Calculating the free energy of transfer of small solutes into a model lipid membrane: Comparison between metadynamics and umbrella sampling

Davide Bochicchio, Emanuele Panizon, Riccardo Ferrando, Luca Monticelli, and Giulia Rossi

Citation: The Journal of Chemical Physics **143**, 144108 (2015); doi: 10.1063/1.4932159
View online: http://dx.doi.org/10.1063/1.4932159
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/143/14?ver=pdfcov
Published by the AIP Publishing

Calculating the free energy of transfer of small solutes into a model lipid membrane: Comparison between metadynamics and umbrella sampling

Davide Bochicchio,[1] Emanuele Panizon,[1] Riccardo Ferrando,[1] Luca Monticelli,[2] and Giulia Rossi[1,a)]

[1]*Physics Department, University of Genoa and CNR-IMEM, Via Dodecaneso 33, 16146 Genoa, Italy*
[2]*Bases Moléculaires et Structurales des Systèmes Infectieux (BMSSI), CNRS UMR 5086, 7 Passage du Vercors, 69007 Lyon, France*

(Received 17 June 2015; accepted 18 September 2015; published online 12 October 2015)

We compare the performance of two well-established computational algorithms for the calculation of free-energy landscapes of biomolecular systems, umbrella sampling and metadynamics. We look at benchmark systems composed of polyethylene and polypropylene oligomers interacting with lipid (phosphatidylcholine) membranes, aiming at the calculation of the oligomer water-membrane free energy of transfer. We model our test systems at two different levels of description, united-atom and coarse-grained. We provide optimized parameters for the two methods at both resolutions. We devote special attention to the analysis of statistical errors in the two different methods and propose a general procedure for the error estimation in metadynamics simulations. Metadynamics and umbrella sampling yield the same estimates for the water-membrane free energy profile, but metadynamics can be more efficient, providing lower statistical uncertainties within the same simulation time. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4932159]

I. INTRODUCTION

In biochemistry and biophysics, the knowledge of free energy landscapes[1] is the key to the comprehension of a plethora of phenomena. Changes in free energy regulate fundamental processes like protein folding,[2] solvation,[3,4] enzymatic reactions,[5] protein-ligand association,[6] and membrane-water partitioning.[7] Computer simulations, and molecular dynamics (MD) in particular, can provide microscopic detail and comprehension of such processes. Nevertheless, despite the growing availability of computational resources and efficient parallel MD software tools, the computational prediction of the free energy landscape of realistic biological systems is still a challenging task.

MD simulations sample the phase space according to the Boltzmann distribution, and the rare events corresponding to the sampling of high-energy states may not occur within the attainable simulation times. Therefore, during the last decades, several computational tools have been developed specifically to address the problem of the calculation of free energy.[8–11] Among the enhanced-sampling methods[12] developed to increase the sampling of all the important regions of free energy surfaces (FESs), we can mention Umbrella Sampling (US),[8] the Wang-Landau algorithm,[9] the adaptive biasing force method,[10] and Metadynamics (MT).[11] Each of these methods has been originally developed having a specific target application, on which its efficiency and limitations have been thoroughly tested. Comparative studies focused on the performance of different computational tools on the same benchmark case can be obviously helpful to guide the users to the choice of the best method.[13]

The membrane permeability and water-membrane partitioning of small chemical compounds and more complex molecular structures are often difficult to measure experimentally,[7] and therefore, computer simulations represent a very useful tool to address this problem. In the last years, simulations have been extensively used for studying the interactions between various kinds of molecules and lipid membranes.[14–19] Most of these studies offer detailed free energy profiles for the transfer of a solute from the water phase to the center of the membrane. The US method has been, by far, the most widely used for this kind of calculation. To our knowledge, MT has been used for similar calculations only in a few cases: to calculate the permeation through the cell membrane of a boron-based β-lactamase inhibitor,[20] to explore the FES of aspirin, diclofenac, and ibuprofen embedded in lipid bilayers,[21] and to get the free energy profile of potassium ions permeating the selectivity filter of the KcsA channel.[22]

In the following, we compare the efficiency of the US and MT methods at calculating the potential of mean force associated to the transfer of hydrophobic oligomers from the water phase to the hydrophobic core of a phospholipid membrane. We will consider polypropylene (PP) and polyethylene (PE) oligomers, modeling them at atomistic and coarse-grained (CG) level. In order to properly compare method efficiencies, adequate estimates of statistical uncertainties are needed. The calculation of the statistical uncertainty for the US method is not trivial, but general strategies have been developed and implemented in the most advanced softwares, such as the weighted histogram analysis method (WHAM).[23] For MT, both theoretical[24] and the empirical[25] equations have been proposed for the calculation of statistical errors, but they are difficult to apply to practical cases. The difficulty in providing error estimates is generally believed to be the main drawback of

a)giulia.rossi@gmail.com

MT, and the information on the statistical uncertainty is often omitted when MT is used. To overcome this issue, here, we propose a general method to calculate statistical uncertainties in MT simulations. We then discuss the similarities and the differences in the statistical uncertainties obtained from standard bootstrap analysis (for US calculations) and from our method (for MT calculations). Our results show that US and MT give virtually identical free energy profiles, and the same statistical uncertainty can be achieved by MT using less computational resources.

II. MODEL AND METHODS

A. Force fields and system setup

Our system consists of oligomers of two common polymers, PE and PP, partitioning into a model membrane made of 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) lipids.

We performed our tests at two different levels of description: the united-atom (UA) level, represented with the UA OPLS force field,[26] and the CG level, represented by the MARTINI[27] force field. In the UA model, hydrogen atoms covalently linked to carbon atoms are not explicitly represented: they are grouped together with the adjacent carbon. The MARTINI force field is a coarser level of approximation. It has been developed by Marrink and co-workers for molecular dynamics simulations of lipids,[27] and later extended to proteins,[28] polymers,[29,30] and other species.[31,32] In MARTINI, a small number of heavy atoms (usually 4) is mapped onto one CG interaction site. The model is parameterized to reproduce the free energy of transfer of each building block between polar and non-polar phases, as well as densities.

We use a small number of POPC lipids in each case: 72 in the atomistic case and 128 in the MARTINI case. POPC membranes are centered along the z axis (normal to the membrane plane) in the simulation box and surrounded by a layer of water of ~2.5 nm on the two sides (as shown in Fig. 1). Periodic boundary conditions are used, and all simulations have been performed in the NPT ensemble at atmospheric pressure and physiological temperature.

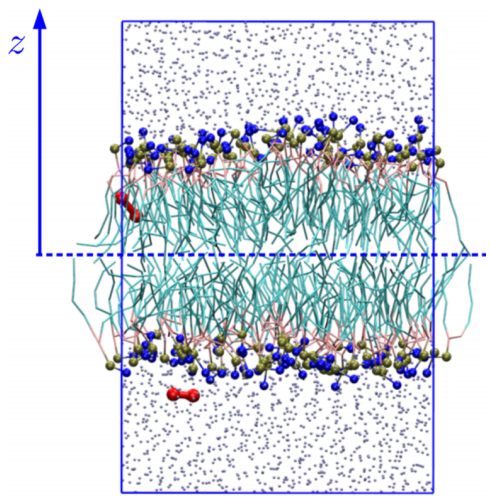

FIG. 1. System setup for the MARTINI US simulations. Two solutes are placed in each US window far apart from each other, and their z coordinate is constrained around a certain q_i value by a harmonic potential.

Our oligomers are (a) an 8-carbon fragment of PE, CH_3–$(CH_2)_6$–CH_3, and (b) a 7-carbon fragment of PP, CH_3–$(CH$–$CH_3)_2$–CH_3. At CG level, we represented these oligomers as dimers, accordingly to the models we recently developed.[33] The topology of the CG dimers does not include any angle or dihedral potentials; thus, the only differences between the two molecules at this level of coarse graining are the equilibrium bond distance between the beads, the strength of the bond, and the type of bead for non-bonded interactions (C_1 for PE and SC_1 for PP, following the MARTINI notation).

B. Umbrella sampling

US is a biased MD method. It attempts at overcoming the sampling problem by biasing the original potential energy surface, so as to enhance the sampling of the unfavourable states. In the following, we will show only the key equations for the case of a single reaction coordinate q. The potential energy is modified as follows:

$$U(\boldsymbol{R}) \rightarrow U(\boldsymbol{R}) + h_i(q), \tag{1}$$

where $\boldsymbol{R}$ is a vector of all coordinates and

$$h_i(q) = \frac{1}{2}k(q - q_i)^2, \tag{2}$$

with k being a harmonic constant and q_i the position along q around which the sampling is enhanced.

A certain number N of simulations is needed, varying q_i in each simulation to sample all the relevant regions of the parameter space. The result is a set of N partially overlapping histograms, each of them providing a probability distribution function $\rho_i(q)$. These probability distributions are biased by the presence of the $h_i(q)$. The free energy, on the other hand, can be written as a function of the unbiased probabilities, which makes necessary to recover the total unbiased probability distribution $\rho^u(q)$ from the biased ones. Here, we use the iterative WHAM,[23] implemented in the GROMACS tool *g_wham*.[34]

Once $\rho^u(q)$ is known, the free energy along q is obtained by

$$F(q) = F(q_0) - \kappa_B T \ln \frac{\rho^u(q)}{\rho^u_0(q_0)}, \tag{3}$$

where q_0 is a reference point. For a complete description of the US method and more details about the unbiasing procedure, the reader is addressed to the original literature.[8,23]

C. Umbrella sampling: Simulation setup

All US runs were performed with GROMACS 4.[35] We have defined our reaction coordinate as the z distance between the center of mass of the oligomer and the center of mass of the lipid membrane (as shown in Fig. 1). We restrained the distance between the oligomer and the center of mass of the membrane, in the direction of the bilayer normal, through the use of harmonic potentials (see Equation (2)). Two molecules of the same kind are put in each simulation box far apart from each other, so that they do not interact (see Fig. 1), allowing to double the sampling at no computational cost.

A different number of windows and a different force constant for the restraining potential are used in UA-OPLS and MARTINI simulations. In the UA case, k is set to 3000 kJ mol^{-1} nm^{-2}, and 35 windows with a width of 0.1 nm are used, thus sampling the distance between the polymers and the center of mass of the membrane between 0 and 3.5 nm. In the CG case, instead, k is set to 1000 kJ mol^{-1} nm^{-2}, and 40 windows with a width of 0.1 nm are used. This increase in the sampled distances with respect to the atomistic case is needed due to the greater thickness of the CG membrane. For both the UA and CG cases, our choice of k parameters leads to optimal overlapping of the US histograms, as shown in Figure S1 of the supplementary material.[42]

D. Umbrella sampling: Error estimation

The GROMACS tool *g_wham*[34] allows an estimation of the error on the recovered free energy profile through a bootstrap analysis. Bootstrapping is a resampling technique which permits to estimate the statistical error of a quantity which is computed from a large set of observations, without repeating the observations multiple times.[36] It works by estimating the probability distribution underlying the observations using a single set of real observations, and then generating fake observations based on that distribution. In our umbrella sampling windows, each position along the reaction coordinate can be considered as a set of independent observations. Alternatively, each complete histogram can be considered as an individual observation. Both options are implemented in *g_wham*, and they are called, respectively, bootstrapping of trajectories (b-traj) and bootstrapping of histograms (b-histo). However, one has to be careful because to be able to use the bootstrapping of histograms without severely underestimating the error, multiple independent simulations for each window would be needed.[34] In our case, we have only two histograms for each window (due to the two images of the dimer in each simulation run); hence, the only feasible choice is the b-traj option.

To bootstrap the trajectories, an estimation of the integrated autocorrelation time (IACT) of the reaction coordinate in each window is needed.[34] We used the GROMACS tool *g_analyze* to estimate the IACTs, with the option -ee, which performs a block averaging:[37] the set of points is divided into blocks and averages are calculated for each block. The error e for the total average is calculated from the variance between averages of all the blocks. An analytical block average curve can also be defined, assuming that the autocorrelation is a sum of two exponentials. Two characteristic times τ_1 and τ_2 are obtained by fitting this curve to e^2. A complete derivation can be found in Ref. 38. In our approximation, the IACT of each window is τ_2 (the longest one). The IACTs calculated in this way (see Figure 2) are then used with *g_wham* to collect independent observations and obtain the free energy profile with its relative error.

E. Metadynamics

In metadynamics, a single biased MD run is performed to derive the free energy landscape of the system under study. Contrary to what happens in US, the bias of metadynamics

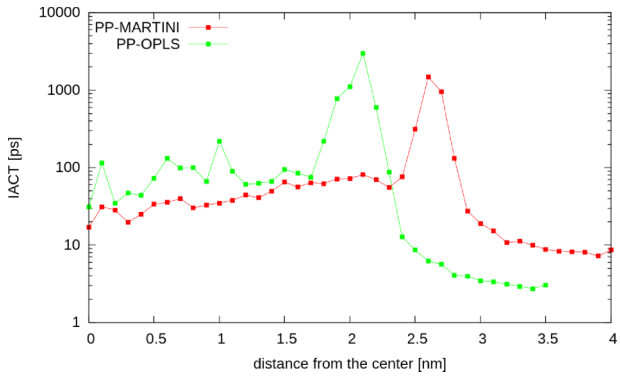

FIG. 2. Integrated autocorrelation times for US simulations. The peak position roughly corresponds to the lipid head region. The different positions of the UA and CG peaks are coherent with the different thickness of the membrane in the two models.

is thought to encourage the exploration of the unfavorable states by kicking the system out of the most favorable ones. Once one or more reaction coordinates are defined, a history dependent bias potential $V_G(t)$ is constructed as a sum of gaussian functions of the selected reaction coordinates. The potential acts at every time step directly on the microscopic coordinates of the system, with the aim to push the system away from already visited configurations. When a single collective coordinate q is used, the potential has the form,

$$V_G(q,t) = w \sum_{t'<t} \exp\left(-\frac{(q-q(t'))^2}{2\sigma^2}\right), \tag{4}$$

where w is the height of each gaussian function, σ its width, and $q(t')$ the position at which the gaussians are centered at the discrete times t' (which are separated by Δt). So for every Δt, the history dependent potential is updated, and while the simulation goes on, it fills the underlying free energy surface helping the crossing of the energy barriers. When the motion of the system along q has become diffusive (in other words, when the free energy profile has been flattened), it means that the simulation has converged, and the free energy profile is given by (minus) the total biasing potential,

$$F(q) \simeq -V_G(q,t). \tag{5}$$

After convergence, the free energy profile should not change anymore, presenting only a global shift due to the continuous process of adding gaussians. However, due to the finite size of the gaussians and the finite value of the diffusion coefficient along q, the profile will always present spurious oscillations around the real FES.[39] This issue could be addressed via the use of the Well-Tempered Metadynamics (WT-MT).[40] In this upgraded version of the MT algorithm, the height of the gaussians is decreased during the MT run, according to the assignment of a further parameter, called bias factor. We thus tried to find an optimal set of parameters to be used with the WT-MT algorithm. Overall, our conclusion is that for our specific case, it is the standard MT algorithm to offer the best performances. More details on the optimization of the WT-MT algorithm can be found in the supplementary material, and all the metadynamics results presented in the following are obtained with the standard MT algorithm.[42]

F. Metadynamics: Simulation setup

GROMACS 4, patched with the PLUMED version 2.1 plugin,[41] is used for MT calculations. A single solute molecule is placed in each simulation box, and the same reaction coordinate as in US is used: the distance along the z axis between the center of mass of the molecule and the center of mass of the lipid membrane. Some important parameters must be set before launching a simulation: the height (w) and the width (σ) of each gaussian, and the time interval at which the gaussians are inserted (Δt). The choice of these parameters is not straightforward, and they influence both the precision and the speed of the calculation.[24,25] The smaller is the ratio $w/\Delta t$ the smaller will be the error, but the corresponding computational cost will be higher. Our final choice is the result of an optimization procedure that is described in detail in the supplementary material.[42] We identified for the MARTINI calculations the values of $w = 0.2$ kJ/mol, $\sigma = 0.12$ nm, and $\Delta t = 1$ ns as a good compromise between speed and accuracy. As for the UA runs, w is the same as in the MARTINI case, but σ is reduced to 0.1 nm due to the smaller z edge of the box.

G. Metadynamics: Convergence and error estimation

The feature that is commonly accepted as an indication of convergence for MT simulations is the diffusive behaviour, i.e., in our case, a situation in which the solute freely moves between the water slab and the membrane without any preferential site. In Fig. 3(a), the z component of the solute trajectory, in a CG run, is shown. Initially, the solute is confined in the lipid tails region (the center of mass of the membrane is set to $z = 0$). After about 2 μs, the biasing gaussians let the solute overcome the energy barrier represented by the polar head region and explore the water phase.

Unfortunately, it is not trivial to identify precisely when the diffusive regime has been reached. A practical method to assess the convergence of MT simulations is based on monitoring the free energy difference between the bulk water region and the center of the membrane. In Fig. 3(b), one can see that in the CG simulations, convergence is achieved after about 3 μs.

Following the example of Ref. 20, we want to calculate the free energy profile as an average of a certain number of different profiles obtained after convergence. Each of the profiles can be seen as the sum of the real free energy profiles and random oscillations. If these profiles are independent, they can be averaged and the error on each point can be assumed to be the standard error of the mean. The problem is that, given the nature of the MT continuous algorithm, two consecutive profiles calculated after convergence will be correlated. Thus, to be able to calculate the standard error of the mean, the decorrelation time (expressed as number of gaussians deposited) has to be identified. We estimate this decorrelation time calculating the time evolution of free energy difference between the bulk water region and the center of the membrane (after convergence), and then doing a block averaging of this quantity. The error coming from the block averaging is initially increasing with the size of the blocks, due to the above mentioned correlations (see Fig. 3(c)). The size of the blocks from which the error stops increasing is our estimation of the decorrelation time. For the CG case, it corresponds to 500 deposited gaussians.

We considered each of the two sides of the bilayer as independent profiles, doubling the statistics. This is somehow the equivalent of using two molecules at a fixed distance in the US simulations. Before averaging the independent profiles, each of them is shifted to have a mean value of zero. Then the result of the average is shifted again setting the free energy minimum at zero (to compare it with the results from US simulations).

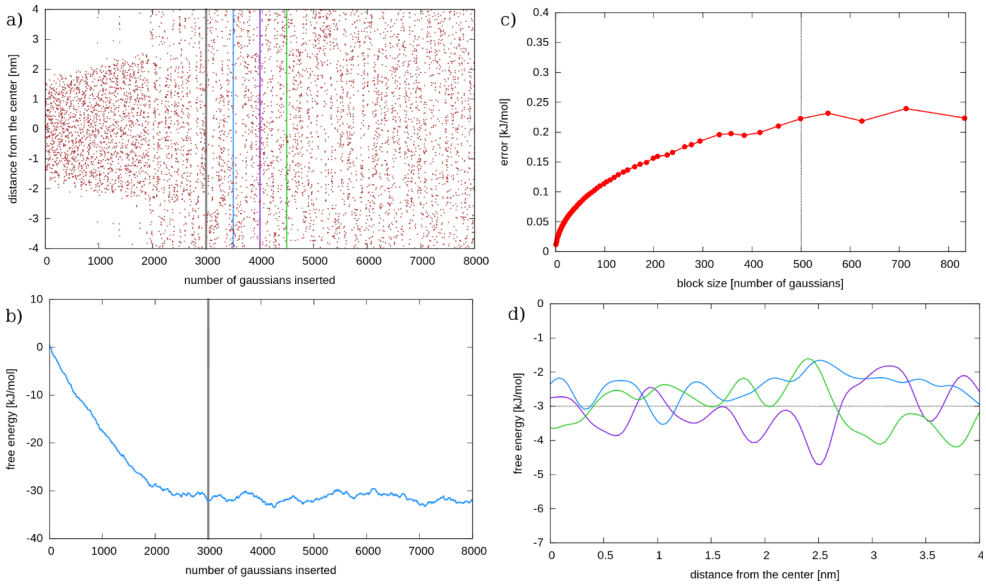

FIG. 3. MT error estimation illustrated in the case of the PP MARTINI dimer. (a) z component of the distance between the solute and the membrane center as more and more gaussians are deposited. Coloured lines indicate the points at which the first three free energy profiles are calculated after the diffusive regime has been reached; (b) free energy difference between water and the center of the membrane. The grey vertical line indicates convergence; (c) block averaging of the free energy difference of panel (b) (after convergence). When the block size exceeds 500 gaussians, the error stops increasing with the block size, indicating decorellated values; (d) free energy deposited between the profiles calculated every 500 gaussians (colours corresponding to panel (a)).

We propose also a simple visual criterion to estimate how much sampling is needed (i.e., how many gaussians have to be added) to obtain independent free energy profiles. If we calculate the free energy deposited between each of the calculated profiles, we find curves that are overall flat, presenting local random oscillations (see Fig. 3(d)). The average free energy deposited (around 3 kJ/mol in the example) is more than the amplitude of the oscillations (around 2 kJ/mol). This means that enough "sand" has been deposited to completely cover the previous hills, obtaining an uncorrelated curve. This happens approximately at the same number of gaussians (500) found with the previous criterion.

III. RESULTS AND DISCUSSION

Let us start comparing the MT and US free energy profiles obtained using the coarse grained MARTINI model. We report in Fig. 4 free energy profiles for PE and PP CG dimers, obtained using a total simulation time of 8 μs with both techniques (the total time in US is the sum of the time duration of all simulation windows). It can be noticed that the energy profiles obtained with the two methods are fully compatible, both for PE and PP. When looking at the error bars, one easily notices that the MT ones are on average smaller. The profile obtained with US looks generally smoother, but this is an intrinsic characteristic of the WHAM, which shifts the free energy profile obtained from each window to make the whole function continuous.[23]

To understand how the statistical uncertainties (errors) behave in the two methods, we discuss the dependence of the error bars from the total simulation time and the position along

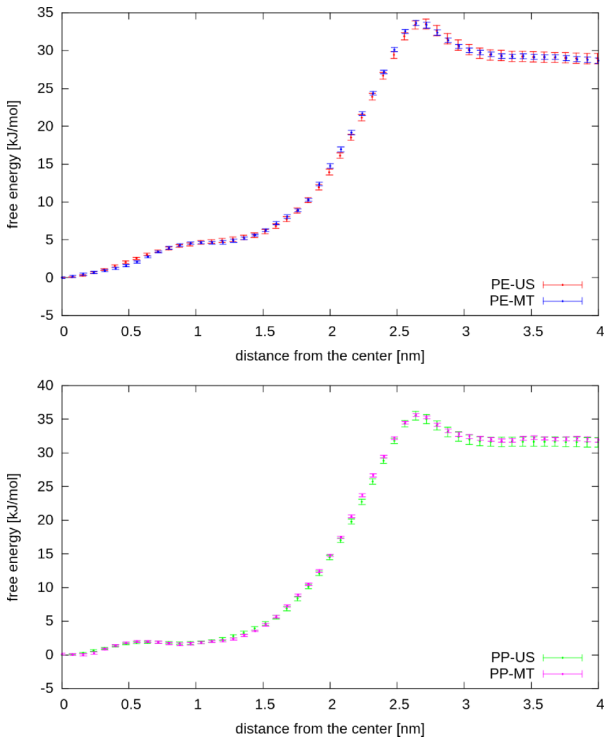

FIG. 4. Free energy profiles obtained with the MARTINI coarse-grained energetic model for PE (top panel) and PP (bottom panel). Here the error bar in each point is the standard error of the mean of the shifted profiles used for the average.

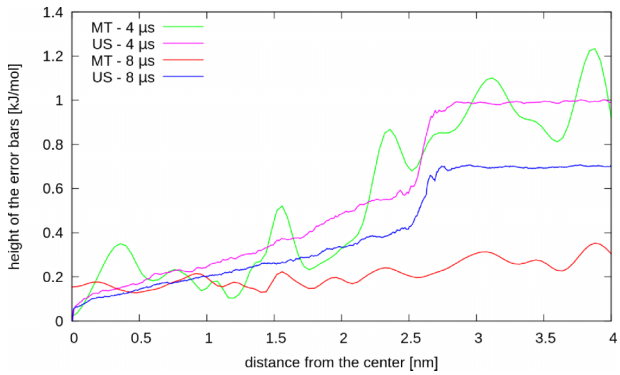

FIG. 5. Dependence of the error bars on the total simulation time and on the position along the reaction coordinate in MT and US, in the case of a PP MARTINI dimer.

the reaction coordinate. The result of this test is presented in Fig. 5 for the case of the PP MARTINI dimer. Since the MT simulation is considered to be converged after 3 μs, a minimum time of 4 μs is required to start estimating the free energy profile. This corresponds to the average of 4 independent profiles (since we consider the two sides of the membrane as independent). The very low statistics is reflected in large standard errors for the MT profile and in the corresponding large fluctuations, as can be seen in Fig. 5. The US error bars, in contrast, present a much smoother behaviour, but their average amplitude is not much different from the MT ones. Increasing the simulation time to 8 μs, the statistics for the MT grows to 20 independent profiles, and the errors of the MT simulation become much smaller than those of US (especially in the water region).

These results indicate that by increasing the total simulation time, the errors in MT, calculated in the way we proposed, decrease much faster than in US. This is true for the set of optimized gaussian parameters reported in Section II. For instance, if one uses $w = 1$ kJ/mol, the MT error bars result bigger than the US ones. However, the time required to roughly fill the free energy well decreases, giving a faster (but less precise) estimation of the free energy difference between water and the membrane interior. Therefore, the appropriate choice of the parameters depends very much on the aim of the calculation: the parameters suggested in this work are suited for obtaining a smooth and precise free energy profile in every region.

The behaviour of the error bars as a function of the reaction coordinate value deserves a separate discussion. For US, in the WHAM method, the value of the free energy in the center of the membrane is set to zero, and thus, it has zero error by default. Then the error increases monotonically going further from the center, with a considerable jump around 2.5 nm. This distance corresponds to the water-membrane interface. The jump in the error is due to the fact that the IACTs are much bigger than in the other regions, as shown in Fig. 2, thus reducing the number of independent samples available to the bootstrapping procedure. At the water-membrane interface, auto-correlation times are in fact more than one order of magnitude bigger than in the inner part of the membrane, while in the water region, they fall to much lower values. This last behaviour is expected, since the mobility in water is much bigger than inside the membrane. Correspondingly, the error of US simulations stops increasing once in the water region.

The MT error is in general more evenly distributed over the whole range of the reaction coordinate. This is especially true when one looks at the curves corresponding to the total simulation time of 8 μs. From this analysis, we can comment that the error in US has some criticality where the system presents strong anisotropy (e.g., the membrane-water interface), hence higher autocorrelation times. The error in MT is less heavily affected by such features (after convergence is reached).

Let us now discuss the results for the case of OPLS-UA simulations. Fig. 6 shows the results obtained with the united-atom model and a total simulation time of 1.8 μs for both US and MT. The time interval between two consecutive gaussian depositions in the MT runs has been set ten times smaller (Δt = 0.1 ns) than in the CG case, to compensate for the reduced MD time step used.

In this case, the free energy difference between water and the center of the membrane shows large fluctuations and correlation times are much longer. The block averaging analysis does not produce clear results as in the case of MARTINI simulations. Thus, we decided to average the profiles without caring for the amount of correlation, and then consider the standard deviation as an upper limit for the error. In Fig. 6, the error bars reported are the standard deviations.

We note that the MT simulations results are once again comparable with the US results, and that the average uncertainty is similar, despite the MT errors are certainly overestimated. In fact, error bars are bigger for MT than for US in the inner parts of the membrane; they are similar in the central part of the profiles and smaller in the water region.

A comparison between Figs. 4 and 6 shows that the errors for both techniques are significantly larger in OPLS-UA than in MARTINI simulations. This is a direct consequence of the

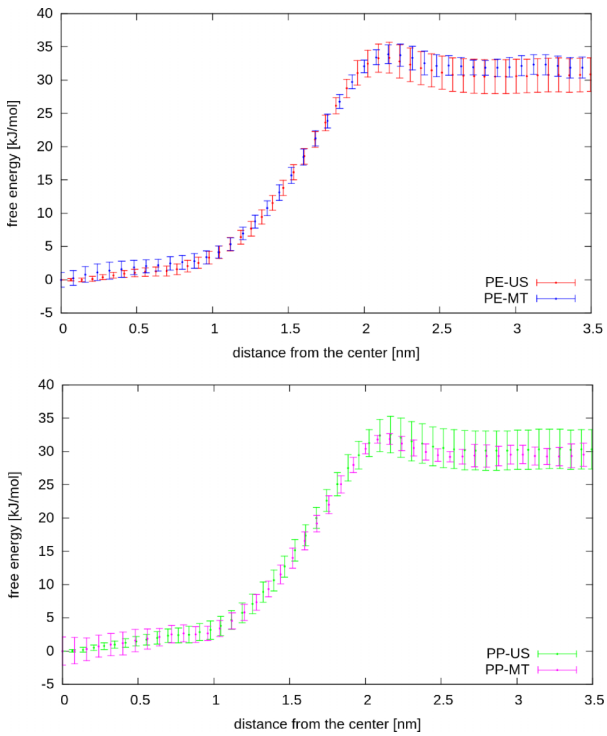

FIG. 6. Free energy profiles obtained with the OPLS united atoms energetic model for PE (top panel) and PP (bottom panel). Here the error bar in each point is the standard deviation of the shifted profiles used for the average.

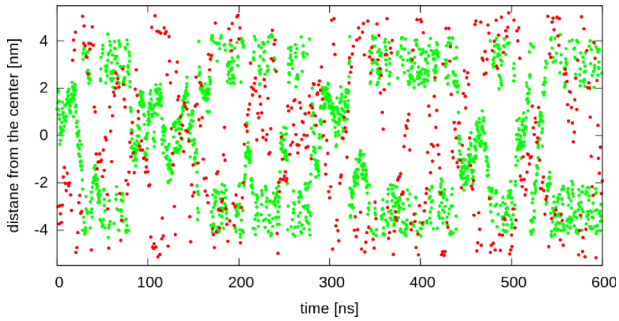

FIG. 7. The solute reaction coordinate in the last 600 ns of a MARTINI simulation (red) and of a UA-OPLS simulation (green) is shown. In both cases, convergence has been achieved and the reaction coordinate is in the diffusive regime. The slower dynamics in UA-OPLS leads to less frequent crossing of the water-membrane interface in the same time intervals.

total simulation time used, which is much smaller for UA-OPLS simulations. The use of a smaller total simulation time is imposed by the bigger computational effort required for simulating a system in which all heavy atoms are explicitly taken into account. In US, the bootstrapping involves less independent points because of the shorter trajectories but also because the IACTs are almost everywhere higher (see Fig. 2).

In MT, a key parameter governing the error is the rate at which the gaussians are inserted. The need to use a smaller total time forces us to use a higher rate of insertion, and thus leads to less precise MT calculations. There is also another problem: the dynamics itself is slowed down with respect to the MARTINI case, and, as a consequence, crossing the water-membrane interface is less frequent. Thus, the solute remains stuck in the water region or in the membrane region for relatively long times, even after convergence (see Fig. 7). Increasing the total simulation time for MT OPLS-UA simulations would allow a more rigorous estimate of the statistical uncertainties (as used in the MARTINI simulations), leading to smaller uncertainties. Overall, our results suggest that the performance of MT is superior to US even in the OPLS-US case, if the choice of the simulation parameters is appropriate.

IV. CONCLUSIONS

In the present work, we have compared the efficiency of two common methods for free energy calculations, US and MT. We applied such methods to the calculation of the partitioning of small molecules, namely, PE and PP dimers, between water and a POPC lipid membrane. Calculations were performed at two different levels of description, united atoms (UA-OPLS) and coarse-grained (MARTINI). Our results indicate that both methods can produce reliable free energy profiles with a reasonable computational effort. The US technique, that is the most commonly used for this type of calculations, is straightforward for what concerns the calculation of the profile through the WHAM method. The calculation of the error is rather straightforward too, once the integrated correlation times for each window have been estimated, and consist in a bootstrapping of trajectories as already implemented in the *g_wham* GROMACS tool. On the contrary, MT requires the proper setting of some key parameters to really shine. In particular, a proper choice of the height and the width of the

gaussians, as well as their rate of insertion (which should be a compromise between precision and time consumption), has to be made. Here, we have proposed good simulation parameters for MT, as well as a general methodology for the calculation of the statistical uncertainties, which makes use of a single MT run. We could not apply the same rigorous procedure for the calculation of uncertainties in the atomistic case, due to the significantly higher computational cost required to reach convergence in atomistic simulations. Overall, the MT method proved to be able to recover the free energy profiles with a smaller uncertainty within the same total simulation time (or alternatively with the same error but with a smaller total simulation time). The present work offers useful guidelines for the efficient use of MT for the calculation of the water-membrane free energy of transfer.

ACKNOWLEDGMENTS

Giulia Rossi acknowledges funding from the FP7 Marie Curie Career Integration Grant No. PCIG13-GA-2013-618560. Part of the calculations were performed at CINECA under the NANOPLAS (No. HP10CM4EYY) grant.

[1]P. Kollman, "Free energy calculations: Applications to chemical and biochemical phenomena," Chem. Rev. **93**, 2395–2417 (1993).
[2]D. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, UK, New York, 2003).
[3]B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff, and B. Honig, "New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects," J. Phys. Chem. **100**(28), 11775–11788 (1996).
[4]D. L. Mobley, C. I. Bayly, M. D. Cooper, M. R. Shirts, and K. A. Dill, "Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations," J. Chem. Theory Comput. **5**, 350–358 (2009).
[5]A. Warshel, "Calculations of enzymic reactions: Calculations of pka, proton transfer reactions, and general acid catalysis reactions in enzymes," Biochemistry **20**, 3167–3177 (1981).
[6]V. Limongelli, L. Marinelli, S. Cosconati, C. La Motta, S. Sartini, L. Mugnaini, F. Da Settimo, E. Novellino, and M. Parrinello, "Sampling protein motion and solvent effect during ligand binding," Proc. Natl. Acad. Sci. U. S. A. **109**, 1467–1472 (2012).
[7]R. V. Swift and R. E. Amaro, "Back to the future: Can physical models of passive membrane permeability help reduce drug candidate attrition and move us beyond qspr?," Chem. Biol. Drug Des. **81**, 61–71 (2013).
[8]G. Torrie and J. Valleau, "Nonphysical sampling distributions in Monte Carlo free energy estimation: Umbrella sampling," J. Comput. Phys. **23**, 187–199 (1977).
[9]F. Wang and D. P. Landau, "Efficient, multiple-range random walk algorithm to calculate the density of states," Phys. Rev. Lett. **86**, 2050–2053 (2001).
[10]E. Darve and A. Pohorille, "Calculating free energies using average force," J. Chem. Phys. **115**, 9169–9183 (2001).
[11]A. Laio and M. Parrinello, "Escaping free-energy minima," Proc. Natl. Acad. Sci. U. S. A. **99**, 12562–12566 (2002).
[12]L. Monticelli, *Biomolecular Simulations Methods and Protocols* (Humana Press Springer, New York, 2013).
[13]J. C. Gumbart, B. Roux, and C. Chipot, "Standard binding free energies from computer simulations: What is the best strategy?," J. Chem. Theory Comput. **9**, 794–802 (2013).
[14]H. D. Herce and A. E. Garcia, "Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes," Proc. Natl. Acad. Sci. U. S. A. **104**, 20805–20810 (2006).
[15]L. M. Lichtenberger, Y. Zhou, V. Jayaraman, J. R. Doyen, R. G. O'Neil, E. J. Dial, D. E. Volk, D. G. Gorenstein, M. B. Boggara, and R. Krishnamoorti, "Insight into NSAID-induced membrane alterations, pathogenesis and therapeutics: Characterization of interaction of NSAIDs with phosphatidylcholine," Biochim. Biophys. Acta **1821**, 994–1002 (2012).
[16]M. B. Boggara, M. Mihailescu, and R. Krishnamoorti, "Association of nonsteroidal anti-inflammatory drugs with lipid membranes," J. Am. Chem. Soc. **134**, 19669–19676 (2012).
[17]G. Rossi, J. Barnoud, and L. Monticelli, "Partitioning and solubility of c-60 fullerene in lipid membranes," Phys. Scr. **87**(5), 058503 (2013).
[18]J. Barnoud, G. Rossi, and L. Monticelli, "Lipid membranes as solvents for carbon nanoparticles," Phys. Rev. Lett. **112**(6), 068102 (2014).
[19]G. Rossi, J. Barnoud, and L. Monticelli, "Polystyrene nanoparticles perturb lipid membranes," J. Phys. Chem. Lett. **5**(1), 241–246 (2014).
[20]M. Minozzi, G. Lattanzi, R. Benz, M. P. Costi, A. Venturelli, and P. Carloni, "Permeation through the cell membrane of a boron-based beta-lactamase inhibitor," PLoS One **6**(8), e23187 (2011).
[21]J. P. M. Jambeck and A. P. Lyubartsev, "Exploring the free energy landscape of solutes embedded in lipid bilayers," Phys. Chem. Lett. **4**, 1781–1787 (2013).
[22]E. Piccinini, M. Ceccarelli, F. Affinito, R. Brunetti, and C. Jacoboni, "Biased molecular simulations for free-energy mapping: A comparison on the KcsA channel as a test case," J. Chem. Theory Comput. **4**, 173–183 (2008).
[23]S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenbergl, "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," J. Comput. Chem. **13**(8), 1011 (1992).
[24]A. Laio, A. Fortea-Rodriguez, F. Gervasio, M. Ceccarelli, and M. Parrinello, "Assessing the accuracy of metadynamics," J. Phys. Chem. B **109**, 6714–6721 (2005).
[25]M. Parrinello, G. Bussi, and A. Laio, "Equilibrium free energies from non equilibrium metadynamics," Phys. Rev. Lett. **96**, 090601 (2006).
[26]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids," J. Am. Chem. Soc. **118**, 11225–11236 (1996).
[27]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The martini force field: Coarse grained model for biomolecular simulations," J. Phys. Chem. B **111**, 7812–7824 (2007).
[28]L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, "The martini coarse-grained force field: Extension to proteins," J. Chem. Theory Comput. **4**, 819–834 (2008).
[29]H. Lee, A. H. de Vries, S. J. Marrink, and R. W. Pastor, "A coarse-grained model for polyethylene oxide and polyethylene glycol: Conformation and hydrodynamics," J. Phys. Chem. B **113**(40), 13186 (2009).
[30]G. Rossi, L. Monticelli, S. R. Puisto, I. Vattulainen, and T. Ala-Nissila, "Coarse-graining polymers with the martini force-field: Polystyrene as a benchmark case," Soft Matter **7**, 698–708 (2011).
[31]C. A. Lopez, A. Rzepiela, A. H. de Vries, L. Dijkhuizen, P. H. Hunenberger, and S. J. Marrink, "The martini coarse-grained force field: Extension to carbohydrates," J. Chem. Theory Comput. **5**, 3195–3210 (2009).
[32]L. Monticelli, "On atomistic and coarse-grained models for c60 fullerene," J. Chem. Theory Comput. **8**(4), 1370 (2012).
[33]E. Panizon, D. Bochicchio, L. Monticelli, and G. Rossi, "MARTINI coarse-grained models of polyethylene and polypropylene," J. Phys. Chem. B **119**, 8209 (2015).
[34]J. S. Hub, B. L. de Groot, and D. van der Spoel, "g-wham: A free weighted histogram analysis implementation including robust error and autocorrelation estimates," J. Chem. Theory Comput. **6**, 3713–3720 (2010).
[35]B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," J. Chem. Theory Comput. **4**(3), 435–447 (2008).
[36]B. Efron, "Bootstrap methods: Another look at the jackknife," Ann. Stat. **7**, 1–26 (1979).
[37]H. Flyvbjerg and H. G. Petersen, "Error estimates on averages of correlated data," J. Chem. Phys. **91**, 461 (1989).
[38]B. Hess, "Determining the shear viscosity of model liquids from molecular dynamics simulations," J. Chem. Phys. **116**(1), 209–217 (2002).
[39]A. Barducci, M. Bonomi, and M. Parrinello, "Metadynamics," Comput. Mol. Sci. **1**(5), 826–843 (2011).
[40]A. Barducci, G. Bussi, and M. Parrinello, "Well-tempered metadynamics: A smoothly converging and tunable free-energy method," Phys. Rev. Lett. **100**, 020603 (2008).
[41]M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, "Plumed: A portable plugin for free-energy calculations with molecular dynamics," Comput. Phys. Commun. **80**, 1961–1972 (2009).
[42]See supplementary material at http://dx.doi.org/10.1063/1.4932159 for a detailed description of the US and MT parameter optimization procedure (Figures S1 and S2) and for a comparison between the performances of the standard MT algorithm vs. the well-tempered MT (Figure S3).