50–100 annular lipids. This kind of protein–lipid binding is observed even in lipid membranes composed of a single lipid type and therefore clearly nonspecific. On the basis of CG simulations of a large variety of membrane proteins, Sturgis and collaborators show that even simple mixtures of nonspecifically bound lipids already give rise to a complex perturbation pattern around the protein that is not easily described by elastic membrane deformations and can be long-range in nature.432 In multicomponent membranes, nonspecific lipid binding will lead to an inhomogeneous distribution of lipids around the protein in general. One of the first examples in this respect is the study of Grossfield et al. on the enrichment of polyunsaturated lipid chains around rhodopsin433 in mixed SDPC/SDPE/Chol bilayers. A recent extension of this work points at a dependence of these interactions on the conformational state of the protein.434 Other examples of nonspecific lipid binding include the recruitment of short-tail lipids around OmpA,435 of anionic (POPA) lipids at the gap junction hemichannel connexion-26,436 the nonspecific binding of POPs to TM and juxtamembrane domains of cytokines437 and to integrin,438 weak binding of annular PC/PE lipids to MsbA flippase,439 accumulation of cholesterol at the C-terminal helix of a phospholipase scramblase,440 accumulation of GM1 around WLP peptides441 and aquaporin442 preferential binding of PE over PC lipids in secondary transporters443 and VDAC444 and redistribution of PC lipids around the gramicidin A channel dependent on tail length and unsaturation.445

A related topic is the membrane binding of peripheral proteins, which often requires the presence of specific lipids that provide the necessary driving force for stable protein–membrane interactions. Simulation studies in this category reveal the CL mediated membrane binding of creatine kinase (Mtk),444 the enrichment of phosphatidic acid (PA) and PIP2 at the membrane binding spot of actin capping protein (CP),445 the increased propensity of negative lipids to interact with a glycosyltransferase,446 and PAs clustering around the acylated pleckstrin homology domain-containing protein (APH).447 The signaling lipids PI and PIPs seem to play a particularly important role. Many examples can be found in the recent simulation literature in which PIs or PIPs drive the protein binding and orientation, for example, the phosphohomology domain (PX),448 auxilin,449 BIN1/M-Amphiphysin450 MIM I-BAR,451 GTPases,452 Kidney- and BRAINexpressed protein (KIBRA),453 the HIV-1 matrix protein,454 as well as a variety of actin binding proteins.455–457 Other examples of lipid-mediated membrane binding are the ganglioside mediated binding of cholera toxin458 competitive binding of bis(monoacylglycerol)phosphate and SM to Niemann-Pick Protein C2,459 the role of cation-pi interactions in stabilizing the binding of phospholipases,460 and the modulating effect of cholesterol on the depth, orientation, and conformation of the membrane binding fragment caveolin-1.461

Nonspecific clustering as well as specific binding of lipids by membrane active peptides is another area where simulation studies are increasingly being used.462–464 Simulations addressing lipid-peptide interplay in case of amyloid peptides also remains a hot topic, in particular with respect to membrane-mediated fiber formation.465–468 For a more elaborate description of modeling of protein and peptide-lipid interactions, we refer to previous reviews by Hedger and Sansom,469 Grouleff et al.,470 Wen et al.,471 as well as to Corradi et al. in this issue.472

3.1.3. Lipid-Mediated Protein Oligomerization. Lipid-mediated interactions are of key importance in driving the clustering of membrane proteins. Clustering forces of this kind include release of membrane curvature stress, capillary condensation, lipid depletion effects, and Casimir type forces (arising from perturbed fluctuations in, for example, lipid density or thickness). Depending on the system details, such lipid-mediated effects can dominate direct protein–protein interactions and can be long-range in nature.473–475 Understanding the molecular driving forces that are ultimately responsible for the sorting and clustering of membrane proteins is currently an active field of research in which simulation studies play a key role.476

To systematically study clustering of membrane proteins, basically two approaches are followed: spontaneous self-assembly, which allows also for the formation of higher-order oligomers, or biased simulations to determine the protein–protein dimerization free energy. In self-assembly simulations, the complexation of proteins is simply followed over time. Self-
assembly simulations are typically performed using many independent replicas to probe the reproducibility of the interfaces formed. The behavior of WT and mutant proteins can be compared, providing additional insights into the packing motifs. In the case of single TM helices, a growing number of studies demonstrate that the experimentally determined interfaces can be reproduced, even when using CG models to speed up the sampling (reviewed by Psachoulia et al.\(^{477}\)). Such approaches require a high-throughput approach to obtain statistically relevant results.\(^{478,479}\) The ability to predict packing of TM helices has paved the way for CG modeling studies of self-assembly of larger protein complexes, in particular, protein complexes in which the interface is formed by single TM helices.\(^{480−487}\) For instance, in a joint experimental/modeling effort, van den Boogaart et al.\(^{483}\) revealed the molecular organization of syntaxin clusters and showed that syntaxin clustering is mediated by electrostatic interactions with the strongly anionic lipid phosphatidylinositol-4,5-bisphosphate (PIP2). The mediating role of PIP2 is also apparent from a simulation study of the binding of FERM to L-selecin.\(^{437}\) Not only is PIP2 required for efficient binding of FERM to the plasma membrane but also PIP2 induces a conformational change of L-selecin which allows the formation of the heterocomplex. Computational evidence for the role of PIP2 in activating integrin, by stabilizing an integrin-talin heterocomplex, has also been obtained,\(^{486}\) as well as revealing the role of PIPs in modulating EphA2, a receptor tyrosine kinase.\(^{487}\) As another example of lipid-mediated protein complex formation, PS lipids are found to steer the formation of the RAS/RAF complex in a multiscale study by Travers et al.\(^{486}\)

Self-assembly studies of polytopic membrane proteins are still hampered by slow kinetics, but the onset of protein oligomer formation can be simulated. Following the pioneering studies of Periole et al.,\(^{487}\) the oligomerization tendency of GPCRs as well as other membrane proteins has been simulated through self-assembly by a number of groups.\(^{488−492}\) These studies, mostly based on the Martini FF, often indicate formation of stringlike clusters of proteins with preferred protein–protein interfaces. A clear effect of cholesterol was reported in steering the dimer interface formation in case of the beta2-adrenergic receptor\(^ {493}\) and chemokine receptors.\(^ {494,495}\) Similarly, both cholesterol and PIP2 lipids were found to affect the interfaces formed in large-scale self-assembly simulations of human serotonin transporters.\(^ {496}\) Self-assembly simulations of respiratory chain complexes by Arnarez et al.\(^ {482}\) show complexation between cytochrome bc1 and cytochrome c oxidase and point toward a specific role for CL in bridging the proteins together. Clustering of mitochondrial translocases was also shown to depend on the presence of CL in large-scale CG MD simulations.\(^ {502}\) In another recent example, gangliosides were observed to bridge tetranspan CD81 proteins into higher-order aggregates (Figure 6).\(^ {497}\) However, on the multimicrosecond time scale accessible with current simulations, equilibration of the protein–protein interfaces of polytopic membrane proteins has not yet been achieved. Reversible sampling of protein–protein binding/unbinding events is extremely challenging for models that retain chemical specificity, the more so in the crowded environment of real cells (see below).

In addition to affecting the kinetics, protein crowding can also impact the thermodynamic behavior of the system. Domanski et al.\(^ {498}\) found that, at lipid/protein ratios characteristic of real membranes, TM peptides can induce coalescence of their annular lipid shells triggering large-scale domain segregation. This is reminiscent of capillary condensation, predicted by Mouritsen and co-workers already in the late nineties based on MC simulations using a highly simplified membrane model.\(^ {509}\) Likewise, Ackerman and Feigenson\(^ {480}\) observed growth of nanodomains induced by WALP TM peptides, in agreement with the experiment. They furthermore showed that WALPs can induce registration of domains. Lipid-mediated protein crowding was also observed in the study of Guixà-González et al.\(^ {501}\) Here, the presence of lipids containing omega-3 polyunsaturated fatty acids (PUFA) drives the oligomerization of adenosine A2A and dopamine D2 receptors, again via coalescence of the annular protein shells enriched in the PUFA containing lipids (Figure 7).

In addition to self-assembly approaches, biased simulations can be used to predict binding interfaces and obtain insight into the thermodynamic driving forces for protein–protein aggregation by computation of the PMF. Although in certain

**Figure 6.** Example of lipid-mediated protein–protein oligomerization. The snapshots show how gangliosides mediate cluster formation of tetranspan CD81 proteins. Reproduced from ref 497. Copyright 2014 American Chemical Society.

**Figure 7.** Example of lipid-mediated protein clustering. Snapshots of clustering of GPCRs in a healthy membrane (left), containing DHA, and a nonhealthy membrane (right) depleted of DHA. In the healthy case, higher order oligomers, stabilized by DHA, are more present. Reproduced with permission from ref 501. Copyright 2016 Nature (http://creativecommons.org/licenses/by/4.0/).
cases protein–protein PMFs can be extracted from self-assembly simulations (when multiple binding/unbinding events are spontaneously observed),\textsuperscript{502} in most cases calculation of the PMF requires forced binding using, for example, umbrella sampling or metadynamics type approaches. Like soluble proteins, the large number of possible interfaces still poses a major sampling problem. On the one hand, due to the two-dimensional nature of the membrane, prediction of membrane protein interfaces is easier. On the other hand, sampling of the desolvation of the protein–protein interface is problematic due to trapping of lipids. Even between simple TM helices convergence of the PMF requires sampling on the microsecond time scale, often necessitating the use of CG models.\textsuperscript{503–507} These studies provide important insights on the driving forces for helix–helix association and, in particular, the contribution of lipid-mediated effects versus direct protein contacts. In a systematic study on the dimerization of WALP peptides under different mismatch conditions, Castillo et al.\textsuperscript{504} concluded helix–helix association to be enthalpically favorable in all cases, while the entropic contribution appears favorable only in the presence of significant positive hydrophobic mismatch. The interpretation of this requires care given the coarse-grained nature of these simulations, but the sign of the thermodynamic quantities agrees with experimental measurements on dimerization of (AALALAA)\textsubscript{3} peptides, and the observed association free energies are within the experimental range. The work of Benjamini and Smit is also noteworthy.\textsuperscript{508} The authors use a CG DPD model to challenge the notion that packing of TM helices is determined by specific interactions. Instead, the authors show that hydrophobic mismatch, through its effect on helix tilt, can explain many experimental cross-angle distribution features. This notion is supported by all-atom simulations of glycoporphin A dimers embedded in different membrane environments, showing robust packing motifs despite the poor hydrophobic match, using mechanisms based on dimer tilting or local membrane thickness perturbations.\textsuperscript{509} In a combined experimental/computational study, Cybulski et al. show that mismatch-induced tilting of TM helices forms the basis for the thermosensing mechanism of DesK in bacterial membranes.\textsuperscript{510}

Due to the sampling issues mentioned above, calculation of PMFs between polytopic membrane proteins has thusfar been limited to specific interfaces only. The first PMFs between fully solvated polytopic membrane proteins, GPCRs, was reported by Periolo et al.\textsuperscript{488} The authors show that sampling times exceeding 1 ms are required to obtain converged profiles for specific binding interfaces. Remarkably, it was found that the amount of protein burial (i.e., number of protein–protein contacts not exposed to lipids) does not correlate with the binding strength of the interface. This finding challenges the potential utility of buried accessible surface area as a predictor of the strength of membrane-embedded protein–protein interfaces, a strategy that works well for soluble proteins. This view was not confirmed, however, in the case of another membrane protein, NanC.\textsuperscript{511} Here, the strength of binding was found to be proportional to the number of protein–protein contacts. Clearly, more work is needed in this area. In general, the above, and other recent studies on GPCRs by Filizola and co-workers\textsuperscript{512,513} on the dopamine transporters by the group of Stockner\textsuperscript{514} and on the human serotonin transporter\textsuperscript{515} reveal specific, favorable, association interfaces stabilized by energies of the order of 30–60 kJ mol\textsuperscript{-1}. Considering the quantitative predictive capability of CG models, a warning is in place, however. Overstabilization of TM helix dimer formation has been reported for the widely used Martini model with respect to the all-atom OPLS FF,\textsuperscript{516} as well as compared to experimental data.\textsuperscript{514} On the contrary, other studies show a much better match, either with atomistic data\textsuperscript{516} or in comparison to available experimental data.\textsuperscript{504} Given the high sensitivity of dimerization free energies to the exact mismatch conditions (see for instance the work of Benjamini and Smit\textsuperscript{508}), and the importance of a proper choice of reaction coordinates,\textsuperscript{85} more systematic studies are needed to solve this controversy. Further progress in the use of enhanced sampling techniques will be very valuable in this respect. For instance, Lelimousin et al. recently showed that metadynamics can be used to induce reversible binding/unbinding of the TM domain of EGFR to obtain free energy landscapes directly.\textsuperscript{514} Domanski et al.\textsuperscript{517} used replica exchange umbrella sampling to speed up convergence of the PMF between glycoporphin TM domains. Application of these methods to polytopic membrane proteins should, in principle, be possible. Another example is the combination of MD with Markov state models, as used in the study of Filizola and co-workers to elucidate the association kinetics of \(\mu\)-opioid receptors.\textsuperscript{513} Advanced protein–protein docking tools such as HADDOCK are currently being extended into the realm of membrane proteins and could provide an alternative route toward prediction of protein–protein complexes.\textsuperscript{517}

For more detailed information on the topic of membrane protein oligomerization, see for instance the general review on protein–protein interactions by Baaden and Marrink,\textsuperscript{518} reviews on GPCRs by Periolo,\textsuperscript{519} Gabhauer and Böckmann,\textsuperscript{520} and Meng et al.\textsuperscript{521} and a review focusing on driving forces by Johannes et al.\textsuperscript{522}

### 3.1.4. Membrane Curvature Generation and Sensing

Curvature generation and sensing is important for many cellular processes that involve membrane remodeling, such as fusion and fission, and shaping of internal cellular compartments.\textsuperscript{523–525} In general, membrane curvature may result from the presence of nonlamellar forming lipids (e.g., DOPE) or arise from any asymmetry between the membrane monolayers. In vivo, generation of large curvatures typically requires the action of specialized proteins. Protein-induced membrane curvature could be an activated process, for example, making use of molecular motors or polymerizing actin filaments but also arise from direct protein–lipid interplay. In the latter case, three mechanisms can be distinguished: scaffolding, crowding, and insertion.\textsuperscript{523} In scaffolding, the proteins adhere to the bilayer and induce curvature through their curved interaction interface. In crowding, proteins located at the membrane surface generate a pressure that produces a bending moment acting on the membrane. In the insertion mechanism, proteins generate a curvature stress by asymmetric insertion of hydrophobic or amphipatic domains in the membrane. The three mechanisms are not mutually exclusive, however, and may act together. The number of simulation studies that address these mechanisms is steadily growing, following the pioneering simulations of Reynwar et al.\textsuperscript{287} on large-scale membrane remodeling using a generic CG model, the all-atom simulation by Blood and Voth,\textsuperscript{526} showing membrane curvature generation by a BAR domain, and the four-scale description of membrane sculpting of BAR domains by Arkhipov et al.\textsuperscript{527}

A recent example in this area is the work of Davies and co-workers,\textsuperscript{528,529} probing the role of F1F0-ATP synthase dimers
in shaping the mitochondrial cristae. On the basis of large-scale CG MD simulations, the authors propose that the assembly of ATP synthase dimer rows is driven by the reduction in the membrane elastic energy, rather than by direct protein contacts, and that the dimer rows enable the formation of highly curved ridges in mitochondrial cristae. De Oliveira Dos Santos Soares et al. used all-atom and CG MD simulations to investigate membrane-bending forces in the Dengue virus envelope. The structural organization of three heterotetramers EM proteins (EM3 unit) serves as an anisotropic bending unit for the Dengue virus envelope because it is able to locally decrease the thickness of the membrane with its short transmembrane helices. The simulations show that the specific arrangement of the EM membrane proteins inflict a curvature stress on the membrane. The resulting elastic energy is minimized by the systematic migration of lipids from the lower into the upper layer. Membrane undulations induced by the NS4A domain of Dengue virus have also been reported and were linked to the U-shape of this membrane spanning protein. Simulations of the curvature field induced by α-synuclein, by Sachs and co-workers, are also good examples of the power of near-atomistic membrane modeling in this field. On the basis of simulations involving 48 copies of the N-terminal membrane-binding domain of α-synuclein, together with more than 85000 lipids, the onset of membrane tubulation could be observed due to the collective action of the proteins (Figure 8). Another recent example of curvature generation due to surface-bound proteins is a simulation study of Li and Gorfe on asymmetrically bound H-ras proteins. Imposing curvature stress on membranes is not limited to proteins; small amphipatic peptides that adsorb at the membrane/water interface potentially have the same effect. Buckling of model bilayers can, for instance, be induced by antimicrobial peptides (AMPs) as is shown in simulations of Woo and Wallqvist. Sodt and Pastor have quantified the curvature stress generated by a model amphipatic peptide, showing that the peptide induces positive curvature in line with the conclusions from a simulation study on fusion peptides. The extent of curvature induction was found to depend sensitively on the molecular interactions and cannot be explained using simple shape-based concepts. Pannuzzo et al. proposed an efficient approach to simulate the bending power of peptides based on the use of lipid bicelles that are stabilized by short-chain lipids.

As mentioned before, membrane curvature may also arise from an asymmetric distribution of lipids between the leaflets. MD simulations of multicomponent membranes show, for example, that the ganglioside GM1 induces curvature. Conversely, curvature leads to lipid sorting as is demonstrated in a number of simulation studies. Again, simple shape-based concepts do not suffice to explain lipid-induced curvatures and sorting effects, in particular in multicomponent systems where effects are nonadditive. Noteworthy are also simulations that show how electrostatic fields can induce membrane curvature (flexoelectric effect).

In addition to curvature generation, an important question is how proteins may sense different curvatures. Cui et al. used the concept of membrane-packing defects as measure for curvature sensing. The idea is that curved membranes expose a larger fraction of hydrophobic defects to which the hydrophobic domains of proteins can bind. Indeed, the authors demonstrate, based on all-atom MD simulations, an increasing number of defects with increasing curvature. Another study showed that the ability of lipid tails to backfold to the membrane/water interface also increases with curvature. Vamparys et al. furthermore show that the size and number of such defects increase with the number of monounsaturated acyl chains and with the introduction of conical lipids. Moreover, the size and probability of the defects promoted by conical lipids resembled those induced by positive curvature, thus explaining why conical lipids and positive curvature can both drive the adsorption of surface active peptides and proteins. This hypothesis was confirmed by subsequent studies in which experimental data and simulation data were combined to explain the binding affinity of peripheral proteins as a function of lipid composition and curvature as well as membrane tension as another determining factor. A simulation study of a buckled membrane demonstrated differences in sensing characteristics between different AMPs. Thus, proteins can sense curvature, and induce curvature, but can also undergo conformational changes in response to curvature. A number of MD studies have demonstrated that membrane curvature can indeed shift the conformational equilibrium in peptides, as well as affect peptide folding kinetics.

Fusion and fission are key cellular processes that involve extensive membrane curvatures. The main question remains to what extent fusion and fission are lipid-driven or protein-mediated. Simulation studies have contributed significantly in this area, and protein-free fusion pathways between lamellar membranes and between vesicles are now quite well-established (reviewed in refs and ). Current efforts are directed to calculate the energetics and kinetics of the various intermediates, the importance of hydration forces in the initial approach, stalk formation between multicomponent phase separated membranes, the role of calcium and PEG in mediating fusion, and carbon nanotube-mediated fusion. Computational modeling of peptide and protein-induced fusion or fission is still in its pioneering phase. A number of researchers investigated the ability of small amphipatic peptides, including HA fusion peptides, to stabilize cubic phases and stalk/pore complexes that are relevant as fusion intermediates. Moiset et al. found that certain AMPs, that are traditionally associated with forming transmembrane pores, can also induce stalk formation between juxtaposed membranes. Stalk formation in this case is initiated by the ability of multiple lysine residues to form a bridge between the apposing bilayers and trigger the flipping of lipid tails between the proximal leaflets. A similar mechanism was recently observed in all-atom simulations of stalk formation in the presence of arginin-rich cell penetrating

Figure 8. Membrane curvature generation by proteins. Onset of membrane tubulation induced by 48 copies of α-Syn100 (yellow) interacting with a membrane composed of 85296 POPG lipids (blue tails, red headgroups). Water is not shown for clarity. Snapshot is obtained at 300 ns simulation time with the Martini model. The budding tubule extends ∼25 nm above the bulk lipid bilayer. Adapted from ref Copyright 2014 American Chemical Society.
peptides. Baoukina and Tieleman simulated the fusion of small unilamellar vesicles mediated by lung surfactant protein B (SP-B). They found SP-B monomers capable of triggering fusion events by anchoring two vesicles, facilitating the formation of a lipid bridge between the proximal leaflets. In a series of breakthrough papers, Risselada et al. simulated neuronal SNARE-mediated membrane fusion. The simulations reveal that SNARE complexes operate in a cooperative and synchronized way. In the postfusion state, zipping of the SNAREs extends into the membrane region, in agreement with the recently resolved X-ray structure of the fully assembled state. Additional details of the fusion pathway were resolved in simulations of SNARE mediated fusion between bilayers and nanodisks. In the work of Pinot et al., a combination of in vivo, in vitro, and in silico experiments were used to show the combined effect of lipids and proteins in shaping membranes during fission. In particular, the role of polyunsaturated lipids in membrane vesiculation by dynamin and endophilin was revealed. The simulations provided the molecular mechanism: polyunsaturated lipids can backfold in the membrane and thereby adapt their conformation to the change in membrane curvature during vesiculation. This plasticity of polyunsaturated lipids was already noted in earlier simulations of small liposomes.

### 3.2. Realistic Cell Membranes

Recent advances in computation and molecular FFs have allowed for more faithful modeling of realistic biological cell membranes. Focusing on the lipid component only, recent models are approaching realistic complexity of biological bilayers with respect to the number of different lipid types, bilayer asymmetry, and geometry. Marked differences are found in the lipid composition and dynamics of bilayers from different organisms, cell and tissue types, organelle, as well as dependent on environmental factors and cell cycle, therefore, a diverse set of bilayer models is needed. Here we list some of the different types of membrane models that have been developed.

#### 3.2.1. Plasma Membranes

In cells, the plasma membrane (PM) defines the boundary, separating the cell interior from the outside environment. A typical PM contains hundreds, if not thousands, of different lipid species that are actively regulated by the cell and nonuniformly distributed in the membrane plane. In particular, the role of polyunsaturated lipids in membrane vesiculation by dynamin and endophilin was revealed. The simulations provided the molecular mechanism: polyunsaturated lipids can backfold in the membrane and thereby adapt their conformation to the change in membrane curvature during vesiculation. This plasticity of polyunsaturated lipids was already noted in earlier simulations of small liposomes.

![Figure 9](image_url). An example of a complex plasma membrane model. Corradi et al. simulated ten different membrane proteins in a 63 lipid PM mixture. Each protein’s different TM shape and lipid–protein interactions resulted in a unique lipid fingerprint. Here AQP1 is depicted showing the simulation setup, snapshot of the outer membrane as well as lipid enrichment/depletion and bilayer properties around the protein. Adapted from ref. Copyright 2018 American Chemical Society.
phase separation is believed to be important, a three-component mixture is often used (with a high and low melting temperature phospholipid and cholesterol, see above). Recently, however, models attempting to approximate realistic PMs have emerged, some of which we will discuss here.

In terms of lipid composition, Ingólfssson et al. average mammalian PM model\textsuperscript{593} is the most complex simulation to date. The model is Martini-based and contains 63 different lipid types, with 14 different types of lipid headgroups and 11 different tails that are asymmetrically distributed across the leaflets. A large-scale simulation of the model membrane, \(\sim 20000\) lipids and simulated for 80 \(\mu\)s,\textsuperscript{93,596} gives a high-resolution view of the dynamic interplay of all lipid species and overall organization. PM nonideal lipid mixing, membrane properties, lipid flip-flop dynamics, leaflet coupling, and domain formation were explored. At the microsecond time scale cholesterol, ceramide and diacylglycerol lipids domain formation were explored. At the microsecond time scale curvature gradients, together forming unique lipid fingerprints, were explored. Vo et al.\textsuperscript{604} investigated the effect of cytoskeletal immobilization on protein and lipid mobility,\textsuperscript{605} how loading with model transmembrane helices or GPCRs effect membrane dynamics (such as bilayer undulation and lipid diffusion),\textsuperscript{606} and lipid binding to receptor tyrosine kinases (RTKs)\textsuperscript{607} and the epidermal growth factor receptor (EGFR).\textsuperscript{608}

A number of other average or specific tissue type PM models have been developed. Jeevan et al. made an average PM model using Martini that is asymmetric and has six lipid components.\textsuperscript{609} The model has been used to explore Ebola virus protein V\textsubscript{P40} PM binding.\textsuperscript{609} Hedger et al. explored the cholesterol interaction of the Class F G protein-coupled receptor Smoothened in a number of bilayers, including an asymmetric five-component Martini lipid bilayer containing PC, PE, PS, PIP\textsubscript{2}, and cholesterol.\textsuperscript{380} Kalli et al. constructed an asymmetrical five lipid type PM model using Martini to explore integrin receptor dynamics, showing how the receptor altered lipid organization especially that of cholesterol and PS.\textsuperscript{610} Domicevica, Koldso, and Biggin constructed a five-component asymmetrical epithelial brain PM to explore the lipid interaction of P-glycoprotein.\textsuperscript{597} They used both CG Martini and atomistic Slipid models and found enrichment of charged PS lipids next to the protein and specific cholesterol interaction sites. A five-component Slipid model was also used to explore the effect of curvature on PM properties by Yesylevskyy et al.\textsuperscript{611} Yesylevskyy and co-workers recently made a variant of the model to mimic a cancerogenic PM.\textsuperscript{612} Klähn and Zacharias build asymmetric five component PM models representing a cancerogenic and normal eukaryotic PM and simulated them using the atomistic CHARMM FF.\textsuperscript{513}

Ueoka and co-workers made compositionally complex asymmetrical PM models of normal and cancerogenic thymocyte membranes using the CHARMM FF and containing 23 and 25 different lipid types, respectively.\textsuperscript{600} Flinner and Schleff constructed an asymmetrical ten component Martini bilayer model of the red blood cell (RBC) PM to explore the dynamics of glycoporfin A dimers.\textsuperscript{95} Characteristic of RBC PMs, the model mixture is high in cholesterol and contains PE plasmalogon lipids. Kalli and Reithmeier constructed asymmetric six component RBC PMs both using Martini and GROMOS to study the interactions between the Band 3 and glycoporfin A proteins and the lipids.\textsuperscript{614} Kadri et al. constructed symmetric epithelial cell membrane models with 10 different Martini PC lipids to study how the increased tail saturation associated with lipotoxinication effects bilayer properties.\textsuperscript{615} Ingólfssson, Carpenter, and co-workers assembled a human neuronal PM model, based on Martini, with an asymmetric lipid distribution and 58 different lipid types.\textsuperscript{596} Compared to the 63-lipid type average PM model the bilayer properties of the neuronal PM are overall strikingly similar, despite significant difference in lipid composition. The effects of the higher cholesterol content of the neuronal bilayer are somewhat compensated by the higher tail unsaturation. Interestingly, the domain sizes fluctuations in both the neuronal brain and average PM mixtures were sensitive to the level of bilayer undulation. Guixá-González et al. constructed two six component symmetrical brain models with high and low docosahexaenoic acid (DHA) concentration to explore DHA role in GPCR oligomerization.\textsuperscript{610}
Klauda and co-workers created PM models of the soybean hypocotyl (the stem of the germinating seedling) and root using the atomistic CHARMM FF; the models are symmetrical and contain 9 and 10 lipid types, respectively. Soybeans lipid composition differs significantly from eukaryotic membranes, containing different sterols (sitosterol and stigmasterol instead of cholesterol) as well as a large fraction of di- and tripolyunsaturated fatty acid tails. Jo et al. made a six-lipid type symmetric average yeast membrane using CHARMM and characterized its properties with more tail saturation, less sterol content, and imposed surface tension.

### 3.2.2. Organelle Membranes

The lipid composition varies widely between organelles requiring a large set of membrane models, which to date is significantly under-represented. One of the more studied organelles is mitochondria, the “powerhouse” of the cell, generating most of the ATP that the cell uses. Cardiolipin (CL) is the signature lipid of mitochondria. It is anionic, with two phosphate groups and four acyl tails. Cardiolipin is present at a high concentration in the inner membrane of mitochondria (up to 20%) and is required to stabilize the respiratory chain supercomplexes. Mitochondrial membranes have mostly been modeled as symmetric binary or ternary mixtures consisting of PC or PC/PE with cardiolipin. At the CG Martini level, inner mitochondrial mixtures have been used to explore cardiolipin protein binding for the respiratory chain supercomplexes, the rotor of the metazoan ATP synthases, and the adenine nucleotide translocase (ANT). Vähäheikkilä et al. explored the inner mitochondrial bilayer properties using the atomistic OPLS FF with different cardiolipin variants and levels of cardiolipin tail peroxidation.

Thylakoid compartments are the sites of lipid-dependent photosynthetic reactions in chloroplasts and cyanobacteria. The thylakoid membranes are rich in galactolipids, their lipids have a high fraction of polysaturated tails, and many of their lipids are nonlamellar phase lipids. Van Eerden et al. created two thylakoid membrane models: a five-lipid type cyanobacterial model and a seven-lipid type plant model, both at the CG Martini level and atomistic GROMOS level. The bilayer properties of both membrane models were evaluated as well as the dynamics of two photosynthesis cofactors (plastoquinone and plastoquinol) inserted in the membranes. Later studies have used thylakoid membranes to explore the dynamics of the Photosystem II (PSII) complex, focusing on the protein lipid interaction and the entry and exit of membrane embedded cofactors to the protein, as well the membrane interaction of the cold-regulated (COR) protein COR15A.

Ray et al. constructed symmetric four to six component membrane models for the endoplasmic reticulum (ER), Golgi apparatus, and mitochondria using the CHARMM FF and analyzed the distribution of forces within the membranes. Su et al. modeled a peroxisomal membrane from the yeast *Pichia pastoris*, as a symmetrical five component mixture using the Martini model, which they used to explore the lipid association and aggregation of the N-terminal helix of the peroxisome elongation protein. Monje-Galvan and Klauda modeled the PM, ER, and trans-Golgi Network (TGN) bilayers of yeast and compared their properties with the previously constructed average yeast membrane model. The models were built using the atomistic CHARMM FF. The membrane is kept symmetric, with 6–11 different lipid types, and for each organelle, two models were made with different levels of tail unsaturation. Simulations of the model membranes highlight differences in bilayer properties (e.g., thickness, area per lipid, compressibility) between the different organelar membranes.

### 3.2.3. Bacterial Membranes

The lipid composition of different bacteria is quite diverse. Gram-negative bacteria, such as *E. coli* and *S. aureus*, have an inner and outer cell membrane, separated by a viscous periplasm. The outer leaflet of the outer membrane is mainly composed of lipopolysaccharide (LPS) lipids. LPS consist of Lipid A, with 4−7 fatty acid tails attached to a sugar backbone and a polysaccharide forming an inner and outer core and a variable length O-antigen. The different constituents of LPS can vary significantly both within and between bacterial species. A range of different LPS variants and fragments have been parametrized. For CHARMM, the CHARMM-GUI web portal now has an LPS Modeler that as of May 2018 has 15 bacteria species, 37 lipid A types, 52 core oligosaccharide types, and 304 O-antigen polysaccharide types.

A number of models of the outer bacterial membrane has been constructed, both at the CG and atomistic level of resolution. Typically, these models contain 2−5 lipid types and are asymmetric, with the outer leaflet consisting mostly of different variants of LPS, and the inner leaflet either DPPPE or a mixture of PE, PG, and sometimes cardiolipin. The models have been used to explore and characterize different basic properties of the outer bacterial membrane such as density, packing, average area per lipid, diffusion, and divalent cation binding. In addition, partitioning and permeation of molecules into and through the membrane, the influence and packing of membrane proteins, and the effects of Lipid A structural variations from different pathogenic bacterial species and within species. Other bacterial membranes have also been modeled. Models for chlamydia’s (*C. trachomatis*) two main life cycles, the elementary body and reticular body, have been developed within the CHARMM FF. Both simulated membranes are asymmetrical and contain nine different lipid types, corresponding to the most prevalent lipids of the different life cycles, including three lipid types with methyl branched tails. Klauda and co-workers constructed a cytoplasmic *E. coli* membrane in the CHARMM atomistic FF. The membrane model is symmetrical, containing the six most prominent lipid types in the inner membrane, including a lipid containing a cyclopropane ring within the acyl chain tail that they parameterized. In a later study, they modeled the *E. coli* inner membrane at different stages along the growth cycle, showing significant differences in average area per lipid and rigidity.

Hwang et al. also used the CHARMM FF to study the effect of stress on the *E. coli* cell envelope, modeling both the outer and inner membrane as well as the cell wall. Berglund et al. explored the interaction of the antimicrobial peptide polymyxin B1 with both the outer and inner *E. coli* membranes, using the GROMOS atomistic FF; their inner membrane model was a symmetric three component mixture of mostly PE with some PG and cardiolipin. Hsu et al. also modeled both membranes using the Martini model and simulated them with various native membrane proteins embedded, including the outer/inner membrane spanning AcrAB-TolC complex (Figure 10). For a review focusing on simulations of bacterial membrane channels, see ref 654.
3.2.4. Skin Models. The outer layer of the skin (stratum corneum, SC) consists of dead cells (corneocytes). The lipid “structure” of the stratum corneum is a mixture of long-chain saturated ceramides, free fatty acids, and cholesterol, in a 1:1:1 ratio. Due to their relevance for skin barrier properties, numerous simulation efforts have studied the properties of these lipid mixtures. Here, we mention a few of the more recent studies.

McCabe and co-workers have modeled the SC using mixtures of ceramides and fatty acids as well as ceramides, fatty acids, and cholesterol, using both the CHARMM FF with modified ceramide parameters and the Berger lipid FF as well as a customized CG FF for lipid self-assembly. Wang and Klauda characterized bilayer properties of pure ceramide bilayers and SC ternary models using the CHARMM FF at different temperatures and ceramide tail length. Ho et al. modeled the SC with a combination of fatty acids and cholesterol using the GROMOS FF. Das et al. constructed a SC model containing 15 ceramide variants as well as a fatty acid and cholesterol using the atomistic Berger lipid FF and found a preference for the inverse micellar phase. Del Regno and Notman modeled SC at two different lipid concentrations and two levels of hydrations using the Berger FF. They suggest a permeation path for small polar molecules through the SC lamellae that avoids pockets of water between the bilayers. Wennberg et al. simulated glycosylceramides and ceramides at varying levels of hydration using the Martini FF. They showed that glycosylceramides can maintain a cubiclike bilayer structure while the ceramides collapse into a stacked lamellar structure, which might be an important step for SC. In two studies, Gupta and Rai explored fullerene C_{60} permeation through SC bilayers using the Martini FF, and using an atomistic SC model they studied electroporation by imposing a varying external electric field.

3.2.5. Complications of Complexity. Membrane models need to be complex enough for the question at hand, but additional complexity comes with a prize. Before adopting a more realistic, more complex model, the price of doing so should be carefully evaluated. Here we discuss some of the caveats that need to be considered.

Bilayer models with more lipid species require longer sampling times, especially if rare lipid species are included. Proteins can affect their local lipid environment, promoting lipid sorting and/or bilayer perturbation, see for example, refs 603, 610, and 667, also extending the required sampling time. The sampling challenge is even bigger when considering more realistic conditions characterized by a high protein density. Domanski et al., Goos and Sansom, and Javanainen et al. simulated membranes under such crowded conditions, with formation of extended clusters and networks of proteins dramatically slowing down the lateral diffusion rates of the components. Interested readers are pointed to a recent review on protein crowding. In fact, under crowded conditions, diffusion becomes anomalous and may lead to deviations from the Saffman-Delbruck model at physiological levels. Although, in the latter case, these claims are not substantiated as a proper correction of periodicity artifacts on hydrodynamics has not been taken into account.

Many biological membranes are asymmetric; therefore, more physiologically relevant models of those membranes may need to include asymmetry. To model an asymmetrical membrane, first, it is necessary to determine the lipid concentrations in each leaflet. This is not trivial as our current knowledge about lipid asymmetry is incomplete. For specific lipid classes and membranes, the asymmetry has been determined, see for example, ref 3, but for most membranes, many lipid classes, and most individual lipid types, the asymmetry is not well-determined or not known at all. Second, including asymmetry in a periodically constrained system, a primary concern is to determine the relative number of lipids in each leaflet. Several different criteria have been proposed to determine what a “correct” balance of outer/inner leaflet lipids should be, such as matching the average area per lipid (APL) in both leaflets, the leaflets surface tension, the lateral pressure profile across the two leaflets, and the lipid’s chemical potential. Generating a “well-balanced” asymmetric membrane using one of these criteria can be quite involved. Recent simulation setup protocols for asymmetrical bilayers include using prior estimates of idealized APL, biased self-assembly, an iterative building procedure, or zeroing bilayer leaflet tension. Additional complexity also arises when including lipids that can flip-flop between the leaflets at time scales relevant for the simulation at hand. Cholesterol is a good example of a fast flip-flopping lipid; it has been shown to flip-flop on the microsecond time scale.

Realistic bilayers, depending on lipid mixture, protein content, and cell attachment, can undulate significantly. Allowing for larger bilayer undulations is computationally very expensive. The simulation box has to be large both in the plane of the bilayer (to reduce undulation dampening due to periodic image constraints) as well as perpendicular to the plane, increasing simulation cost. Longer simulations are also needed to capture the longer length scale bilayer undulation modes and lipids redistribution, as lipids have been shown to organize in the plane of the bilayer based on curvature. All analysis of undulating bilayers also becomes more complex as the curved bilayer surface needs to be fitted and accounted for.
With increased model complexity, the sampling required goes up exponentially. Analysis therefore also becomes a bottleneck, as the amount of generated data mirrors the sampling and the number of interactions between components to analyze also goes up with the square of the number of species in the model. For the more complex models, they fast become intractable for manual analysis, requiring reduction in complexity (e.g., combining lipid species into classes), automated analysis methods, and/or use of unsupervised machine learning methods for identifying possible hidden correlations.

3.3. Toward Full Cell Models

With the use of the CG and multiscale approaches described above, it is possible to perform very large-scale simulations of cell membranes, at the level of, for example, the envelope of a complete virus particle.676,677 While appealing as a tour de force, one should ask what might be learned from such simulations. A major motivation is to overcome barriers between simulations and experiments. Thus, very large scale simulations may allow us to approach the length and/or time scales of experimental studies of biological membranes, which in turn will enable direct comparison between experiments and simulation, permitting rigorous molecular interpretations of mesoscopic observations. This is especially important in linking molecular structures of membranes and their components through to cell biological investigations using a variety of imaging modalities, for example, cryoelectron tomography and super-resolution optical microscopies.

There are spatial and temporal challenges in matching mesoscale experimental data while not losing molecular specificity in the underlying models. How we can address such challenges is illustrated via a number of examples of increasing scale: (i) viral envelope membranes via very large-scale CG simulations; (ii) bacterial outer membrane protein (OMP) clustering by large scale CG simulations enabling parametrization of simple mesoscale models; and (iii) CG and meso scale simulations to study processes of remodeling of eukaryotic cell membranes.

3.3.1. Viral Envelopes. CG-MD has been used to explore the membranes of a number of enveloped viruses, providing examples of very large-scale (ca. 5 million particles) simulations of biological membrane assemblies. A ground-breaking study of the membrane envelope of the immature HIV-1 virion578 combined electron cryotomography data and multiscale simulations to provide insights into the Gag lattice assembly process in the immature HIV-1 virion. These simulations employed a multiscale approach in which multiple CG parameters were explored in critical regions, with the aim of identifying those interactions that are critical to maintaining the structure of the virion. Subsequently, the CG simulation results were used to guide all-atom MD simulations of selected regions in order to refine the model.

Simulations of a complete virion envelope model for influenza A combined X-ray structures and TM domain models for the hemagglutinin (HA) and neuraminidase (NA) proteins, an NMR structure for the TM domain of the M2 protein, and a lipid bilayer composition based on the experimentally determined lipidome of the viral membrane.679 The prevalence of glycolipid headgroups on the outer surface of the influenza A viral membrane suggested that access of therapeutic compounds to the M2 proton channel may have to overcome substantial steric barriers. The influenza A envelope proteins moved slowly within the cholesterol-rich membrane, with diffusion constants matching previous NMR measurements. Lipid molecules had reduced diffusion coefficients (D) and exponents (α) less than 1, the latter indicative of anomalous diffusion. The spacing between membrane glycoprotein molecules on the influenza A surface suggested that polyvalent interactions between HA and/or NA on the viral surface and sialic acid residues on the host cell membrane are likely to occur. This would enable strong virus-host association despite relatively weak (~2–3 mM affinity) viral HA-single host receptor interactions in vitro.

The membrane envelope of the dengue virus has been simulated in two recent studies680,681 using the Martini FF. Reddy and Sansom681 used a combination of CG modeling and simulation to “add back” the lipid bilayer to the cryo-EM structure of the Dengue virus envelope proteins. These simulations revealed that the crowding of protein TM domains and the enclosure of the outer leaflet of the lipid bilayer within a protein shell resulted in lipid diffusive properties similar to those in the “raftlike” influenza A membrane, despite the absence of cholesterol from the dengue membrane model. Bond and colleagues680 used a novel protocol to embed the cryo-EM structure of the envelope protein complexes of the DENV-2 icosahedral shell within a spherical lipid vesicle, the composition of which was guided by lipidomics data. Microsecond-time scale simulations of the virion envelope enabled refinement of the lipid/protein complex, assessed by comparing density maps calculated from simulations with those determined by cryo-EM. The refined structures revealed locally induced curvature resulting from specific interactions with phosphatidylserine molecules. These lipids may facilitate subsequent fusion of the viral envelope with the host membrane inside the endosome during infection. A subsequent study,682 based on targeted MD simulations, provided evidence that the low pH structures obtained with cryo-EM are biologically meaningful intermediates of the fusion process with the endosomal membrane.

A hybrid multiscale approach using the SIRAH FF has been used to study the envelope of zika virus (ZIKV).683 Those parts of the system of particular interest were modeled using atomistic and/or CG resolution, while those of less direct interest used a supra-CG resolution. This hybrid multiscale approach allows for efficient simulations of large-scale biological membrane systems to be run on modest computational resources, thereby making computational virology accessible to a wider range of researchers. These and other studies demonstrate the potential of very large-scale simulation of viral envelopes. Future challenges for such studies include development of a full CG model of glycosylation of viral surface proteins, which will enable more realistic and hence predictive modeling of virions binding to models of target cell membranes.

3.3.2. Large-Scale Membrane Organization. Large-scale simulations can be used to probe the structural and dynamic consequences of protein–lipid and protein–protein interactions in complex and crowded cellular membranes. For example, simulations of a mitochondrial inner membrane protein model indicate how cardiolipin may “glue” together respiratory proteins into supercomplexes.818 Analysis of the free energy landscape of interactions of the bacterial outer membrane protein (OMP) NanC have also revealed how intervening lipids may stabilize a membrane protein dimer.514 Such protein–lipid–protein interaction may underlie functionally
important larger scale membrane organization. Earlier work in this area\textsuperscript{64,685} provided a theoretical framework for our understanding of the role of lipids in mediating membrane protein interactions and for extracting appropriate parameters from large scale MD simulations. Recently, highly coarse-grained simulations have been used alongside experiments to explore the interplay of lipids and proteins which underlie clustering of the influenza M2 protein and its possible role in mediating viral budding from infected host cells.\textsuperscript{686}

Large-scale CG-MD simulations can in turn enable more highly coarse-grained (or mesoscopic) simulation approaches to be developed for modeling of emergent behaviors in these complex protein–membrane systems (Figure 11). For example, the spatiotemporal organization of membrane proteins is often characterized by the formation of large protein clusters. In the outer membrane of \textit{E. coli}, protein clustering leads to OMP islands, the formation of which underpins membrane protein turnover and drives organization across the cell envelope. By combining CG simulations with in vitro and in vivo experimental studies, it has been possible to suggest how protein–protein interactions enable formation of large clusters of bacterial OMPs which may play a key role in the formation of these membrane “islands”.\textsuperscript{687}

However, a detailed mechanistic understanding of how OMP islands form has been confounded by the difficulties of simulating very large number of OMPs on experimentally addressable time scales. To address this limitation, Chavent, Duncan, and colleagues recently developed a mesoscale model which they trained on large scale CG-MD simulations.\textsuperscript{688} In the meso model, each OMP molecule was represented by a single particle within a 2D membrane model. The meso model was used to run simulations of ca. 5000 copies of an OMP on multimillisecond time scales, thus allowing direct comparison of simulated and in vitro experimental single tracking measurements of OMPs. These studies revealed that specific interaction surfaces between OMPs were the key to formation of OMP clusters and that mesoscale simulations captured the restricted diffusion characteristics of OMPs. This agrees well with recent measurement of the glasslike behavior of crowded membranes.\textsuperscript{689} The OMP clusters in turn presented a mesh of moving barriers that confine newly inserted proteins within membrane islands. Such “corralling” of newly inserted proteins is likely to be of importance for OMPs newly inserted by the BAM machinery. Thus, this type of model enables us to provide a nanoscale molecular mechanism for mesoscale experimental observations. Future refinements of this approach could include using large-scale CG-MD simulations of realistic models of the lipid composition of \textit{E. coli} outer membranes\textsuperscript{258,653} in order to allow meso models to explore the behavior of OMP islands in vivo.

In addition to providing data for parametrization of mesoscale models, very large-scale simulations can provide insights into the emergent behavior of complex and crowded biological membranes which may in the future be included in more biorealistic mesoscale models of complex in vivo membranes. This approach builds upon pioneering work in large scale simulations of crowding of proteins in models of the cytoplasm.\textsuperscript{690,691} These emergent properties include large scale dynamic fluctuations of membranes which may be used to derived mesoscale mechanical parameters of membranes such as the bending rigidity.\textsuperscript{41,109} Application of such analysis to large-scale CG simulations has revealed a complex dependence of the membrane-bending rigidity on both protein contents and lipid composition.\textsuperscript{692} Inclusion of simple models of cytoskeletal tethering of integral membrane proteins also modulates membrane bending rigidity.\textsuperscript{604} Large-scale CG simulations may also be used to explore, for example, the influence of lipid bilayer composition and of specific protein–lipid interactions on patterns and dynamics of membrane protein clustering.\textsuperscript{604} These and other emergent properties from CG simulations will need to be included in a next generation of mesoscale models in order to address the picture emerging from current dynamic experimental measurements which are suggesting cell membranes to be heterogeneous and “scale rich”.\textsuperscript{8}

3.3.3. Membrane Remodeling. In addition to the large-scale dynamic organization of cell membranes “at rest”, large-scale molecular simulations have been used to explore dynamic events including, for example, membrane fusion and remodeling of membranes.\textsuperscript{693,694} In particular, Voth and colleagues have taken a multiscale approach\textsuperscript{226} to explore the biologically important question of how BAR domain proteins interact with lipid bilayers to bring about membrane remodeling.\textsuperscript{695} Using a supra-CG model, they demonstrated how multiple copies of N-BAR domain proteins on a membrane surface form linear aggregates at high protein densities can lead to formation of budlike deformations of the membrane.\textsuperscript{696} Combining CG simulations with microscopy data was used to develop a model of a BAR-domain scaffold, emphasizing the key role of amphipathic helices in the formation of these scaffolds.\textsuperscript{697} These CG studies have in turn fed into mesoscale approaches\textsuperscript{288,295} allowing membrane remodeling to be explored on submicron length and microsecond time scales. Some other reviews covering mesoscale modeling of curvature generation can be found elsewhere.\textsuperscript{267,698}

Figure 11. Developing a mesoscale model for simulation of bacterial outer membrane protein islands. The top panel shows a schematic diagram of an \textit{E. coli} cell, with the areas of outer membrane studied via CG simulation (yellow square), by mesoscale simulation (blue square), and by experimental single molecule tracking (green circle) shown to scale. The lower two panels are snapshots from CG (left) and meso (right) simulations of OMP clustering (see main text and ref 688 for details).
3.3.4. In Silico in Vivo. It is clear that large multiscale simulations of cell membranes can now be used to simulate complex dynamic events in cell and organelle membranes. By integrating such computational approaches with a growing wealth of cryo-EM and optical microscopy data, there is the prospect for future "in silico in vivo" studies of the cell biology of membranes, relating underlying structural and biophysical properties to cellular level events. A number of computational tools will facilitate simulation studies of increasingly complex membrane systems, including tools for semiautomated setup of complex mixed lipid bilayers. On a larger scale, for example, cellPACK provides mesoscale packing algorithms to generate and visualize three-dimensional models of complex biological environments. This has been evaluated on, for example, models of synaptic vesicles and of an HIV virion. Future developments are likely to further integrate a range of tools for setup, running, visualization, and analysis of larger and more complex membrane and cellular systems in addition to development of databases for storage and dissemination of the results of membrane simulations (e.g., MemProtMD and Limonada).

4. OUTLOOK

Thirty years of computer modeling of cell membranes have provided a wealth of information on the lateral organization principles underlying these fascinating quasi two-dimensional systems. From the detailed dynamics of individual lipid tails, via collective processes such as pore formation, protein–lipid sorting, and membrane remodeling, we have now reached a stage where the full complexity of real cell membranes is being captured. Referring back to Figure 1, the obvious question is, what stage comes next?

On the one hand, the quest for more realism has certainly not ended (Figure 12). Detailed models for cell envelopes of most cell types, as well as the many internal organelles, are still very sparse. A key bottleneck is the availability of experimental data concerning their exact lipid composition. Although advanced lipidomics can provide a wealth of data in this respect, it is often not trivial to isolate specific cell fractions. Moreover, any information on membrane asymmetry is lost, hampering realistic modeling efforts. A level up in realism can be achieved by putting cell membrane models into a more realistic environment. Addition of the cytoskeleton would be an obvious example but also the interaction of cell membranes with the crowded environment of the cytoplasm. In fact, not only the membrane leaflets are asymmetric but also the solvent facing both sides, including differences in pH, ionic strength, and electric potential. This has hardly been considered at all in current simulations studies. A major challenge of increased complexity is the increase in required sampling time, as discussed above. Here, there is a need for enhanced sampling algorithms that can deal with crowded and very heterogeneous environments. Data analysis becomes another bottleneck. Whereas waiting for a simulation to complete used to be the bottleneck until some ten years ago, nowadays producing terabytes of data occurs overnight. To make sense of this source of big data, the use of machine learning techniques is promising, but still at its infancy.

On the other hand, there will be a continuing demand for simulations of model membranes containing few components only. Even simple systems can give rise to rich and complex behavior; many of the simulation studies discussed in this review are proof of this. In principle, simplified model bilayers are ideally suited to connect computational and experimental data. From the experimental side, it would be helpful to have more systematic data on some of the basic properties, such as lipid mixing, the effect of membrane curvature, the effect of leaflet asymmetry, effect of ions and pH, as well as the behavior of dyes. Related challenges on the computational side are to provide high throughput data on, for example, multi-component lipid phase behavior and protein sorting and clustering, and to more systematically explore the effect of curvature gradients. Furthermore, efficient constant-pH algorithms need to be developed. Enforcing the connection between experiment and simulation will benefit the ongoing validation of both existing and novel lipid types, and the careful calibration of protein–lipid interactions. Machine learning techniques could also be used to improve the parametrization of FFs; pioneering efforts are already taking place in this direction.

Considering the progress not only in complexity but also in system sizes that can be simulated with particle-based models, it is not too bold to predict that a full cell simulation at near-atomic resolution is feasible within the next ten years. Although such a simulation, featuring many billion atoms, would certainly be very impressive and aid our understanding of how cells are structured at the molecular level, this is by no means the final aim. Real cells, in contrast to equilibrated pieces of cells in a simulation box, are inherently out-of-equilibrium. Incorporating the constant energy flow into nonequilibrium simulations is one of the major challenges for the future.

AUTHOR INFORMATION

Corresponding Author

*E-mail: s.j.marrink@rug.nl.

ORCID

Siewert J. Marrink: 0000-0001-8423-5277
Mark S.P. Sansom: 0000-0001-6360-7959

Notes

The authors declare no competing financial interest.

Biographies

Siewert J. Marrink received his Ph.D. in Chemistry in 1994 from the University of Groningen. His academic career continued with postdoc...
positions at the Max Planck Institute of Tuebingen and the Australian National University. Since 2005, he has been full professor, heading a research group in Molecular Dynamics at the University of Groningen. He also acts as director of the Berendsen Center for Multiscale Modeling at the same university. His main research interest is on multiscale modeling of (bio)molecular processes, with a focus on unraveling the lateral organization principles of cell membranes.

Valentina Corradi holds a Master of Science in Pharmaceutical Chemistry and Technology from the University of Parma and obtained a Ph.D. in Pharmaceutical Sciences from the University of Siena under the supervision of Dr. Maurizio Botta. In 2009, she joined Dr. D. P. Tieleman’s lab at the University of Calgary as a postdoctoral associate, to work in the area of biomolecular simulation. Since 2013, she has worked as a Research Associate in the same group. Her research interests focus on understanding the correlation between structure, function, and dynamics of membrane proteins and their interactions with lipids.

Paulo C.T. Souza is a senior postdoc at the University of Groningen. He received his BSc (2007) and MSc in Chemistry (2009) from the University of Campinas. At the same university, he completed his Ph.D. in Physical-Chemistry (2013) with the guidance of Munir S. Skaf. Since his undergraduate studies, Paulo has dedicated his efforts on the use of molecular dynamic simulations to study proteins as nuclear receptors, cyclin-dependent kinases, and toll-like receptors. Currently, he is in the group of Siewert-Jan Marrink at the University of Groningen, where he works on the development of the MARTINI force field.

Helgi I. Ingólfsson obtained his B.S. degree in Computer Science from the University of Iceland in 2002 and his Ph.D. degree in Computational Biology and Medicine from Weill Cornell Medical College in 2010. He worked as a postdoctoral fellow at the University of Groningen from 2011 to 2016, after which he joined Lawrence Livermore National Laboratory as a Staff Scientist. His current research is primarily focused on the lipid organization of cellular membranes and their perturbation by small molecules, as well as lipid protein interactions and how membrane proteins can be functionally regulated by their lipid environment.

D. Peter Tieleman studied Chemistry and Philosophy at the University of Groningen, where he obtained his Ph.D. under the supervision of Herman Berendsen. After a postdoctoral position with Mark Sansom at Oxford University he moved to the University of Calgary in 2000. He holds the Canada Research Chair in Molecular Lipid Composition. His current research interests include methods for biomolecular simulation and the biophysics of membranes.

Mark S.P. Sansom is the David Philips professor of molecular biophysics in the Biochemistry Department, University of Oxford. He obtained his D.Phil. in Molecular Biophysics from Oxford in 1983. After a postdoctoral fellowship at the University of Nottingham, he was a Lecturer at Nottingham from 1984 to 1991. In 1991, he joined the Biochemistry Department at Oxford. His research interests are in the use of molecular simulations to understand structure, dynamics, and function of membrane proteins. His current interests are focused on protein–lipid interactions and on water in ion channels and nanopores.

ACKNOWLEDGMENTS

Part of this work was performed under the auspices of the U.S. Department of Energy under contract DE-AC52-07NA27344 (LLNL-JRNL-755168). Work in the M.S.P.S. group is supported by the Wellcome Trust and by BBSRC. Work in the S.J.M. group was supported by an ERC Advanced Grant “COMP-MICR-CROW-MEM”. Work in the D.P.T. group was supported by the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research. D.P.T. holds the Alberta Innovates Technology Futures Strategic Chair in (Bio)Molecular Simulation. This work was undertaken, in part, thanks to funding from the Canada Research Chairs program. For help with the figures, we thank Timothy Carpenter (Figure 3), Durba Sengupta (Figure 4), George Khelashvili (Figure 5A), César A. López (Figure 5B), Tao Jiang and Emad Tajkhorshid (Figure 5C), Thorsten Lang (Figure 6), Hector Martinez-Seara (Figure 7), Jonathan Sachs (Figure 8), and Firdaus Samsudin and Syma Khalid (Figure 9).

REFERENCES

(14) Van Der Wel, P. C. A. Lipid Dynamics And Protein–Lipid Interactions In Integral Membrane Proteins: Insights From Solid-State NMR. Emagres 2014, 3, 111−118.
(38) Oloo, E. O.; Tieleman, D. P. Conformational Transitions Induced By The Binding Of MgATP To The Vitamin B12 ATP-binding Cassette (ABC) Transporter BtuCD. J. Biol. Chem. 2004, 279, 45013–45019.


in Giant Unilamellar Vesicles. Poolman, B. Lipid Phase Separation in The Presence of Hydrocarbons


2010, 54, 2683–2687.


