Biochimica et Biophysica Acta xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Validating lipid force fields against experimental data: Progress, challenges and perspectives☆

David Poger [a,*], Bertrand Caron [a], Alan E. Mark [a,b,**]

[a] School of Chemistry & Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia
[b] Institute for Molecular Bioscience, The University of Queensland, Brisbane QLD 4072, Australia

ARTICLE INFO

Article history:
Received 18 November 2015
Received in revised form 7 January 2016
Accepted 27 January 2016
Available online xxxx

Keywords:
Lipid bilayer
Liquid-crystalline phase
Experimental model
Force field
Molecular simulation
Model validation

ABSTRACT

Biological membranes display a great diversity in lipid composition and lateral structure that is crucial in a variety of cellular functions. Simulations of membranes have contributed significantly to the understanding of the properties, functions and behaviour of membranes and membrane–protein assemblies. This success relies on the ability of the force field used to describe lipid–lipid and lipid–environment interactions accurately, reproducibly and realistically. In this review, we present some recent progress in lipid force-field development and validation strategies. In particular, we highlight how a range of properties obtained from various experimental techniques on lipid bilayers and membranes, can be used to assess the quality of a force field. We discuss the limitations and assumptions that are inherent to both computational and experimental approaches and how these can influence the comparison between simulations and experimental data. This article is part of a Special Issue entitled: Membrane Proteins edited by J.C. Gumbart and Sergei Noskov.

1. Introduction

Biological membranes are found throughout all the kingdoms of life and regulate a myriad of critical cellular processes. This is achieved through the modulation of essential properties such as membrane elasticity [1,2], membrane fluidity and the formation of lipid microdomains [3,4,5]. Such differences in turn affect the sensitivity to toxic agents such as antimicrobial peptides [6] and organic pollutants [7]. The principal component of cellular membranes is a lipid bilayer that consists of possibly over a thousand unique lipid species [8,9]. The diversity of lipids can be characterised in terms of structural families (*e.g.* acylglycerols, glycerophospholipids, sphingolipids and sterols), the nature of the headgroups (*e.g.* choline, glycerol and carbohydrates) and the structure of hydrocarbon chains (chain length, (un)saturation, branching groups) [10,11]. A key property of lipid systems is their polymorphism, that is the ability for lipids to aggregate in a range of supramolecular structures. Lipid assemblies can be lamellar (*e.g.* gel phases L_{β} and $L_{\beta'}$, liquid-crystalline phase L_{α} and ripple phase $P_{\beta'}$), spherical (micelles, vesicles—both uni- and multilamellar), cylindrical (hexagonal phases H_{I} and H_{II}) or cubic to name just a few. The formation and the stability of lipid phases depend on a subtle balance between a range of factors, including the lipid composition, the intrinsic properties of individual lipids and the environmental conditions (*e.g.* temperature, osmotic pressure and hydration level) [12]. Amongst the lipids found in cell membranes, phosphatidylcholines assemble into stable bilayers under physiological conditions whereas phosphatidylethanolamines, cardiolipins and diacylglycerols tend to promote non-lamellar structures such as hexagonal and cubic phases [13,14]. However, the behaviour of especially charged lipids such as cardiolipins [15,16] and phosphatidylserines [17,18] is dependent on the concentration of cations (Na^{+}, Ca^{2+}, Mg^{2+}) and variations in pH.

Computer simulations, in particular molecular dynamics (MD) simulations, are a powerful way to examine the properties of membrane systems; in particular to study how specific lipids interact with other lipid species, how they segregate into domains within membranes, how they modulate the structure and the dynamics of membranes, how they associate with proteins and peptides, and how they affect the functions of proteins and peptides. Using MD simulations, it is possible to probe atomic-scale details of the structure and dynamics of lipids and lipid assemblies in any phase. In contrast, the accurate determination of such properties experimentally is difficult. This is especially true for fully hydrated lipids in liquid-crystalline phase which is the most biologically relevant. In fact, while the structure of proteins and protein complexes can be determined with high precision (within a few ångströms of resolution) using experimental techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, determining the structure of a lipid system remains challenging,

☆ This article is part of a Special Issue entitled: Membrane Proteins edited by J.C. Gumbart and Sergei Noskov.

* Corresponding author.
** Correspondence to: A.E. Mark, School of Chemistry & Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia.
E-mail addresses: d.poger@uq.edu.au (D. Poger), a.e.mark@uq.edu.au (A.E. Mark).

http://dx.doi.org/10.1016/j.bbamem.2016.01.029
0005-2736/

even in the case of a pure lipid bilayer composed of a single kind of lipid. This is due to the high fluidity of such systems (*e.g.* high thermal fluctuations), their small thickness, their high water content and the lack of mid- and long-range order [19]. As a result, MD simulations have become the primary means not only to investigate the structure of membranes but also to interpret information from experiment (often X-ray and neutron scattering; NMR, infrared, Raman and fluorescence spectroscopies) [20,21,22,23]. Simulations can also be used to explore the spatial and temporal organisation of lipid bilayers and rearrangements that they undergo during dynamic processes such as membrane fusion, fission, tubulation and poration. However, the predictive value of any MD simulation depends on the accuracy of the underlying model that describes interatomic interactions (or interparticular interactions in the case of coarse-grained simulations) commonly referred to as the force field. The accuracy and the reliability of a lipid force field is determined by its ability to reproduce not only the structure of lipids and lipid phases correctly, but also the individual and collective behaviour of lipids under a given set of conditions: temperature, hydration, compositional heterogeneity (number of different lipid species, presence of protein) and stoichiometry. This is challenging [24,25]. The validation of the lipid force field used is thus critical in determining its capacity to capture specific properties qualitatively and quantitatively and establishing the limitations of the associated simulations. Several well-validated force fields for lipids have been proposed recently. These include AMBER Lipid14 [26], CHARMM36 [27], GROMOS 54A7 [24] and Slipids [28].

Force-field validation is based on the comparison of properties calculated from simulations with reference data obtained experimentally. Reference data can be classified as either direct of indirect. Direct or primary experimental data are experimentally measured properties (*e.g.* X-ray and neutron form factors from scattering experiments, ^{2}H, ^{13}C and ^{31}P relaxation rates and quadrupolar splittings in NMR spectroscopy, and fluorescence intensity in fluorescence spectroscopy and microscopy). Indirect or secondary experimental data are properties that have been inferred from primary experimental data based on a given model (*e.g.* cross-sectional area per lipid, lipid bilayer thicknesses, NMR bond order parameter, and lipid diffusion coefficient). Ideally, force-field validation should be based on the comparison of the simulated data with primary experimental data. This is however not always possible. In the case of fluorescence microscopy experiments, calculating the fluorescence intensities from a classical MD simulation of a lipid bilayer is impossible as fluorescence is a quantum phenomenon. Furthermore, the comparison between simulations and experimental data can be limited because of fundamental differences between the experimental and simulation systems. Experimental systems are macroscopic. Measurements are often based on a stack of bilayers and the data is acquired over minutes or hours. In contrast, a simulation system generally consists of a single bilayer composed of a few hundred to a few thousand lipid molecules that are simulated for a few microseconds at best. It is prudent to treat any comparison between theoretical and experimental data with caution.

In this review, we focus on recent advances in lipid force-field development. Specifically, we show how lipid force fields can be validated using a range of properties obtained from experiment on lipid bilayers and membranes. We have chosen to focus on a limited number of properties that have been measured (or estimated) and which have been used in force-field validation strategies extensively in the literature. These include properties pertaining to the structure of lipid bilayers (namely the scattering form factor of the lipid bilayer, the cross-sectional area per lipid, the bilayer thickness and the lipid chain order) and lipid dynamics in bilayers (lipid lateral diffusion coefficient).

2. Validation of the structure of model membranes

2.1. Scattering form factor

The scattering form factor $F(q_z)$ of a system can be obtained from X-ray and neutron scattering experiments through the analysis of the scattering intensity (q_z is the component of the scattering vector $\mathbf{q}$ along the normal to the lipid bilayer taken as the z-axis). Note, given that lamellar phases have high in-plane thermal disorder, only information on their transverse lipid organisation is accessible from X-ray and neutron scattering measurements. An atomic distribution $\rho(z)$ of all the atoms in a lipid bilayer along z (referred to as a number density profile) can be readily computed from simulations, and converted to a scattering electron density profile in the case of X-ray scattering, and a neutron scattering length density profile in the case of neutron scattering ($\tilde{\rho}(z)$) using:

$$\tilde{\rho}(z) = \sum_{\alpha} f_{\alpha}(q_z)\rho_{\alpha}(z) \tag{1}$$

where $f_\alpha(q_z)$ is the atomic form factor of all atoms of type α of density $\rho_\alpha(z)$. In the case of X-ray scattering, $f_\alpha(q_z)$ can be approximated using tabulated parameters fitted for each atom and ions [29]. For neutron scattering, $f_\alpha(q_z)$ is independent of $\mathbf{q}$ and corresponds instead to the scattering length of the isotope of atom α [30]. $F(q_z)$ is then obtained from the Fourier transformation of $\tilde{\rho}(z)$:

$$|F(q_z)| = \left| \int -d/2\, d/2 \left(\sum_{\alpha} \tilde{\rho}\alpha(z) - \tilde{\rho}s \right) (\cos(zq_z) + i\sin(zq_z))\, dz \right| \tag{2}$$

where $\tilde{\rho}_s$ is the scattering density of the solvent molecules along z and d is the dimension of the simulation box along z (or the repeat distance of a multilamellar stack in an experimental system). Using this procedure (illustrated in Fig. 1), the quality of a force field can be compared in the same reciprocal space as the experimental measurement. This is important as, although a force field may appear to reproduce a range of properties of a lipid bilayer in the real space such as the thickness of the bilayer, significant differences in reciprocal space may be evident. For example, Benz et al. [31] showed that a range of factors are critical for the reproduction of an experimental $|F(q_z)|$ profile; including the value of d at equilibrium (which is related to the hydration level), the frequency with which configurations of the system are sampled ($|F(q_z)|$ should be calculated using statistically independent, meaning uncorrelated, samples), the depth of the midplane trough in the scattering profile, and the distance between headgroups. While the comparison of experimental and simulated scattering form factors have been reported in multiple studies in the literature (for example in Refs. [31,32,33,34,35]), the assessment of the overall agreement is often qualitative, that is whether the profiles overlay to a degree that the authors feel is appropriate. Different authors use different criteria. For example, it has been proposed that the positions of the minima and maxima in the experimental data are the essential features to be captured by a force field [32], whereas other studies have focussed on the position of the minima and whether the relative peak heights in the $|F(q_z)|$ profile from simulation coincide with experimental data [34,35]. Alternatively, the capability of a force field to reproduce the scattering form factor profile of a lipid bilayer has been quantified by calculating a root mean square deviation (RMSD) or a reduced χ^2 value between the simulation and experimental data sets [21,32,23]. However, the threshold within which the two data sets are judged similar remains subjective and its utility limited by experimental uncertainties [31,32].

2.2. Membrane thickness

The most common real-space parameter derived from X-ray and neutron scattering experiments is the bilayer thickness. However, there is no unique definition of the bilayer thickness. One way of determining the thickness of a lipid bilayer is to neglect the interfacial region, that is to divide the repeat spacing d of the stack of lamellae into a water reservoir of thickness d_w that contains only water and a lipid layer of thickness $d_L = d - d_w$ that consists only of lipid (Fig. 2) [36]. In this formalism developed by Luzzati and co-workers, it is necessary to know the lipid concentration and the specific volume of both the lipid and

Please cite this article as: D. Poger, et al., Validating lipid force fields against experimental data: Progress, challenges and perspectives, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbamem.2016.01.029