Biochemical and Biophysical Research Communications xxx (2017) 1–6

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Efficient potential of mean force calculation from multiscale simulations: Solute insertion in a lipid membrane

Roberto Menichetti*, Kurt Kremer, Tristan Bereau

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

ARTICLE INFO

Article history:
Received 16 June 2017
Received in revised form
21 July 2017
Accepted 24 August 2017
Available online xxx

Keywords:
Multiscale simulations
Drug-membrane systems
Potential of mean force
Enhanced sampling

ABSTRACT

The determination of potentials of mean force for solute insertion in a lipid membrane by means of all-atom molecular dynamics simulations is often hampered by sampling issues. Recently, a multiscale method has been proposed to leverage the conformational ensemble of a lower-resolution model as starting point for higher resolution simulations. In this work, we analyze the efficiency of this method by comparing its predictions for propanol insertion into a lipid membrane against conventional atomistic umbrella sampling simulation results. The multiscale approach is confirmed to provide accurate results with a gain of one order of magnitude in computational time. We then investigate the role of the coarse-grained representation. We find that the accuracy of the results is tightly connected to the presence of a good configurational overlap between the coarse-grained and atomistic models—a general requirement when developing multiscale simulation methods.

1. Introduction

Investigating the behavior of compounds embedded in lipid membrane environments is of fundamental importance for biological and pharmaceutical applications, with a variety of interesting physicochemical phenomena occurring in a large interval of time- and length-scales, ranging from microscopic to meso- and macroscopic. Just to mention a few examples, in addition to the traditional analysis of the permeation of a specific compound through the membrane, which is crucial for drug-delivery applications and will be the subject of this work [1], it has been shown that the presence of interfacial proteins or small molecules can affect the curvature [2], the phase diagram [3] and reshape the lipid domains of bilayer systems [4], or even modify the kinetics of ion gating [5].

In probing the thermodynamic partitioning of small molecules in lipid bilayers, a central role is played by the potential of mean force $G(z)$, which describes how the free energy of a compound changes as a function of the normal distance z between its center of mass and the membrane midplane. A qualitative picture of the features of a potential of mean force, together with a set of representative atomistic configurations with the compound at different z is presented in Fig. 1.

Besides providing structural resolution about the equilibrium properties of bilayer insertion [6–8], $G(z)$ is directly linked to the partitioning coefficient, an experimentally measurable quantity that has been systematically analyzed for a large set of compounds [9,10]. Moreover, $G(z)$ can be combined with dynamical properties such as the position dependent diffusion coefficient $D(z)$ in order to gain insights into the compound permeation kinetics, quantified in terms of the permeability coefficient [11,12].

An accurate determination of the potential of mean force—more generally, of a free-energy surface—by means of classical atomistic molecular dynamics (MD) simulations is often hampered by sampling issues [13,14]. These may arise as a consequence of high free-energy barriers separating metastable states along the collective coordinates chosen (the distance z in our case). Many strategies in the form of enhanced-sampling techniques have proven to be very successful in solving this problem, such as umbrella sampling simulations, metadynamics, adaptive biasing force, and generalizations [15–17]. A second source of sampling errors which is more difficult to identify—and consequently overcome—arises from the so-called hidden free-energy barriers [18–21]. Such barriers are embedded into the potentially large set of degrees of freedom

List of abbreviations: DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine; DOPC, 1,2-dioleoyl-sn-glycero-3-phosphocholine; POPC, 1-palmitoyl-2-oleoylphosphatidylcholine; MD, molecular dynamics; AA, all-atom; CG, coarse-grained; MACS, multiscale approach to conformational sampling; M4B/M5B, MARTINI DOPC 4/5-beads models.

* Corresponding author.
E-mail address: menichetti@mpip-mainz.mpg.de (R. Menichetti).

http://dx.doi.org/10.1016/j.bbrc.2017.08.095
0006-291X/

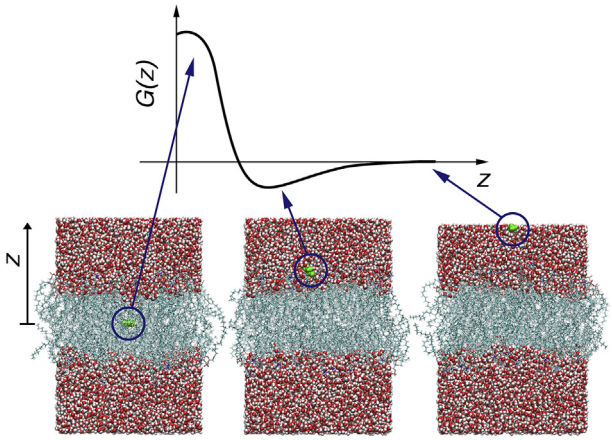

Fig. 1. Qualitative features of a potential of mean force $G(z)$, in relation to the distance of the small molecule (in green) from the bilayer midplane. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

orthogonal to the ones under consideration, and are typically characterized by long relaxation times. Because these variables are not visible along the chosen ones, they can trap the simulation into disguised metastable states, causing a severe undersampling even if enhanced techniques are employed.

These problems would be absent in the ideal case of infinite sampling. In finite MD simulations, they could be clearly mitigated by longer simulations [13] and/or increasing the number of (relevant) collective variables sampled [21]. However, despite the rapid increase in computer power, all-atom (AA) MD simulations are still limited to relatively small systems and short timescales. This limitation does not specifically apply to drug-membrane systems, but characterizes the *in silico* investigation of soft matter as a whole.

One possible way of gaining access to larger simulation time and length scales is to employ coarse-grained (CG) models, which describe the system at a lower level of resolution. In these effective models, subsets of atoms are grouped together into elementary units, or beads, with interactions that are tuned in order to reproduce a set of target properties of the original system, either structural or thermodynamic [22]. Besides reducing the number of degrees of freedom, grouping a set of atoms into an effective interaction site usually corresponds to a smoothing of the rough atomistic energy (and consequently free-energy) landscape, thus reducing some of the difficulties of an adequate conformational sampling.

Among the variety of CG models proposed in the literature, the MARTINI force field [23,24] has proven to be very successful in reproducing the thermodynamic properties of compounds embedded in lipid bilayers [25,26]. MARTINI maps a chemical fragment onto a bead chosen among a set of 18 distinct types, depending on the fragment's water/octanol partition coefficient, hydrogen bond capability and charge. An automated scheme for determining the MARTINI representation of small molecules has been recently proposed in Ref. [26].

The downside of such coarse-grained descriptions is the loss of local, chemical detail, which may have an impact on the properties of drug-membrane systems. More generally, the use of a single level of resolution—either atomistic or coarse-grained—is often insufficient to comprehensively cover the large variety of physical phenomena occurring in soft matter, resulting from a continuous and delicate interplay between different time and length scales.

The necessity of dealing simultaneously with different levels of detail paved the way to the introduction of multiscale simulation approaches, in which several resolutions are concurrently employed in the description of a system, within a single framework. It is beyond the scope of this work to comprehensively discuss such methods, excellent reviews being available in the literature [27,28].

However, it is important to stress that the prerequisite of an accurate multiscale approach is the development of reliable CG models starting from accurate AA ones. Besides capturing the correct large scale behavior of the system, this is fundamental when a reconstruction of the high-resolution detail starting from the low-resolution description is needed, a procedure commonly referred to as backmapping [29–31]. Indeed, one must ensure that the backmapped AA configurations are truly representative of the equilibrium ensemble of the atomistic system. This is far from trivial especially in the case of top-down coarse-grained models as MARTINI, which are often parametrized in terms of little—or no—structural information.

The multiscale approach to conformational sampling (MACS) for determining the potentials of mean force of solute insertion in a membrane was recently proposed in Ref. [32] by Bereau and Kremer. The MACS method aims at accelerating the exploration of the complex atomistic conformational space by means of a systematic backmapping of uncorrelated coarse-grained snapshot onto the atomistic detail, followed by short AA simulations. This procedure partially mitigates the effect of high free-energy barriers that could prevent an adequate sampling when a single, long atomistic simulation is used.

MACS was first applied to determine the atomistic potential of mean force of the protein backbone and side-chain units into a POPC membrane. The convergence properties of the method were then analyzed, including its sensitivity with respect to both the atomistic and coarse-grained models. Moreover, the predictions of the method in the case of the insertion of a WALP16 peptide into a POPC membrane were compared with those obtained by means of long atomistic simulations, finding a reasonable agreement given the complexity of the molecule.

In this work, we extend previous analyses and probe the overall accuracy of the MACS method in assessing the stability of small molecules in a membrane environment, taking propanol as a test case. We first analyze the case of a 1,2-dimyristoyl-sn-glycero-3-phosphorylcholine membrane (DMPC), for which reference atomistic results with conventional umbrella sampling simulations are available. We then investigate the role of two different MARTINI coarse-grained representations of a 1,2-dioleoyl-sn-glycero-3-phosphocholine lipid (DOPC) in the accuracy of the resulting MACS potentials of mean force, providing insights into the delicate relationship occurring between the coarse-grained and atomistic conformational ensembles.

2. Methods

Coarse-grained molecular dynamics simulations were performed in GROMACS [33], and with the MARTINI force field [23–25,34]. The integration time step was $\delta t = 0.02\ \tau$, where τ is the model's natural unit of time, and we relied on the standard force-field parameters reviewed in Ref. [35]. Sampling from the *NPT* ensemble at $P = 1$ bar and $T = 300$ K was achieved by means of a Parrinello-Rahman barostat [36] and a stochastic velocity-rescaling thermostat [37], with coupling constants $\tau_P = 12\ \tau$ and $\tau_T = \tau$. In CG simulations, membranes of $\approx 36\ \mathrm{nm}^2$, containing $N_L = 128$ lipids (64 per layer) were generated by means of the INSANE building tool [38] and subsequently minimized, heated up, and equilibrated. The number of water molecules surrounding the lipids were $N_W = 2150, 1890, 1590$ for DMPC, DOPC M4B and M5B, respectively (see Results section). As usual when using non-polarizable MARTINI water, we added an additional $\approx 10\%$ of

Please cite this article in press as: R. Menichetti, et al., Efficient potential of mean force calculation from multiscale simulations: Solute insertion in a lipid membrane, Biochemical and Biophysical Research Communications (2017), http://dx.doi.org/10.1016/j.bbrc.2017.08.095