

Comparison of Umbrella Sampling and Steered Molecular Dynamics Methods for Computing Free Energy Profiles of Toluene Molecules through Phospholipid Bilayers

Sang Young Noh^{1, a)} and Rebecca Notman^{1, b)}
Department of Chemistry, University of Warwick

(Dated: 24 September 2019)

Understanding the permeation of molecules through lipid membranes is fundamental to understanding and predicting the cellular uptake of solutes and drug delivery mechanisms. In molecular simulations the usual approach is to compute the free energy (FE) profile of a molecule across a model lipid bilayer, which can then be used to estimate the permeability of the molecule. Umbrella sampling (US), which involves carrying out a series of biased simulations along a defined reaction coordinate (usually the bilayer normal direction), is a popular method for the computation of such FE profiles. However, US can be challenging to implement because the results are dependent on the strength of the biasing potential and the spacing of windows along the reaction coordinate, which, in practice, are usually optimized by an inefficient trial and error approach. The steered molecular dynamics implementation of the Jarzynski Equality (JE-SMD) has been identified as an alternative to equilibrium sampling methods for measuring the FE change across a reaction coordinate. In the JE-SMD approach, equilibrium FE values are evaluated from the average of rapid non-equilibrium trajectories, thus avoiding the practical issues that come with US. Here, we use three different corrections of the JE-SMD method to calculate the FE change for the translocation of a toluene molecule across a lipid bilayer, and compare the accuracy and computational efficiency of these approaches to the results obtained using US. We show that when computing the free energy profile of a small molecule across a model membrane, the JE-SMD approach suffers from sampling issues that may be alleviated through the use of a slower pulling velocity, but at the cost of computational efficiency. We deduce that, despite its drawbacks, US remains the more viable approach of the two for computing the FE profile.

I. INTRODUCTION

The translocation of molecules through lipid membranes is a subject of much interest, due to the cell membrane being the main barrier for nutrient or drug molecules to enter a cell. Depending on the physiochemical properties of the permeating molecule and the composition of lipids and proteins in the membrane, transport may be active (requiring energy expenditure by the cell) or passive, where the permeant diffuses along its concentration gradient^{1,2}. Here we consider passive membrane transport. Key factors that affect the passive permeability include the lipid composition of the membrane (which includes the chemical nature of the lipid head and tail groups, the packing of the lipids and thus the free area and/or volume and the fluidity of the membrane) and the size, shape and hydrophilicity/phobicity of the permeant³⁻⁶. It is important to be able to determine the ease by which molecules permeate membranes in order to predict the cellular uptake of solutes, the bioavailability of drug molecules and the effectiveness of drug delivery vehicles at getting encapsulated molecules into their target cells. The use of molecular simulation to rationalise and predict the transport properties of molecules and nanomaterials across cell membranes remains an field of

significant growth⁷⁻¹⁰. In these simulations the cell membrane is typically represented by a lipid bilayer with one or more lipid components. The majority of these studies aim to compute the potential of mean force (PMF) (or free energy (FE) profile) along the permeation pathway, from which one can determine the free energy of transfer from the water to the membrane interior. If the local diffusivity of the permeant along the permeation pathway is also calculated, the permeability coefficient can be estimated using the inhomogeneous solubility-diffusion model¹¹.

In molecular simulations, a number of enhanced sampling methods are routinely used to compute the FE profile along a reaction coordinate. In particular, the *umbrella sampling* (US) method has been ubiquitous in its use, due to its intuitive nature and relative simplicity of implementation¹²⁻¹⁴. In the US method, the reaction coordinate is divided into windows, and a series of independent simulations are carried out where the system is restrained by a Hook's potential to sample within each window. The resulting histograms of the sampled values of the reaction coordinate in each simulation are then recombined to reconstruct the unbiased FE profile. The weakness of the US method comes from its relatively high computational cost - the high energy parts of the phase space need to be sampled adequately and there needs to be adequate overlap between each sampling window. Thus, in practice, implementation of US usually requires a time-consuming trial and error process to identify the most suitable spacing of simulation

^{a)} Electronic mail: s.y.noh@warwick.ac.uk

^{b)} Electronic mail: r.notman@warwick.ac.uk

windows and associated force constants for the biasing potential. As an alternative way to compute the free energy profile, Jarzynski¹⁵ demonstrated a revolutionary equality (known as *Jarzynski's equality* (JE)), which showed that, independent of the velocity of the process (i.e. it can be a non-equilibrium process), the force-distance curve along the reaction coordinate can be used to compute the work, a path-dependent function, to compute the FE of the process, an *equilibrium* state function.

The computation of the FE using JE-based methods introduced so-called *fast-switching* events (a series of simulations where the system is pulled along the reaction coordinate at a relatively high constant velocity) as a possible efficient alternative to the US method. Should the validity of the equality be widely applicable, the advantages of JE-based methods compared to US are such that one would not need to sample discrete bins along the reaction coordinate, which would overcome the challenges associated with US and potentially reduce the overall computational resources required. In principle this would enable the exploration of more complex reaction coordinates, as it would not be an explicit requirement for the system to equilibrate before extracting the FE value. Schulten *et al.* suggested the *steered molecular dynamics* (SMD) method¹⁶ whereby a moving Hook's potential drives the system along the reaction coordinate, which enables the computation of the the work distribution that is required for the calculation of the FE. Hence, we abbreviate the practical implementation of JE as JE-SMD.

Previous studies comparing the US and JE-SMD methods have established that the JE-SMD can give comparable results to US for an accurate estimate of the PMF, depending on the velocity chosen for the moving harmonic potential¹⁶. However, there is also evidence to suggest that sampling issues arise when applying JE-SMD to complex many body systems¹⁷⁻¹⁹. Verification of the JE involve relatively simple processes, such as the extension of a DNA residue¹⁶ and the unfolding of RNA hairpins²⁰, which does not therefore, dismiss the possibility of error in the JE-SMD method when applied to many body systems. The trial system which was used to verify the JE method used a Langevin dynamics parameter to mimic thermal fluctuations, and its small size allowed the repeat simulations (of the order 10,000 runs and more) to be carried out with ease. This would not be feasible for the vast majority of complex systems, where the simulation costs relating to the level of atomic detail and system size are the limiting factors for making a comparable number of simulations. The application of the JE-SMD to larger systems (of 10 - 100 nm scale) have shown mixed results; Kuyucak *et al.*^{18,19} investigated the flow of water molecules across a carbon nanotube and a gramicidin A channel in a bilayer and showed that FE profiles computed using JE-SMD showed a large discrepancy compared to the benchmark US results. This was suggested to be due to the slow relaxation of the surrounding system, which is dependent on the pulling velocity. Thus, while the JE-SMD method shows considerable promise as

a FE sampling method, further optimization is required for its application to simulations of realistic systems. A recent review by Pienko and Trylska²¹ illustrated that SMD was a good method for identifying translocation pathways, but had outstanding issues to use as a reliable FE computation. In particular, it was noted that the method deviates from an accurate prediction when the reaction coordinate goes far from linearity. It was also noted that when the pulling force from the JE-SMD is large, it would distort flexible polymers and peptide structures during permeation processes, which would prevent the sampling of the true reaction pathway.

In this article, we compute the FE profile of a toluene molecule translocating through a model 1,2-Dioleoyl-sn-glycero-3-phosphocholine (DOPC) lipid bilayer using the US and the JE-SMD methods. We use the US profile as the benchmark for the true FE profile, and examine the effects of three different interpretations of JE-SMD that have been suggested in the literature and the effect of the pulling velocity. We ascertain the degree to which these modifications to the original JE-SMD method may improve the computation of the true FE profile. We look at the convergence behaviour of the JE-SMD methods towards the US result with the assumption that the FE profile from the US method represents the correct FE profile. In addition we analyse the computational efficiency of the various approaches.

II. COMPUTATIONAL METHODS

A. Simulated Systems

The system contained 128 coarse-grained (CG) DOPC molecules in a bilayer configuration. A total of 4232 water beads in total were placed above and below the leaflets. A single toluene molecule (at atomistic resolution, see below) was placed above the bilayer. Figure 1 shows the schematic of the CG-DOPC molecule and the molecular structure of toluene used in these simulations.

B. Forcefield Parameters

The *ELBA* biomolecule *mixed resolution* coarse-grained and all-atomic (CG-AA) compatible FF was used to model the DOPC bilayer^{22,23}. The ELBA-CG-AA hybrid model offers an intriguing possibility for new insights, with AA-level detail coupled with CG-level simulation scales. The CG-AA compatibility of ELBA allows the CG potentials of the lipid molecules to be used in combination with the AA interatomic potentials of the General Amber FF (GAFF)²⁴, which gives it wide applicability to known libraries of small molecules. Hence, the GAFF FF was used to model the toluene molecule. The details of the ELBA FF have been included in the supplementary information.

C. Simulation Parameters

Each simulation was run using the *LAMMPS* molecular dynamics package²⁵. Each bilayer system was initially run in the *NVT* ensemble with the Langevin thermostat for 100 ns at temperatures of 298 K. A production simulation run with the Berendsen barostat and Langevin thermostat (i.e. a *NPT* equivalent) was then carried out for 100 ns to ensure that the system had equilibrated. The temperature and pressure of the *NPT* simulation was set to 298 K and 1 atm respectively. The pair-wise interactions were a shifted-force type with a 1.2 nm cut-off. The reversible-Reference System Propagator Algorithm (rRESPA) was used to divide the total timestep to compute the AA and CG components on a different timescales²⁶. Outer level timesteps updated the CG-CG interactions (i.e. between the ELBA-beads) every 4.0 fs while the AA-CG and AA-AA interaction were updated in the inner level every 1.0 fs. The LAMMPS input file is included in the supplementary information (Listing 1).

D. Free Energy Calculations

1. Umbrella Sampling

For the US simulations, the reaction coordinate (the z coordinate along the bilayer normal) was divided into 1 Å intervals from the center of the bilayer (0 Å) to the bulk water above the headgroup region (30 Å). The bias potentials of each US window were set to 2.5 kcal mol⁻¹, increased to 5.0 kcal mol⁻¹ for 2 Å above and below the headgroup region (this corresponds to a FE barrier region and follows closely the US procedure followed by Genheden *et al.*²⁷). In total, 30 US windows were simulated with an initial equilibration run of 10 ns, and a production run of 30 ns used to construct the free energy profile using the weighted histogram analysis method (WHAM)²⁸. Further detail on the WHAM procedure is included in the supplementary information.

2. Jarzynski Equality

The Jarzynski equality (JE) is defined as:

$$e^{-\beta \Delta G(z)} = \langle e^{-\beta W(z)} \rangle \quad (1)$$

where β is $\frac{1}{k_B T}$, where k_B is the Boltzmann constant, T is the temperature, $\Delta G(z)$ the FE at the given z coordinate, and W is the work¹⁵. What is intriguing with this equality is that W is a path-dependent property, whereas ΔG is a state function - an equilibrium property. Hence, this value equates a non-equilibrium property to a equilibrium property. Issues regarding the practical use of the JE are the inherent bias related to insufficient sampling and the dissipated work. A direct interpretation of

the JE to map a FE profile can be interpreted from the following equation:

$$\Delta G_J = \frac{1}{\beta} \ln \left[\frac{1}{N} \sum_i^N e^{\beta W_i(z)} \right] \quad (2)$$

Where N is the total number of work samples. The use of the JE for FE calculations may suffer from a significant amount of bias (i.e. the difference between the expected value of the FE and its estimate, $\Delta G - \Delta G_J$) due to insufficient sampling along the reaction coordinate. Also, the exponential average value $\left(\frac{1}{N} \sum_i^N e^{\beta W_i(z)} \right)$ is dominated by rarely occurring small work values. The second order cumulant expansion term for the JE was used to correct the bias due to this sampling problem¹⁶:

$$\Delta G_{\text{cumulant}} = \langle W(z) \rangle - \frac{N}{N-1} \frac{\beta}{2} \left(\langle W(z)^2 \rangle - \langle W(z) \rangle^2 \right) \quad (3)$$

Where the $\left(\langle W(z)^2 \rangle - \langle W(z) \rangle^2 \right)$ term is the variance of the work along the reaction coordinate (where the angular brackets represent the averaged value over N trajectories), and the $\langle W(z) \rangle$ term is the averaged work. This modified JE term for the FE difference is valid on the condition that the work distribution along the reaction coordinate is Gaussian, as this enables the elimination of cumulants higher than that of second order of the JE to equal 0. When using this estimator to compute the FE, another factor to consider is the distribution of the dissipation work, where the dissipated work is defined as the difference between the averaged work and the true FE difference at the reaction coordinate; the magnitude of the dissipated work dictates the width of the Gaussian probability distribution. The dissipation work may therefore be estimated by:

$$W(z)_{\text{dissipation}} = \frac{1}{2} \beta \sigma_{W(z)}^2 \quad (4)$$

Where $\sigma_{W(z)}^2$ represents the variance of the work. We note that this term is identical to the factor in variance of the work in the term $\frac{\beta}{2} \left(\langle W(z)^2 \rangle - \langle W(z) \rangle^2 \right)$, which shows that the variance of the work along the reaction coordinate is equal to the $W(z)_{\text{dissipation}}$. The probability of observing a trajectory with negative dissipation work can be described as²⁹:

$$P \left(W(z)_{\text{dissipation}} < 0 \right) = \frac{1}{2} \left[1 - \text{erf} \left(\sqrt{\langle W(z)_{\text{dissipation}} \rangle / 2} \right) \right] \quad (5)$$

Where $\langle W(z)_{\text{dissipation}} \rangle$ represents the averaged dissipation work along the trajectory. The equation implies that with a larger $W_{\text{dissipation}}$, the probability of observing a negative dissipation work sharply decreases. Hence, an increased magnitude of dissipated work corresponds to a lower probability of observing negative work events, and is undesired. An alternate method for taking into ac-

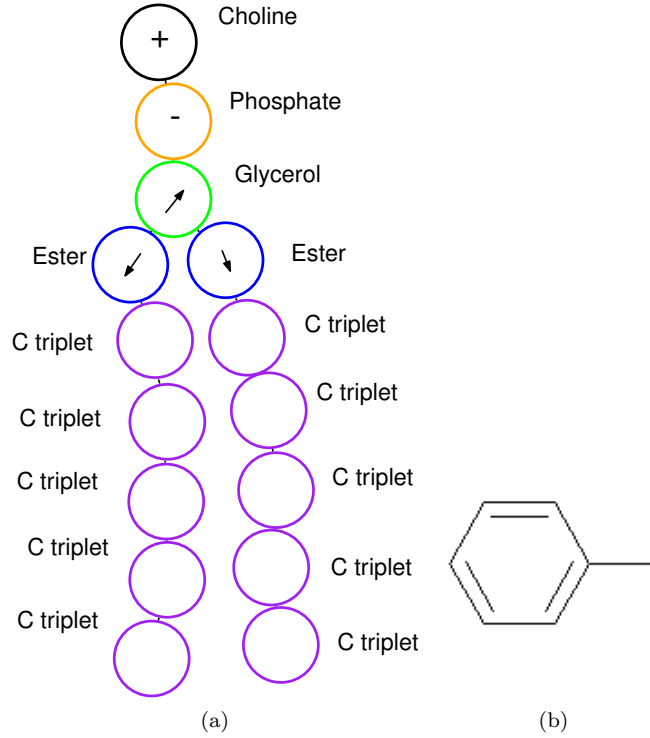


FIG. 1: (a) a schematic of the CG DOPC molecule, where the headgroup is represented by charged choline and phosphate beads, the backbone region is represented by glycerol and ester beads with opposing point dipoles and the hydrophobic tailgroups are represented by C-triplet beads, where each bead represents three carbon atoms (b) molecular structure of toluene.

count the bias of insufficient sampling was suggested by Gore *et al.*^{29,30} who proposed an alternative correction to reduce the sampling bias, following the observation that the bias over a small number of trajectories ($N < 10^2$) has a linear relationship with N , meaning that an approximate bias can be numerically computed. This modified JE sampling bias is defined by Gore as:

$$B_J = \frac{\langle W(z)_{\text{dissipation}} \rangle}{N^{\alpha_b}} \quad (6)$$

where α_b is the rate at which the bias goes to zero with a small number of samples:

$$\alpha_b = \frac{\ln [\beta C_b \langle W(z)_{\text{dissipation}} \rangle]}{\ln [C_b (e^{2\beta \langle W(z)_{\text{dissipation}} \rangle} - 1)]} \quad (7)$$

where C_b is a fitted parameter that determines the boundary of the small N - large N regime. We have utilised $C_b = 15$, as tested by Gore²⁹. From this, we can estimate a bias-corrected Jarzynski estimator:

$$\Delta G_{\text{bias}} = \Delta G_J - B_J \quad (8)$$

When computing the average work along the reaction coordinate, we followed a block-averaging procedure, where every value in a 1 Å interval along the reaction coordinate

was binned and averaged, and the variance within each bin was determined using the mean value of each bin. For each bin, we computed several terms to assess the validity of the JE for the elucidation of an accurate FE profile. We computed the ΔG_J term, $\Delta G_{\text{cumulant}}$, ΔG_{bias} , and the $\sigma_{W(z)}^2$.

For the JE-SMD simulations our approach to compute the work was as follows. After 10 ns of equilibration of the bilayer, the toluene molecule was inserted at 30 Å from the bilayer center of mass, in the normal (z) direction. Four different constant velocities were then assigned to four independent simulations: $8 \times 10^{-6} \text{ Å fs}^{-1}$ (*faster*), $5 \times 10^{-6} \text{ Å fs}^{-1}$ (*fast*), $8 \times 10^{-7} \text{ Å fs}^{-1}$ (*slow*) and $5 \times 10^{-7} \text{ Å fs}^{-1}$ (*slower*). For each simulation, the toluene was pulled from $z = 30 \text{ Å}$ to $z = 0 \text{ Å}$. The faster velocities are consistent with those used in previous studies and slower velocities were also included here to systematically analyse the effect of velocity. For example, Schulten *et al.* used velocities of 0.0001 Å fs^{-1} , 0.01 Å fs^{-1} and 0.1 Å fs^{-1} when validating JE-SMD through stretching deca-alanine¹⁶, where in each case highly exaggerated velocities were used to test the validity of JE in a truly non-equilibrium environment. In other examples, Kucuyak *et al.*^{18,19} used velocities ranging from $1.25 \times 10^{-6} \text{ Å fs}^{-1}$ to $10.0 \times 10^{-6} \text{ Å fs}^{-1}$, which are directly comparable to the velocities used in our simulations. Given this, we can conclude that the velocities we have chosen are reasonable

to test the JE-SMD method. The *faster* and *fast* velocities can be defined as the ‘fast-growth’ regime, while *slow* and *slower* velocities can be recognised as the ‘slow-growth’ regime. The biasing potential (Hook’s potential) to anchor the toluene molecule to the reaction coordinate used a force constant of $25 \text{ kcal mol}^{-1} \text{ \AA}$. The work values were collated over 20 repetitions of the pulling simulation. To analyse the effect of using increased sampling blocks, we measured the ΔG_J term, $\Delta G_{\text{cumulant}}$, ΔG_{bias} and the $\sigma_{W(z)}^2$ term over $N = 10, 15$ and 20 simulation runs. Table I summarises the simulations undertaken in this study.

Index	Method	N	$v(\text{\AA fs}^{-1})$	$k(\text{kcal mol}^{-1} \text{ \AA})$	$t(\text{ns})$
1	JE-SMD	20	8×10^{-6}	25	100
2			5×10^{-6}		120
3			8×10^{-7}		720
4			5×10^{-7}		1200
5	US	30	-	25/50	900

TABLE I: List of simulations run in this study. 1-4 are the JE-SMD simulations at different velocities and 5 is the set of US simulations used as the benchmark/control to compare against. N is the number of simulation trajectories used for averaging the work or the number of sampling windows for US, v is the velocity of the toluene in the JE-SMD simulations, k is the spring constant used for constant velocity pulling in JE-SMD or for restraining the toluene molecule to the reaction coordinate in US and t is the total simulation time.

III. RESULTS

A. JE-SMD/US Comparison

Snapshots from the JE-SMD simulations of the toluene molecule at different positions relative to the bilayer are shown in Figure 2. Table II shows the compiled ΔG values and growth regimes of the different JE-SMD implementations at four different velocities and for the benchmark US simulations. The α_b plot used for computing the ΔG_{bias} values is shown in Figure 3. The results from US are shown in Figure 4. The development of the FE profile with respect to the simulation time of each window is also shown, demonstrating that the calculation had converged by 20 ns (per window). The sampling profile for the US is shown in Figure S1. An overall FE change of $-5.2 \text{ kcal mol}^{-1}$ (corresponding to $-23.0 \text{ kJ mol}^{-1}$) was calculated by US, which is slightly larger favourable FE change than the values obtained with the united-atom Berger and CG MARTINI FFs of $-12.5 \text{ kJ mol}^{-1}$ and $-20.92 \text{ kJ mol}^{-1}$ respectively^{32–34}. The overlap of the plots at 25 and 30 ns for the FE profile of the US indicated convergence. Here, a key feature to compare with the JE-SMD results is the energy minimum near $z = 15 \text{ \AA}$. This energetic preference and localization has been noted

to be due to the cation- π interactions between the uncharged aromatic substrates and the choline groups of the DOPC lipids, which has been shown to be present with substrates without a dipole such as benzene^{35,36}.

The free energy profiles obtained from the ΔG_J , $\Delta G_{\text{cumulant}}$, and ΔG_{bias} implementations of JE-SMD for each velocity are shown in Figure 5. Each JE-SMD plot also shows how the profiles converge as averaging is carried out over successively larger numbers of JE-SMD simulations. For the $v = 8 \times 10^{-7} \text{ \AA fs}^{-1}$ (*faster*) simulations, we see an overall change of -3.0 to $6.0 \text{ kcal mol}^{-1}$, -3.0 to $9.0 \text{ kcal mol}^{-1}$ and -3.0 to $6.5 \text{ kcal mol}^{-1}$ for the ΔG_J , $\Delta G_{\text{cumulant}}$ and ΔG_{bias} values respectively. Similar results are seen with the $v = 5 \times 10^{-7} \text{ \AA fs}^{-1}$ (*fast*) simulations, which show very similar patterns with increasing N simulations - it is clear that the JE-SMD simulations in the ‘fast-growth’ regime (faster and fast velocities), do not reproduce the US results effectively, and that there is minimal effect on the FE profile upon increasing the number of simulations, N .

Interpreting the ‘slow-growth’ regime simulations, for the $v = 8 \times 10^{-7} \text{ \AA fs}^{-1}$ JE-SMD simulations (*slow*), we see a overall improvement in the PMF profiles - we see an overall change of -2.9 to $2.0 \text{ kcal mol}^{-1}$, -2.3 to $4.0 \text{ kcal mol}^{-1}$ and -2.3 to $2.5 \text{ kcal mol}^{-1}$ for the ΔG_J , $\Delta G_{\text{cumulant}}$ and ΔG_{bias} respectively. There is an 0 and 20.0 \AA with the $\Delta G_{\text{cumulant}}$ and ΔG_{bias} , which represents a marginal improvement, but is clearly not close to convergence to the US result. At velocities of $v = 5 \times 10^{-7} \text{ \AA fs}^{-1}$ (*slower*), we see the closest consistency with the US profile at $N = 10$ simulations, with the emergence of the FE minimum at 15 \AA . Within these profiles, we see the key features that are consistent with the US profile, i.e. the minima in the regions 10.0 to 20.0 \AA and 0.0 to 7.0 \AA . The results are marginally improved by averaging over larger number of samples ($N = 10, 20$), where we see an overall FE change of $-3.2 \text{ kcal mol}^{-1}$, $-0.5 \text{ kcal mol}^{-1}$ and $-2.5 \text{ kcal mol}^{-1}$ for the ΔG_J , $\Delta G_{\text{cumulant}}$ and ΔG_{bias} profiles respectively. Here, it is clear that the ΔG_J and ΔG_{bias} shows the clearest consistency with the US results, while the $\Delta G_{\text{cumulant}}$ profile shows a smaller FE minimum near the 10 to 20 \AA region. The $W(z)_{\text{dissipation}}$ used to compute B_J for ΔG_{bias} is shown in the supplementary information (Figure S2). We observe a dramatic trough below $z = 20 \text{ \AA}$, which represents the bilayer region. The trend we see is an overall decrease in the magnitude of the $W(z)_{\text{dissipation}}$, where we see convergence towards values of $10.0 \text{ kcal mol}^{-1}$ (*faster*), $8.0 \text{ kcal mol}^{-1}$ (*fast*), $5.5 \text{ kcal mol}^{-1}$ (*slow*), $4.5 \text{ kcal mol}^{-1}$ (*slower*).

IV. DISCUSSION

From the work distributions along the reaction coordinate (Figure S3), we have shown that the distribution follows a series of Gaussian-like profiles, which satisfies the condition for the cumulant approximation

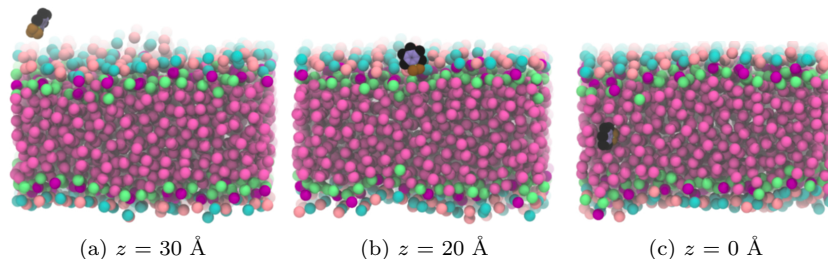


FIG. 2: Snapshots from the simulation trajectories of the US simulations showing the toluene molecule in the DOPC bilayer at a selection of z coordinates of the toluene molecule. The snapshots were produced using VMD³¹.

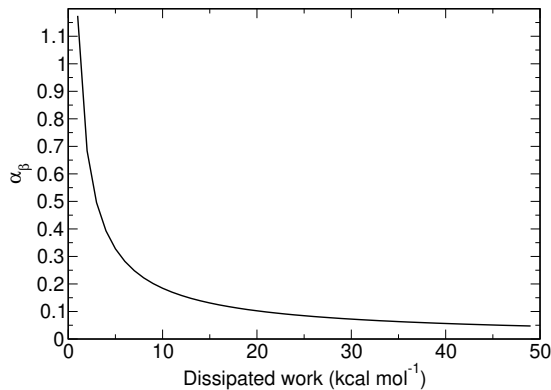


FIG. 3: The plot for $\alpha_b = \frac{\ln[\beta C_b(W(z)_{\text{dissipation}})]}{\ln[C_b(e^{2\beta(W(z)_{\text{dissipation}})} - 1)]}$, for dissipation work between 0 - 50 kcal mol⁻¹. Interpreting the $\frac{\beta}{2}\sigma_W^2$ as the dissipation work, the α_b was selected when computing the $\Delta G(z)_{\text{bias}}$ from the JE-SMD simulations.

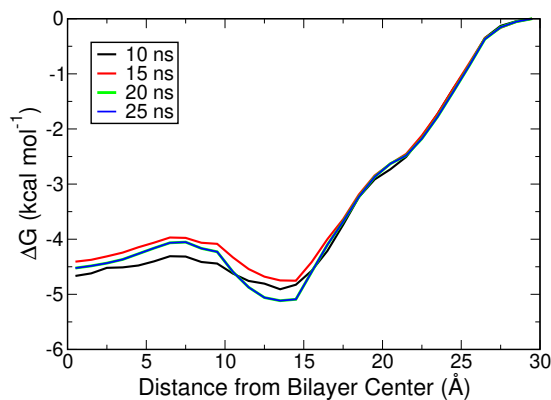


FIG. 4: The US simulations for the toluene. The WHAM collated values at 5, 10, 15 and 25 ns is shown to illustrate the convergence of the FE profile at approximately 30 ns.

of the JE-SMD method to be applicable. In the JE-SMD validation by Schulten¹⁶, the $W(z)_{\text{dissipation}}$ was reported to be between 1.9 and 4.3 kcal mol⁻¹ (3.1 k_BT to 7.1 k_BT), where a smaller estimate corresponded to

FE Method	$v \text{ (\AA fs}^{-1}\text{)}$	$\Delta G(\text{kcal mol}^{-1})$	Type
ΔG_J	8×10^{-6}	6.0	fast-growth
	5×10^{-6}	5.5	fast-growth
	8×10^{-7}	2.0	slow-growth
	5×10^{-7}	-3.2	slow-growth
$\Delta G_{\text{cumulant}}$	8×10^{-6}	9.0	fast-growth
	5×10^{-6}	9.0	fast-growth
	8×10^{-7}	3.8	slow-growth
	5×10^{-7}	-0.5	slow-growth
ΔG_{bias}	8×10^{-6}	6.0	fast-growth
	5×10^{-6}	6.0	fast-growth
	8×10^{-7}	2.5	slow-growth
	5×10^{-7}	-2.5	slow-growth
US	-	-5.2	-

TABLE II: Data for the overall change in ΔG for each JE-SMD simulation and the benchmark US simulations.

a PMF profile that was closer in convergence with the corresponding US example. With our toluene simulations, we observe $W(z)_{\text{dissipation}}$ values in the ranges of 1.0 to 10.0 kcal mol⁻¹, 1.0 to 8.0 kcal mol⁻¹, 1.0 to 5.0 kcal mol⁻¹, and 1.0 to 4.5 kcal mol⁻¹ for the *faster*, *fast*, *slow*, *slower* simulations respectively (Figure S2). The highest values for $W(z)_{\text{dissipation}}$ occur in the bilayer interior ($z < 20 \text{ \AA}$). A greater value of $W(z)_{\text{dissipation}}$ results in a broader probability distribution for the $W(z)_{\text{dissipation}}$ ^{29,30}, and shows that the convergence of an accurate estimate of $P(W(z)_{\text{dissipation}})$ is a concern within the bilayer interior regions. Other bilayer comparison studies of the JE-SMD and US methods by Kuyucak *et al.*^{18,19} show that even a small inclusion of Coulombic interactions and flexible peptide molecules can affect efficient sampling with both fast and slow simulations due to the slow relaxation time. In other studies, Warshel *et al.*³⁷ has shown that, showing that the convergence of a simple gramicidin A channel takes of the order of s, which clearly highlights this issue. Hence to make a valid comparison of the FE methods, we were compelled to select a system that was minimally affected by these factors. In our example, the toluene molecule was explicitly chosen as a rigid, non-ionic, hydrophobic molecule

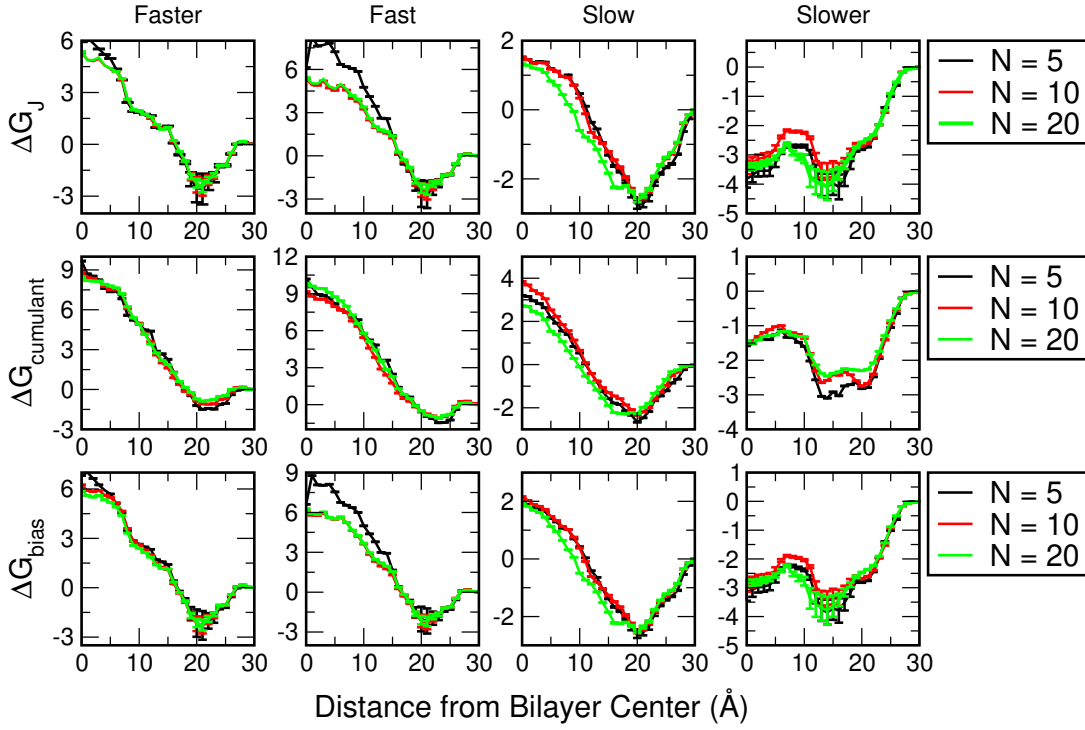


FIG. 5: The FE profiles computed from the $v = 8 \times 10^{-6} \text{ \AA fs}^{-1}$ (faster), $v = 5 \times 10^{-6} \text{ \AA fs}^{-1}$ (fast), $v = 8 \times 10^{-7} \text{ \AA fs}^{-1}$ (slow), and $v = 5 \times 10^{-7} \text{ \AA fs}^{-1}$ (slower) JE-SMD simulations. The first row shows the ΔG_J , the second row shows the $\Delta G_{\text{cumulant}}$, and the third row shows the ΔG_{bias} . For each row, the convergence of JE-SMD simulation profiles was shown from $N = 10, 15$ and 20 simulations respectively. The units for ΔG_J , $\Delta G_{\text{cumulant}}$ and ΔG_{bias} are kcal mol^{-1} .

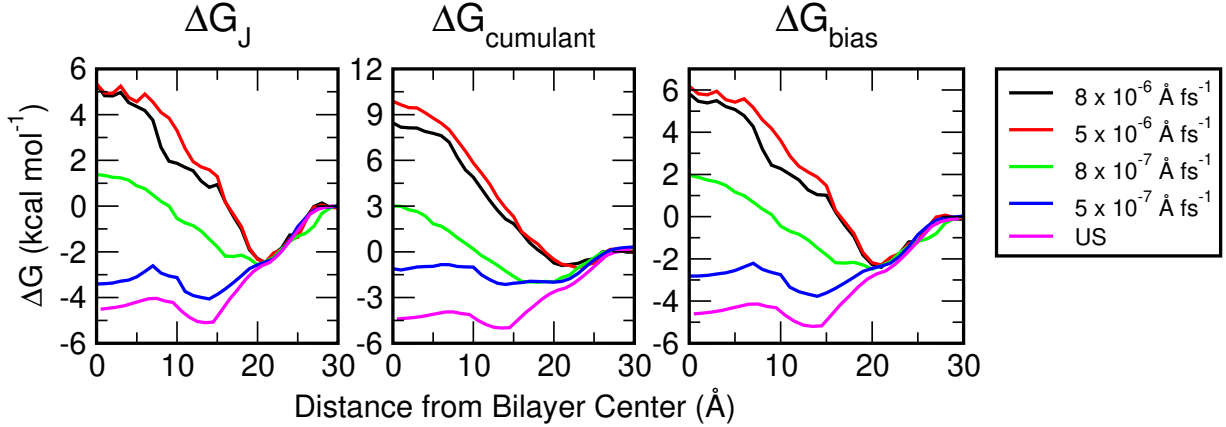


FIG. 6: The collated FE profile JE-SMD simulations, as shown by the $v = 8 \times 10^{-6} \text{ \AA fs}^{-1}$, $5 \times 10^{-6} \text{ \AA fs}^{-1}$, $8 \times 10^{-7} \text{ \AA fs}^{-1}$, $5 \times 10^{-7} \text{ \AA fs}^{-1}$ (corresponding to the faster, fast, slow, and slower velocities respectively) JE-SMD simulations (ΔG_J , $\Delta G_{\text{cumulant}}$, ΔG_{bias}) with the US result as the benchmark profile.

which would circumvent the sampling issues related to the coulombic potentials.

The final comparison between the JE-SMD and US FE profiles are shown in Figure 6. In the *faster*, *fast* and *slow* cases, the region between 0 and 20 Å fails to capture the overall shape or magnitude entirely. It is only with the $5 \times 10^{-7} \text{ \AA fs}^{-1}$ (*slower*) JE-SMD samples that we start to see a pattern of consistency with the US re-

sult in terms of the shape of the profile, though not with the magnitude. The closest estimate is the change of $-3.2 \text{ kcal mol}^{-1}$ with the ΔG_J , with similar changes seen with the ΔG_{bias} result of $-2.5 \text{ kcal mol}^{-1}$. The $\Delta G_{\text{cumulant}}$ result shows the worst agreement with the US result, where the small trough in the range 10 to 15 Å is barely noticeable. The cumulant estimator, which was designed to correct the original JE-SMD estimator, proves infe-

rior in this case compared to both the original and bias-corrected estimator. The overall trend we observe with a JE-SMD simulation is that the slower the velocity, the greater the improvement towards the US result. Therefore, although Schulten *et al.*¹⁶ has commented that the JE is valid for processes at any velocity, even at velocities where it becomes seemingly more practical compared to the US method, the bias due to the under sampling of rare-trajectories makes the JE-SMD method impractical to use, except for the simplest systems.

V. CONCLUSION

We have used a model of a toluene molecule crossing a phospholipid bilayer to establish the validity of JE-SMD methods compared to US for computing the FE profile of a molecule crossing a membrane. We have divided the JE-SMD experiments into two velocity regimes - a ‘fast-growth’ regime where the total simulation time is low compared to the US method, and a ‘slow-growth’ regime where the total simulation time is comparable to or exceeds the total simulation for US sampling. The FE profile from the US simulations confirmed that the FF we have used produces FE data that are consistent with known literature, and hence is valid to use as a benchmark to compare the JE-SMD simulations. To measure the efficiency and accuracy of JE-SMD methods compared to US, we implemented three interpretations of the JE: the raw JE interpreter, ΔG_J , the cumulant second-order interpreter $\Delta G_{\text{cumulant}}$ which corrects the sampling bias, and ΔG_{bias} , which is an alternate method for taking into account high sampling bias. Within the scope of a simple bilayer/toluene simulation, we failed to see a convergence towards the US result for $N = 20$ repeat runs in the ‘fast-growth’ regime, with all interpretations of the JE-SMD. In the ‘slow-growth’ regime the results varied; while we observed significant improvements in terms of the appearance of the shape of the FE profile (particularly the location of the FE minimum), the JE-SMD did not fully converge towards the US result at even the slowest velocities. Hence neither bias corrected interpretation of JE-SMD significantly alleviated the sampling problem.

It is clear that the primary factor in improving the JE-SMD profile is the convergence of the $W(z)_{\text{dissipation}}$ and sampling - slower velocities allows for this, but at the rate which this sampling space is reached, the number of simulations required becomes such that US remains the most efficient option. In the ‘slow-growth’ regime, we have already allocated a longer total simulation time compared to the US simulations, which again indicates that the JE-SMD method needs modification if it is to be practically utilised. As this bilayer system was chosen as a simple benchmark bilayer system, it is likely that the sampling issues seen in this study would apply to larger systems with heterogenous and complex components. Hence, while the JE-SMD method does provide

an alternative approach to computing FE profiles, issues remain in its interpretation, and further work is required to correctly account for significant bias of the region. Recently, alternative modifications using *adaptive stochastic perturbation protocols* (ASPP)³⁸ or *multistep trajectory combination* (MSTC)³⁹ have been suggested for improvements onto the JE. In the case of ASPP method, the method improves the computed FE profile by widening the work distribution, whilst in the case of the MSTC method, small numbers of trajectories are combined in steps. Application of these methods to computing membrane transport properties requires further investigation, and will be explored in future work.

VI. ACKNOWLEDGEMENTS

The authors acknowledge the Scientific Computing Research Technology Platform at the University of Warwick for providing the computational resources to carry out this work. RN thanks the Royal Society for a University Research Fellowship.

- ¹M. Edidin, “Lipids on the frontier: a century of cell-membrane bilayers,” *Nature Reviews Molecular Cell Biology* **4**, 414 – 418 (2003).
- ²D. E. Engelman, “Membranes are more mosaic than fluid,” *Nature* **438**, 578 – 580 (2005).
- ³E. Awoonor-Williams and C. N. Rowley, “Molecular simulation of nonfacilitated membrane permeation,” *Biochimica et Biophysica Acta Biomembranes* **1858**, 1672–1687 (2016).
- ⁴M. Kang and S. M. Loverde, “Molecular Simulation of the Concentration-Dependent Interaction of Hydrophobic Drugs with Model Cellular Membranes,” *Journal of Physical Chemistry B* **118**, 11965–11972 (2014).
- ⁵N. J. Yang and M. J. Hinner, “Getting Across the Cell Membrane: An Overview for Small Molecules, Peptides, and Proteins,” *Methods in Molecular Biology* **1266**, 29–53 (2015).
- ⁶S. J. Marrink and H. J. C. Berendsen, “Permeation Process of Small Molecules across Lipid Membranes Studied by Molecular Dynamics Simulations,” *Journal of Physical Chemistry* **100**, 16729–16738 (1996).
- ⁷W. Shinoda, “Permeability across lipid membranes,” *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1858**, 2254 – 2265 (2016), biosimulations of lipid membranes coupled to experiments.
- ⁸C. Neale and R. Poms, “Sampling errors in free energy simulations of small molecules in lipid bilayers,” *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1858**, 2539 – 2548 (2016), biosimulations of lipid membranes coupled to experiments.
- ⁹R. M. Venable, A. Krmer, and R. W. Pastor, “Molecular dynamics simulations of membrane permeability,” *Chemical Reviews* **119**, 5954–5997 (2019), <https://doi.org/10.1021/acs.chemrev.8b00486>.
- ¹⁰T. Piekot and J. Trylska, “Computational methods used to explore transport events in biological systems,” *Journal of Chemical Information and Modeling* **59**, 1772–1781 (2019), PMID: 30879306, <https://doi.org/10.1021/acs.jcim.8b00974>.
- ¹¹S.-J. Marrink and H. J. C. Berendsen, “Simulation of water transport through a lipid membrane,” *The Journal of Physical Chemistry* **98**, 4155–4168 (1994), <https://doi.org/10.1021/j100066a040>.
- ¹²G. M. Torrie and J. P. Valleau, “Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the subcritical Lennard-Jones fluid,” *Chemical Physical Letters* **28**, 578–581 (1974).

- ¹³J. Shankar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," *Journal of Computational Chemistry* **13**, 1011–1021 (1992).
- ¹⁴J. Kastner, "Umbrella Sampling," *WIREs Comput Mol Sci* **1**, 932–942 (2011).
- ¹⁵C. Jarzynski, "Nonequilibrium equality for free energy differences," *Physical Review Letters* **78**, 2690–2693 (1997).
- ¹⁶S. Park and K. Schulten, "Calculating potentials of mean force from steered molecular dynamics simulations," *Journal of Chemical Physics* **120**, 5946 (2004).
- ¹⁷R. C. Lua and A. Y. Grosberg, "Practical Applicability of the Jarzynski Relation in Statistical Mechanics: A Pedagogical Example," *Journal of Physical Chemistry B* **109**, 6805–6811 (2005).
- ¹⁸T. Bastug, P. Chen, S. M. Patra, and S. Kuyucak, "Potential of mean force calculations of ligand binding to ion channels from Jarzynski's equality and umbrella sampling," *Journal of Chemical Physics* **128**, 155104 (2008).
- ¹⁹T. Bastug and S. Kuyucak, "Application of Jarzynski's equality in simple versus complex systems," *Chemical Physics Letters* **436**, 383 – 387 (2007).
- ²⁰D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. T. Jr, and C. Busamante, "Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems," *Nature Letters* **90** (1998).
- ²¹T. Pienko and J. Trylska, "Computational Methods Used to Explore Transport Events in Biological Systems," *Journal of Chemical Information and Modeling* **59**, 1772 – 1781 (2019).
- ²²M. Orsi and J. W. Essex, "Physical properties of mixed bilayers containing lamellar and nonlamellar lipids: insights from coarse-grain molecular dynamics simulations," *Faraday Discussions* , 249 – 272.
- ²³M. Orsi and J. W. Essex, "The ELBA Force Field for Coarse-Grain Modeling of Lipid Membranes," *Plos One* **6**, e28637 (2011).
- ²⁴J. Wang, R. M. Wang, J. W. Cadwell, P. A. Kollman, and D. A. Case, "Development and testing of a general AMBER force field," *J. Comput. Chem* **25**, 1157–1174 (2004).
- ²⁵S. Plimpton, "Fast Parallel Algorithms for Short-Range Molecular Dynamics," *Journal of Computational Physics* **117**, 1–19 (1995).
- ²⁶M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics," *Journal of Chemical Physics* **97**, 1990 (1992).
- ²⁷S. Genheden and J. W. Essex, *Journal of Chemical Theory and Computation* **11**, 4749 – 4759 (2015).
- ²⁸"WHAM: the weighted histogram analysis method," <http://membrane.urmc.rochester.edu/content/wham>.
- ²⁹J. Gore, F. Ritort, and C. Busamante, "Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements," *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12564 – 12569 (2003).
- ³⁰X. Daura, R. Afontenanger, and A. E. Mark, "On the Relative Merits of Equilibrium and Non-Equilibrium Simulations for the Estimation of Free-Energy Differences," *ChemPhysChem* **11**, 3734 – 3743 (2010).
- ³¹W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *Journal of Molecular Graphics* **14**, 33 – 38 (1996).
- ³²T. V. Pogorelov, J. V. Vermaas, M. J. Arcario, and E. Tajkhorshid, "Partitioning of Amino Acids into Model Membrane: Capturing the Interface," *The Journal of Physical Chemistry B* **118**, 1481–1492 (2014).
- ³³J. L. MacCallum, W. F. D. Bennett, and D. P. Tieleman, *Biophysical Journal* **94**, 3393 – 3404 (2008).
- ³⁴L. 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," *Journal of Chemical Theory and Computation* **4**, 819 – 834 (2008).
- ³⁵W. Yau, W. C. Wimley, K. Gawrisch, and S. H. White, "The Preference of Tryptophan for Membrane Interfaces," *Biochemistry* **37**, 14713–14718 (1998).
- ³⁶K. E. Norman and H. Nymeyer, "Indole Localization in Lipid Membranes Revealed by Molecular Simulation," *Biophysical Journal* **91**, 2046–2054 (2006).
- ³⁷A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, "Modeling electrostatic effects in proteins," *Biochimica et Biophysica Acta* **1764**, 1647 – 1676 (2006).
- ³⁸O. Perisic and H. Lu, "On the Improvement of Free-Energy Calculation from Steered Molecular Dynamics Simulations Using Adaptive Stochastic Perturbation Protocols," *PLoS ONE* **9**, e101810 (2014).
- ³⁹I. Echeverria and L. M. Amzel, "Estimation of Free-Energy Differences from Computed Work Distributions: An Application of Jarzynski's Equality," *Journal of Physical Chemistry B* **116**, 10986 – 10995 (2012).