The Journal of Chemical Physics

Sampling efficiency of the counting method for permeability calculations estimated with the inhomogeneous solubility–diffusion model

Cite as: J. Chem. Phys. **154**, 054106 (2021); https://doi.org/10.1063/5.0033476
Submitted: 16 October 2020 . Accepted: 08 January 2021 . Published Online: 02 February 2021

Samaneh Davoudi, and An Ghysels

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Sampling efficiency of the counting method for permeability calculations estimated with the inhomogeneous solubility–diffusion model

Cite as: J. Chem. Phys. 154, 054106 (2021); doi: 10.1063/5.0033476
Submitted: 16 October 2020 • Accepted: 8 January 2021 •
Published Online: 2 February 2021

Samaneh Davoudi and An Ghysels[a)]

AFFILIATIONS

IBiTech–BioMMeda Group, Ghent University, Corneel Heymanslaan 10, Block B-entrance 36, 9000 Gent, Belgium

[a)]Author to whom correspondence should be addressed: an.ghysels@ugent.be

ABSTRACT

Permeability is a key property in various fields such as membrane technology for chemical separation and transport of substances through cell membranes. At the molecular scale, the counting method uses the number of membrane crossings in a conventional unbiased molecular dynamics simulation to predict the permeability. This contribution investigates under which conditions the counting method has insufficient statistics. An equation is derived for a compartmental model based on the inhomogeneous solubility–diffusion (Smoluchowski) model, giving insight into how the flux correlates with the solubility of permeants. This equation shows that a membrane crossing is a rare event not only when the membrane forms a large free energy barrier but also when the membrane forms a deep free energy well that traps permeants. Such a permeant trap has a high permeability; yet, the counting method suffers from poor statistics. To illustrate this, coarse-grained MD was run for 16 systems of dipalmitoylphosphatidylcholine bilayer membranes with different permeant types. The composition rule for permeability is shown to also hold for fluxes, and it is highlighted that the considered thickness of the membrane causes uncertainty in the permeability calculation of highly permeable membranes. In conclusion, a high permeability in itself is not an effective indicator of the sampling efficiency of the counting method, and caution should be taken for permeants whose solubility varies greatly over the simulation box. A practical consequence relevant in, e.g., drug design is that a drug with high membrane permeability might get trapped by membranes thus reducing its efficacy.

I. INTRODUCTION

Permeability is a key property to describe the kinetics of mass transport of a substance through a membrane. It is highly relevant in various research and application fields, such as membrane technology for chemical separation[1–4] and passive permeation of substances through the cell membrane.[5,6] In biology, permeation through the cell membrane is a vital process that is relevant for drug delivery[5,6] and the regulation of cell death and proliferation.[7] A large number of experimental studies have been carried out to provide information on this process during the past decade.[8–14] However, these methods are either expensive or not informative enough, as they are usually unable to provide data on a molecular level.

On the other hand, computational simulations have been proven to give more insight into permeation at the atomic scale.[15–17] Among simulation methods, molecular dynamics (MD) can lay open more details on the kinetics of the process, the property that experiments often cannot provide directly, under the condition that a proper force field for the molecular interactions is available.[18] Different simulation methods have been applied to calculate the permeability, many of which rely on enhanced sampling methods[18–20] and the inhomogeneous solubility–diffusion (ISD) model based on the Smoluchowski equation.[21–28] In the ISD model, the permeability P is expressed in terms of local diffusivity and free energy gradients.[21] A limitation of this approach, however, is that it is generally not clear *a priori* to which degree and in which situation the diffusive

hypothesis is valid for the system under study. Moreover, accessing the permeability in this model can be challenging regarding the simulation setup and Bayesian analysis (BA).[29]

Permeability can also be determined by counting crossings through the membrane.[17,30] In the case of rare events, there might be an insufficient number of crossings, and the statistics can be poor. To set the mind, imagine aiming for a 10% error on the number of crossings. When crossings are occurring independently according to a waiting time distribution of a Poisson process, at least 100 events would be needed to achieve the 10% standard error.[25] An example of using the counting method is oxygen and water diffusion through phospholipid membranes.[24,25,31–33]

In this paper, the number of crossings expected in a simulation will be investigated as a function of the simulation properties. Several properties are dictated by the simulation setup, such as the number of permeants and box dimensions. A straightforward way to improve the sampling of the unit cell is to increase the system size. When the cross section doubles and the total number of permeants doubles, the concentration in the solvent phase will remain unaltered, while the number of crossings would double. There are unfortunately limits to this approach because the computational cost increases faster than linearly with the box size, and in comparison, it would be more efficient to double the simulation time T_{sim}. Another approach could be to insert more permeants without altering the simulation box size, meaning that all concentrations—both solvent and membrane—increase. One should be wary of the clustering of the permeants, however, and the permeability might become concentration dependent. For instance, oxygen showed no clustering for oxygen concentrations about ten-fold higher than the oxygen solubility in water,[24] while ethanol permeation through phosphatidylcholine (POPC) membranes differs substantially between low and high concentrations.[29,34]

Other properties are inherent to the chosen membrane, solvent, and permeants. This paper considers the solubility of permeants (or, equivalently, the partitioning coefficient K) and the membrane thickness h_m. It is well known that the membrane forms a high free energy barrier when a permeant is nearly insoluble in the membrane, and the associated low permeability prevents good sampling of the unit cell. It will now be shown that a high permeability, for permeants that are highly soluble in the membrane, is another cause for poor sampling. In contrast to low permeability, a high permeability membrane is often regarded as a membrane where the flux is not rate-limiting, but we will show that high permeability cannot guarantee good sampling of the simulation box either.

The membrane thickness follows from the self-aggregation of phospholipid molecules in the bilayer. It is less discussed in the permeability literature than the partitioning coefficient. Still, using the simplified Overton's model $P = KD/h_m$,[35] where D is the diffusion constant, it is clearly of equal importance as the partitioning coefficient. The membrane thickness is however ill-defined since there is no unique answer as to where an atom ends and another atom begins, and moreover, membranes exhibit thermal undulations with lipid protrusions. In experimental work, a roughly estimated membrane thickness is used to estimate the permeability.[10,36] In simulation work, it is often derived in relation to the plane of the phosphate groups.[29] Hence, another aim of this paper is to showcase the sensitivity of the permeability to the considered thickness. Therefore, in addition to the membrane thickness, the thickness of two additional water layers will be treated as an extra parameter in the discussion. Moreover, this parameter will allow us to discuss the overall sampling efficiency. When the limit of the water layer thickness is taken to cover the whole simulation box, the associated number of crossings represents crossings through the whole simulation box.

Section II will use the ISD model to derive a theoretical expression of how the counting method is affected by the solubility of permeants in the membrane. Section III discusses the dependencies and limiting cases of the derived equation. It is shown under which conditions membrane crossings can be considered as a rare event. A composition rule for the flux is derived as well. Section IV presents the number of crossings of 16 different permeant types through the dipalmitoylphosphatidylcholine (DPPC) bilayer membrane, illustrating the derived equation. Section V summarizes our findings.

II. THEORY

The Smoluchowski equation describes the transport of permeants through an inhomogeneous medium on a free energy profile $F(z)$ with a position-dependent diffusivity $D(z)$, where z is the coordinate along the membrane normal.[21] Assume F_{ref} is the free energy at the reference location, usually taken to be in the water phase on both sides of the membrane. The permeability P of a layer of thickness h follows from solving the Smoluchowski equation under steady-state conditions,[21]

$$\frac{1}{P} = e^{-\beta F_{\mathrm{ref}}} \int_h \frac{e^{\beta F(z)}}{D(z)} dz, \tag{1}$$

where the integration runs over a certain thickness h, and $\beta = 1/(k_B T)$ is the inverse temperature with T being the temperature. In a compartmental model, as illustrated earlier for oxygen transport,[24] F and D are piecewise constant functions, and the integral becomes a sum over the several compartments,

$$\frac{1}{P} = \sum_i \frac{1}{P_i} = \sum_i \frac{h_i}{K_i D_i}, \tag{2}$$

where each layer i is characterized by its thickness h_i, diffusivity D_i, and free energy F_i. The factor $K_i = \exp(-\beta(F_i - F_{\mathrm{ref}}))$ is the partitioning constant in layer i compared to the water phase, and it is equal to the concentration ratio between layer i and the water phase, i.e., $K_i = c_i/c_{\mathrm{ref}}$. The composition rule in Eq. (2) shows that the permeation through a series of layers can be seen as a series of local resistances $1/P_i$, if the same reference F_{ref} is used for each layer.

Let us start with a one-compartmental model for a uniform membrane of thickness h_m, for which the integral in Eq. (1) simplifies to

$$P = \frac{KD}{h_m}. \tag{3}$$

The partitioning coefficient K relates to the free energy difference ΔF between the membrane and the water phase, where ΔF is also

referred to as the free energy of transfer to bring a permeant from the water phase into the membrane.

This model is extended to the three-compartmental model depicted in Fig. 1, where a water layer of thickness d is added on both sides around the membrane. For simplicity, the diffusivity is assumed to be constant over the whole simulation box. This model will allow for a more complete discussion of the behavior of the counting method. The permeability of this model is obtained by putting the previous permeation resistance $h_m/(KD)$ in series with the permeation resistance of the water phase $2d/D$, where the partitioning coefficient of the water phase is simply 1. This gives the permeability of the membrane plus water layers equal to

$$P = \frac{D}{2d + \frac{h_m}{K}}. \tag{4}$$

Filling in $d = 0$ in Eq. (4) results back in Eq. (3).

Two regimes exist for P, depending on the partitioning coefficient K of the membrane. In the limit of low K, the concentration in the membrane is much lower than in the water phase, and P reduces to the previous equation, $P \approx KD/h_m$, as if the water layer is unimportant. The permeability is then indeed dominated by the height of the barrier and is a linear function of K. This is expected from the integral form in Eq. (1), which is dominated by high F values. In the limit of large K, at least compared to $h_m/(2d)$, the membrane has a high concentration of permeants, and P becomes approximately a constant, $P \approx D/(2d)$, meaning that the permeability of the membrane is not the limiting factor but rather the resistance coming from the water layer.

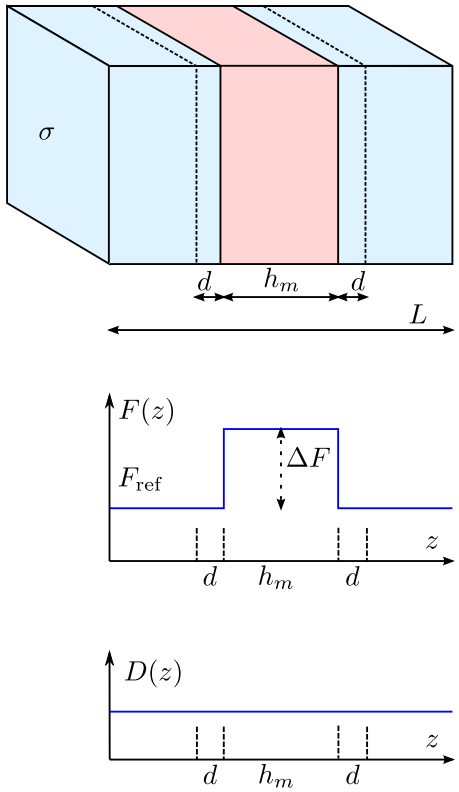

FIG. 1. Simulation box with length L and cross area σ containing the membrane (red) and water (blue). The compartmental model considers the membrane (thickness h_m) and a water layer on both sides (thickness d each). Diffusivity is assumed to be constant, and the free energy profile is piecewise constant.

Given this compartmental model, we will now turn toward the practical situation of a simulation setup. In a simulation, P can be computed using the counting method, by determining the crossing rate J and the reference concentration c_{ref} in the water phase. The number n_{cross} of complete transitions through a membrane with cross section area σ in either direction is counted during a long equilibrium MD simulation of length T_{sim}. Next, the (bidirectional) flux J per unit of time and unit of area in either direction is derived,

$$J = \frac{n_{\text{cross}}}{T_{\text{sim}}\sigma} = |J_{\leftarrow}| + |J_{\rightarrow}|, \tag{5}$$

where $J_{\leftarrow}$ and $J_{\rightarrow}$ are the flux through the membrane in the negative or positive direction, respectively. The well-known formula[17,30] for P from the counting method is

$$P = \frac{|J_{\leftarrow}| + |J_{\rightarrow}|}{2c_{\text{ref}}} = \frac{J}{2c_{\text{ref}}}. \tag{6}$$

The counting method for permeability relies on the number of crossings. The aim is to discuss the number of crossings as a function of the key properties of the membrane and the simulation setup: (1) the free energy of transfer ΔF, (2) considered layer thickness $h = h_m + 2d$, (3) the total number of permeants N_t that are present in the simulation box, and (4) the dimensions of the simulation box, i.e., the height L and cross area σ. For this reason, n_{cross} will be written as a function of these parameters in order to observe their role. Using the first equality in Eq. (5), it is clear that a discussion of J contains similar information as a discussion of n_{cross}, apart from the role of the simulation time T_{sim} and cross area σ, and therefore, the discussion of n_{cross} will be largely based on the discussion of J.

From Eq. (6), it immediately follows

$$J = 2c_{\text{ref}}P, \tag{7}$$

where P is given by Eq. (4). The task at hand is thus to express the reference concentration located in the water phase, as a function of the key parameters K and N_t. The water concentration $c_{\text{ref}} = c_w$ is the average number of permeants $\langle N_w \rangle$ in the water phase divided by the volume of the water phase, which is $\sigma(L - h_m)$ according to Fig. 1. Similarly, the membrane concentration c_m is the average number of permeants $\langle N_m \rangle$ in the membrane divided by the membrane volume σh_m. Note that, while the total number of permeants N_t remains constant during the simulation, the number of particles in the water and membrane, N_w and N_m, respectively, will fluctuate over time. Their averages are given as

$$\langle N_w \rangle = c_w \sigma (L - h_m), \tag{8}$$

$$\langle N_m \rangle = c_m \sigma h_m = K c_w \sigma h_m. \tag{9}$$

Using the equality $N_t = N_w + N_m = \langle N_w \rangle + \langle N_m \rangle$, the reference concentration becomes

$$c_{\mathrm{ref}} = c_w = \frac{N_t}{\sigma(h_m K + L - h_m)}, \tag{10}$$

and we find J as a function of the key parameters

$$J = \frac{2N_t D}{\sigma} \frac{1}{h_m K + L - h_m} \frac{1}{2d + \frac{h_m}{K}}. \tag{11}$$

This is the equation that will be used for the discussion. The number of crossings follows immediately by multiplying J with $T_{\mathrm{sim}}\sigma$ [Eq. (5)].

In a final step, we leave the case of the compartmental model and return to the case of general $F(z)$ and $D(z)$ profiles that can take any shape. It is still assumed that the transport kinetics can be described by the Smoluchowski equation, i.e., the kinetics are Markovian. In the general case, the reference concentration is given by

$$c_{\mathrm{ref}} = \frac{e^{-\beta F_{\mathrm{ref}}} N_t}{\sigma \int_L e^{-\beta F(z)} dz}. \tag{12}$$

P is given by Eq. (1). Using the $\langle \ldots \rangle$ notation for spatial averages in the z-direction, over h or L, these are rewritten as

$$P = \frac{e^{\beta F_{\mathrm{ref}}}}{h \left\langle \frac{e^{\beta F(z)}}{D(z)} \right\rangle_h} \tag{13}$$

and

$$c_{\mathrm{ref}} = \frac{e^{-\beta F_{\mathrm{ref}}} N_t}{\sigma L \left\langle e^{-\beta F(z)} \right\rangle_L}. \tag{14}$$

The flux in Eq. (7) now becomes

$$J = \frac{2N_t}{hL\sigma \left\langle \frac{e^{\beta F(z)}}{D(z)} \right\rangle_h \left\langle e^{-\beta F(z)} \right\rangle_L}. \tag{15}$$

The above equation is the generalization of the compartmental model in Eq. (11) to arbitrary free energy and diffusivity profiles. It gives the flux J as a function of the key parameters in the general case.

III. DISCUSSION

The flux in the compartmental model [Eq. (11)] is first discussed for the two limiting cases where the layer does not comprise any water ($d = 0$, so $h = h_m$) and where all water is included in the layer ($h = L$), and next, it is illustrated for the intermediate cases ($h_m < h < L$).

For $d = 0$, only the membrane of thickness h_m is considered, and the flux simplifies. For large K (free energy well) and small K (free energy barrier), the dependence on K is constant or linear in K, respectively,

$$J(h_m) = \begin{cases} \frac{N_t D}{\sigma h_m^2}, & K \gg 1 \\ \frac{N_t D}{\sigma h (L - h_m)} K, & K \ll 1. \end{cases} \tag{16}$$

When the membrane is highly soluble for the permeants ($\Delta F > 0$), the number of crossings is independent of K and thus ΔF. However, when the permeants' solubility in the membrane is lower than in the water phase ($\Delta F < 0$), the number of crossings will drop exponentially with ΔF. A very high free energy of transfer will give a very few membrane transitions, with poor statistics for the counting method as a consequence.

For $2d = L - h_m$, the whole water layer in the simulation box is included in the counting of membrane transitions. The considered thickness is $h = L$, so $J(L)$ represents the flux of complete crossings of the simulation box,

$$J(L) = \frac{N_t D}{\sigma (L - h_m)^2} \frac{1}{1 + \frac{h_m}{L - h_m} \frac{1}{K}} \frac{1}{1 + \frac{h_m}{L - h_m} K}. \tag{17}$$

Here, $J(L)$ can be interpreted as a measure of the sampling efficiency of permeants over the whole z-direction of the simulation box. It is interesting to see that this equation is symmetric for the replacement of K by $1/K$. Replacing ΔF by $-\Delta F$ does not alter the sampling efficiency $J(L)$. In other words, $J(L)$ is an even function of ΔF by the appearance of K in this expression. This is an important realization: the flux in a simulation can be low because either a high barrier makes it difficult for the permeants to pass through the membrane or a deep well makes it difficult to leave the membrane once they are trapped. In summary, the larger the difference between the free energy minima and maxima across the membrane normal, given by $|\Delta F|$, the lower is the sampling efficiency because J drops with both K and $1/K$. A high variety in solubility hence leads to lower sampling efficiency.

We now approximate $J(L)$ to allow comparison with the previous approximations of $J(h_m)$ in Eq. (16). For large K (free energy well) and small K (free energy barrier), $J(L)$ is inversely proportional to K or linear in K, respectively,

$$J(L) = \begin{cases} \frac{N_t D}{\sigma h (L - h_m)} \frac{1}{K}, & K \gg 1 \\ \frac{N_t D}{\sigma h (L - h_m)} K, & K \ll 1. \end{cases} \tag{18}$$

When K is small, $J(h_m) \approx J(L)$. This can be understood because the exact location of the boundaries, be it at $h = h_m$ or $h = L$, is relatively unimportant when the barrier is high. When K is large, however, we find that $J(h) \gg J(L)$. When the membrane is highly soluble and only few particles are in the water phase, the exact boundary location does matter. Omitting the water phase or including a small or larger part has a large effect on the number of crossings. The reason is that the permeability of the water phase is in this case the limiting resistance for the overall permeability of the considered layer, hence the sensitivity to the included water layer thickness.

For d between 0 and $(L - h_m)/2$, the number of crossings lies between $J(h_m)$ and $J(L)$. Figure 2 visualizes the trends for the toy system (Fig. 1) with numerical values that are typical for MD settings, for instance for a bilayer consisting of 72 phospholipid molecules.[24] The simulation box had sizes $L = 70$ Å and $\sigma = (50\ \text{Å})^2$, the membrane thickness was set to $h_m = 55$ Å, the box contained $N_t = 10$ permeating particles with diffusivity $D = 0.5\ \text{Å}^2/\text{ps}$, and the total simulation time was set to $T_{\text{sim}} = 1\ \mu\text{s}$.

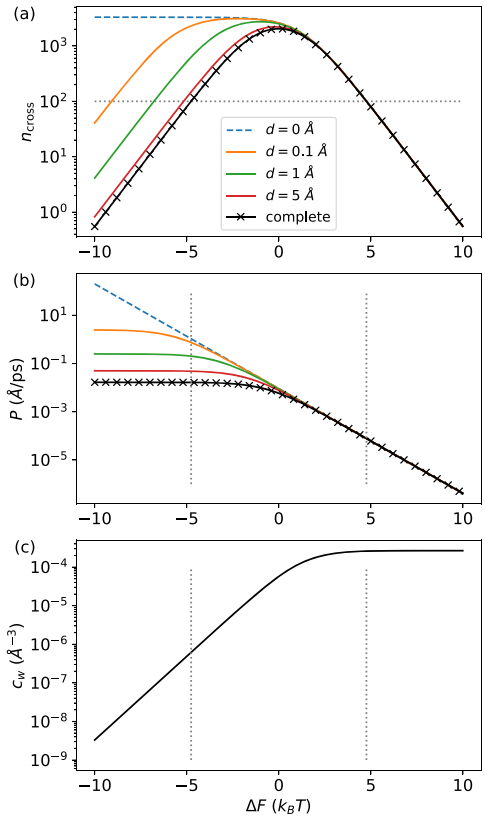

FIG. 2. Number of crossings n_{cross} (a), permeability P (b), and reference concentration c_w (c) in the model system of Fig. 1 as a function of the free energy difference $\Delta F = -k_B T \ln K$, where $\Delta F > 0$ is a barrier and $\Delta F < 0$ is a well. Various water layers d are considered. Blue dashed line: water slab not taken into account. Black line crosses: complete simulation box considered; this is an indicator of overall sampling. The horizontal dashed line in (a) is at 100 crossings, and the vertical dashed lines in (b) and (c) indicate the corresponding free energy difference.

According to Eq. (7), the permeability and the reference concentration are the main contributing factors, which are plotted in Figs. 2(b) and 2(c), for several values of d. The semi-log plots clearly show the constant (zero slope), linear (positive slope), and inversely linear (negative slope) dependence on K in the limiting cases $K \gg 0$ and $K \ll 0$. The number of crossings is mainly the product of these two graphs, and the semi-log plot in Fig. 2(a) is indeed the sum of the curves except for a vertical shift. The widespread curves for small K (i.e., $\Delta F \ll 0$) in Figs. 2(a) and 2(b) are a signature of the sensitivity to the chosen layer thickness when the membrane is a permeant trap. Let us focus on the number of complete crossings of the simulation box (black line), which refers to the sampling efficiency. The vertical dashed line in Fig. 2(a) indicates 100 crossings corresponding with a standard error of about 10%. This corresponds to a free energy difference of about $|\Delta F| = 5k_BT$, indicated with vertical lines in the other plots. This means that the sampling of the complete simulation box becomes poor when $|\Delta F|$ is larger than $5k_BT$ and one should be wary of statistics. When $\Delta F \gg 0$, the statistics are poor because the membrane forms a high barrier making crossings a rare event. However, when $\Delta F \ll 0$, the poor statistics are caused by the low permeant concentration c_w in the water phase [Fig. 2(c)]. When all permeants are trapped in the membrane, meanwhile not contributing to the crossings, there is barely any permeant left in the water phase as a candidate to perform a crossing.

Figure 3 plots the number of crossings as a function of the layer thickness for several values of K. The curves for low K are almost unaffected by considering some water layer. In contrast, the curves for high K indeed decrease much faster than those for low K as those are more sensitive to the layer thickness. For complete crossings at $h = L$, the curve for a given K coincides with the corresponding curve for $1/K$ (indicated with black crosses in the graph), illustrating the symmetry between ΔF and $-\Delta F$.

For all curves in Fig. 3, n_{cross} decreases with increasing thickness h. This is a consequence of the composition rule in Eq. (2) for permeability. A higher thickness gives a lower permeability according to Eq. (2). Using $J = 2c_{\text{ref}}P$, a similar composition rule holds for the flux,

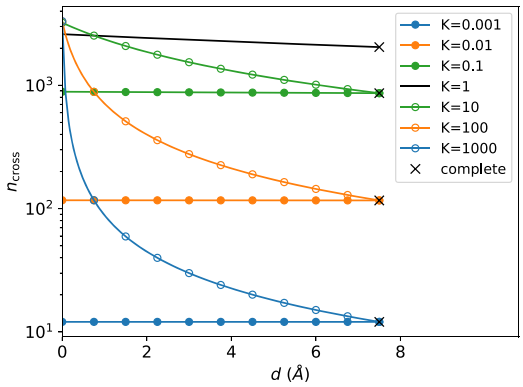

FIG. 3. Number of crossings n_{cross} as a function of the water layer d in the model system of Fig. 1, for various free energy barriers ($K < 0$) and free energy wells ($K > 1$). Black crosses: complete simulation box considered.

$$\frac{1}{J} = \sum_i \frac{1}{J_i}, \tag{19}$$

where J_i is the bidirectional flux through layer i [see Eq. (5)]. An equivalent composition rule holds for $n_{\text{cross}} = T_{\text{sim}}\ \sigma\ J$. Hence, J and n_{cross} also decrease when the considered thickness is increased. For a symmetric bilayer membrane, it immediately follows from the composition rule that the entrance flux through one leaflet is identical to the escape flux through that leaflet, $J_{\text{entr}} = J_{\text{esc}}$, and the complete crossing flux through the entire membrane is half of that, $J = J_{\text{entr}}/2 = J_{\text{esc}}/2$. Likewise, $P = P_{\text{half}}/2$, where P_{half} is the permeability of one membrane leaflet. It is important to note that these composition rules are based on the compartmentalized inhomogeneous solubility–diffusion model and are thus only true under the assumption of perfectly diffusive behavior. Non-Markovian kinetics might distort these composition rules,[37] e.g., when memory effects are at play for ethanol crossings through a phospholipid membrane.[29]

Finally, the symmetry in ΔF is shown in the general case of arbitrary profiles in Eq. (15). When the diffusion profile is taken to be constant, $D(z) = D$, and considering the full simulation box, $h = L$, the flux becomes

$$J(L) = \frac{N_t}{hL\sigma D\left\langle e^{\beta F(z)}\right\rangle_L \left\langle e^{-\beta F(z)}\right\rangle_L}. \tag{20}$$

This highlights that the same number of crossings will be encountered when a profile $F(z)$ is replaced by $-F(z)$, even though the permeability changes. While it may be counterintuitive, a profile with a high free energy barrier will therefore give the same sampling efficiency as a profile with a low free energy barrier. Highly soluble membranes are thus not necessarily improving the statistics in the counting method. This realization is important for drug design, where substances with high permeability might be trapped by the membrane thus reducing their efficacy.

IV. ILLUSTRATION WITH COARSE GRAINED MOLECULAR DYNAMICS

To illustrate the discussion of Sec. III, a membrane of dipalmitoylphosphatidylcholine (DPPC) was simulated at the molecular scale using a coarse grained (CG) model. A range of different permeants was considered with both positive ΔF (barrier) and negative ΔF (well): 16 bead types of the Martini coarse-grained model were used as the permeating molecule. These beads span a wide range of free energy profiles $F(z)$ across the membrane.

A. Simulation details

Using the Gromacs-2019.3 package,[38] simulations were performed on 16 systems consisting of a phospholipid membrane, CG beads, and water. Each system contained 20 permeant beads of one specific CG type, and it was solvated in 2000 water molecules. The membrane consists of 128 DPPC lipid molecules in each simulation, and its initial coordinates were taken from the MARTINI website.[39] The topology file and initial coordinates of the permeant beads were created manually. The MARTINI force field was used in all systems, mapping atoms of each phospholipid molecule and water molecule into 12 beads and 1 bead, respectively.[40]

Water was modeled by applying the SPC/E model.[41] The particle mesh Ewald approach was employed to calculate the Coulombic interactions. Both the Coulombic and the van der Waals interactions were truncated at 1.1 nm, with the potentials shifted to zero at the cutoff using the potential modifiers. The neighbor list length is 1.4 nm updated using the Verlet neighbor search algorithm. With a time step of 30 fs, the leapfrog integrator was used to integrate the equations of motion. A temperature of 320 K was considered for the simulated systems by the coupling of velocity rescale thermostats, with coupling constant set to 1.0 ps. At every time step, the center of mass motion of the system was removed and periodic boundary conditions were applied in all directions (x, y, z). A Parrinello–Rahman barostat was employed to subject the box vectors to semi-isotropic pressure, using a reference pressure of 1 bar, a coupling parameter of 12 ps, and an isothermal compressibility of 3×10^{-4} bar^{-1}. After adding the permeants, an energy minimization was performed. Next, two equilibrations were performed: NVT simulation of 30 ns and NPT simulation of 30 ns. The production run was 270 ns of NPT simulation for each system. Coordinates were stored every 3 ps resulting in 90 000 snapshots in total for each system.

The number of membrane crossings was determined with the Rickflow package code using dividing surfaces at $|z| = h_m/2 = 2.3$ nm.[29] The diffusion tensor with diagonal elements D_{xx}, D_{yy}, D_{zz} was computed with Gromacs.

B. Numerical results

The free energy profile was determined from the histogram $p(z)$ of the position of the CG beads as $F(z) = -k_B T \ln p(z)$. Figure 4 presents $F(z)$ for three exemplary bead types. Bead AC2 has a deep free energy well; hence, it is highly soluble in the membrane ($K \gg 1$).

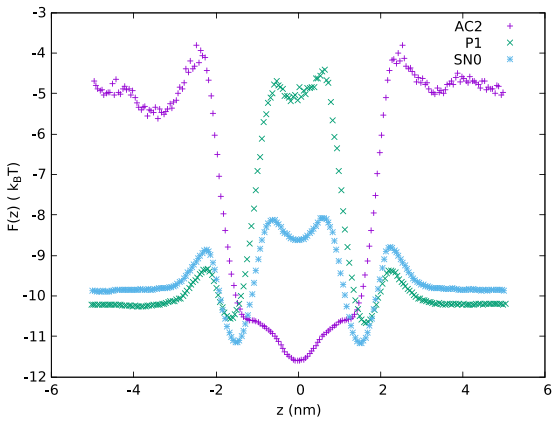

FIG. 4. Free energy of three selected bead types that are representative of a free energy barrier (P1), a free energy well (AC2), and a mix of both (SN0). The profile $F(z) = -k_B T \ln p(z)$ is plotted with $p(z)$, the histogram of the 20 permeants' z-coordinates over all 90 000 snapshots.

Bead P1 has a high free energy barrier and is insoluble in the membrane ($K \ll 1$). Bead SN0 shows a mix of the free energy well and barrier, where the solubility depends on the exact location in the membrane. The other bead types had similar profiles with different elevations.

Figure 5 shows the number of detected crossings vs ΔF for the 16 beads (full symbols), where ΔF is computed as the free energy difference between the membrane center and the membrane border. The graph follows the shape of the curves in Fig. 2(a), where n_{cross} generally drops with $|\Delta F|$. These MD values are compared to the model values of n_{cross} (open symbols) based on the flux in the compartmental model [Eq. (11)] or the more general formulation of the Smoluchowski model [Eq. (15)]. The model parameters were inferred from the molecular dynamics simulations.

For Eq. (11), K is computed as $\exp(-\beta\Delta F)$, $h_m = 4.6$ nm corresponding to the membrane thickness, the average box length $L = 10$ nm, and $T_{\text{sim}} = 270$ ns. The normal diffusivity D_{zz} describes the diffusion normal to the membrane surface, but the measured D_{zz} already incorporates the effects of the barrier or well. Therefore, the lateral diffusivity $D_{\parallel} = (D_{xx} + D_{yy})/2$ was taken as the diffusivity D in the model since diffusion parallel to the membrane is not hindered by free energy barriers/wells. The water thickness d was varied from 0 to $(L - h_m)/2$, giving a range of crossings that is indicated as a dashed vertical line on the graph. The lower end of the line corresponds to the number of crossings through the whole simulation box ($h = L$), while the higher end corresponds to the number of crossings through the membrane only. For $\Delta F \ll 0$, the number of crossings is highly sensitive to the considered membrane thickness. For $\Delta F \gg 0$, the number of crossings is no longer sensitive to the water layer thickness (lines mostly hidden by symbols), in accordance with Fig. 2(a). It was found that a decent correspondence with the MD data could be obtained for a small water layer of thickness $d = 0.1$ nm, which was added to the plot with a gray open symbol. The spread of the vertical lines however indicates that almost any value can be the outcome by tuning the parameter d. This indicates again that the number of crossings, and hence the permeability, is very sensitive to the considered thickness. One should therefore be cautious when interpreting and citing experimental or computational results for highly permeable membranes.

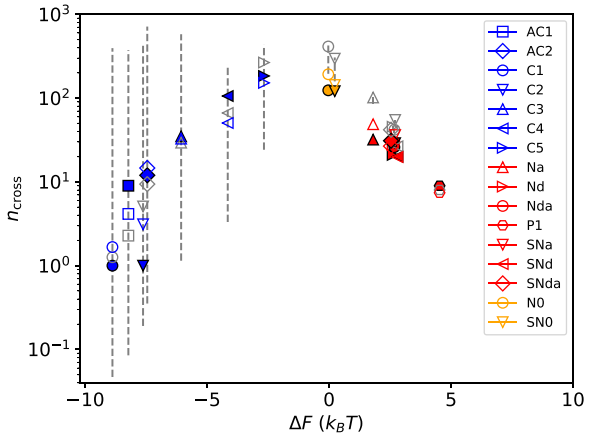

FIG. 5. Number of crossings n_{cross} as a function of ΔF for the 16 permeant types. Filled symbols: value as measured in the MD simulations with color according to the bead type, i.e., well (blue), barrier (red), and mix (orange). Gray open symbols: value based on the compartmental model [Eq. (11)] with $d = 0.1$ nm. Vertical dashed lines show the range between setting $d = 0$ (top of the line) and $d = (L - h_m)/2$ (bottom of the line). Colored open symbols: value based on the general Smoluchowski model [Eq. (15)].

For Eq. (15), the integral is evaluated using the measured free energy profile $F(z)$, the same constant diffusivity, and $d \approx 0$ [not exact because of binning in $p(z)$], giving the values indicated with a colored open symbol. Compared to the compartmental model, this more general model uses a position-dependent solubility, which allows us to describe more details of the membrane. Despite the crudeness in the parameter estimations, a good agreement is found between these model values and the numerical MD data in Fig. 5. The permeants in the present simulations are of course only a single bead without internal degrees of freedom, and the membrane contains only one phospholipid type. For more realistic membranes and realistic permeants, the compartmental model or the more general formulation [Eqs. (11) and (15)] might not have a sufficient level of detail to match the MD results. An extension of the model in Fig. 1 could for instance be a model that has both a free energy well and barrier. However, the complexity of more realistic simulations not only comes from the inhomogeneity in F and D but there might also be memory effects that make the kinetics non-Markovian, while the Smoluchowski equation is based on the Markovian assumption. Another challenging situation is when the permeation is accompanied by hysteresis, which is caused by not observing energy barriers in some orthogonal degree(s) of freedom (orthogonal to the permeation direction z). This is for instance the case when the permeant has an internal degree of freedom that is changed during the permeation, e.g., a dihedral angle in an oligopeptide, or an orientational degree of freedom, e.g., the ring plane in ibuprofen. A two-dimensional or higher-dimensional diffusive equation should then be constructed, which captures all reaction coordinates that are relevant for the permeation.

In these more advanced approaches, based on more elaborate multi-dimensional models with kernels, the free energy would still be one of the main parameters. The present analysis based on a Smoluchowski equation can therefore be seen as a decent first approximation to gain insight into the sampling efficiency. The presented curves clearly illustrate that a higher variation in the free energy profile results in lower sampling efficiency. Both deep wells and high barriers cause a dramatic drop in crossings, and the counting method for permeability will perform inaccurately in those cases.

When the counting method suffers from poor statistics and running much longer MD simulations is not expected to be a sufficient measure, a rare event method for rate calculations can be considered, such as replica exchange transition interface sampling (RETIS)[42–44] or milestoning.[45] In recent work, the connection between RETIS and the permeability was derived,[37] and this method can be a valid alternative for the counting method.[46] Another train of thought is the use of biased simulations, where a large variety of free energy is canceled out artificially by adding a bias potential.[20,47,48]

V. CONCLUSION

This paper investigated the efficiency of counting membrane crossings as a method to determine the permeability. A theoretical relation for the flux was derived based on the ISD model using the compartmentalized toy system in Fig. 1. An additional water layer was included in the model to show the sensitivity to the considered thickness, which causes a large uncertainty for high permeability membranes. It was shown that a high permeability in itself is not a qualified indicator of whether the sampling efficiency of the simulation box will be extensive enough for the counting method to be reliable, and caution should be taken for permeants whose solubility varies greatly over the simulation box. A practical effect of this realization in, e.g., drug design is that substances with high permeability might be trapped by the membrane, thus reducing their efficacy.

Moreover, the composition rule for permeabilities was shown to hold for fluxes as well. The limiting behavior of the number of crossings was determined for a high barrier membrane and a permeant trap membrane, and the dependence on the considered water layer was discussed. It was also shown that the sampling efficiency of the complete simulation box remains unaffected when the free energy profile $F(z)$ is inverted to $-F(z)$, if D is a constant. These theoretical considerations were all made under the assumption of purely diffusive kinetics since the ISD model is based on the Smoluchowski equation for Markovian kinetics.

Besides the theoretical assessment of the counting method, 16 systems were simulated by coarse-grained MD to illustrate the theory. The permeants' crossings were detected and the membrane's free energy profile was characterized as a free energy well, a barrier, or a mix of these two. In both the cases of a high free energy barrier and a deep free energy well, it can be concluded that the permeability calculation by counting crossings will face failure because of the insufficient number of crossing during the simulation. To overcome this problem, one should resort either to running longer MD simulations to observe more crossings using larger simulation boxes with more permeants or implementing other simulation methods that are applicable for rare events, e.g., RETIS.

ACKNOWLEDGMENTS

The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO, and the Flemish Government—department EWI.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] B. Smit and T. L. M. Maesen, Chem. Rev. **108**, 4125 (2008).
[2] M. Tagliazucchi, *Chemically Modified Nanopores and Nanochannels* (William Andrew Publishing, 2017), Chap. 1, pp. 1–25.
[3] Y.-x. Shen, W. Song, D. R. Barden, T. Ren, C. Lang, H. Feroz, C. B. Henderson, P. O. Saboe, D. Tsai, H. Yan, P. J. Butler, G. C. Bazan, W. A. Phillip, R. J. Hickey, P. S. Cremer, H. Vashisth, and M. Kumar, Nat. Commun. **9**, 3304 (2018).
[4] V. A. Poteryaeva, M. A. Bubenchikov, A. M. Bubenchikov, and A. V. Lun-Fu, Sci. Rep. **10**, 15631 (2020).
[5] A. Mandal, V. Agrahari, V. Khurana, D. Pal, and A. K. Mitra, Expert Opin. Drug Delivery **14**, 385 (2017).
[6] B. J. Bennion, N. A. Be, M. W. McNerney, V. Lao, E. M. Carlson, C. A. Valdez, M. A. Malfatti, H. A. Enright, T. H. Nguyen, F. C. Lightstone, and T. S. Carpenter, J. Phys. Chem. B **121**, 5228 (2017).
[7] L. Galluzzi, I. Vitale, S. A. Aaronson, J. M. Abrams, D. Adam, P. Agostinis, E. S. Alnemri, L. Altucci, I. Amelio, D. W. Andrews, and M. Annicchiarico-Petruzzelli, Cell Death Differ. **25**, 486 (2018).
[8] W. K. Subczynski, M. Pasenkiewicz-Gierula, R. N. McElhaney, J. S. Hyde, and A. Kusumi, Biochemistry **42**, 3939 (2003).
[9] J. Widomska, M. Raguz, and W. K. Subczynski, Biochim. Biophys. Acta, Biomembr. **1768**, 2635 (2007).
[10] M. Möller, J. Lancaster, and A. Denicola, Curr. Top. Membr. **61**, 23 (2007).
[11] J. Godyń, D. Gucwa, T. Kobrlova, M. Novak, O. Soukup, B. Malawska, and M. Bajda, Talanta **217**, 121023 (2020).
[12] J. W. Choi, S. H. Bae, M. Kwak, T. G. Lee, M. B. Heo, and D. W. Lee, Sens. Actuators, B **321**, 128624 (2020).
[13] P. Bacchin, D. Snisarenko, D. Stamatialis, P. Aimar, and C. Causserand, J. Membr. Sci. **614**, 118485 (2020).
[14] C. Lin, T.-C. Kuo, J.-C. Lin, Y.-C. Ho, and F.-L. Mi, Int. J. Biol. Macromol. **160**, 558 (2020).
[15] E. Awoonor-Williams and C. N. Rowley, Biochim. Biophys. Acta, Biomembr. **1858**, 1627 (2016).
[16] N. Pokhrel and L. Maibaum, J. Chem. Theory Comput. **14**, 1762 (2018).
[17] R. M. Venable, A. Krämer, and R. W. Pastor, Chem. Rev. **119**, 5954 (2019).
[18] W. Shinoda, Biochim. Biophys. Acta, Biomembr. **1858**, 2254 (2016).
[19] C. Hoffmann, A. Centi, R. Menichetti, and T. Bereau, Sci. Data **7**, 1–7 (2020).
[20] C. T. Lee, J. Comer, C. Herndon, N. Leung, A. Pavlova, R. V. Swift, C. Tung, C. N. Rowley, R. E. Amaro, C. Chipot, Y. Wang, and J. C. Gumbart, J. Chem. Inf. Model. **56**, 721 (2016).
[21] S.-J. Marrink and H. J. C. Berendsen, J. Phys. Chem. **98**, 4155 (1994).
[22] S. J. Marrink and H. J. C. Berendsen, J. Phys. Chem. **100**, 16729 (1996).
[23] K. Gaalswyk, E. Awoonor-Williams, and C. N. Rowley, J. Chem. Theory Comput. **12**, 5609 (2016).
[24] A. Ghysels, R. M. Venable, R. W. Pastor, and G. Hummer, J. Chem. Theory Comput. **13**, 2962 (2017).
[25] O. De Vos, R. M. Venable, T. Van Hecke, G. Hummer, R. W. Pastor, and A. Ghysels, J. Chem. Theory Comput. **14**, 3811 (2018).
[26] H. A. L. Filipe, M. Javanainen, A. Salvador, A. M. Galvão, I. Vattulainen, L. M. S. Loura, and M. J. Moreno, J. Chem. Theory Comput. **14**, 3840 (2018).
[27] R. J. Ferreira and P. M. Kasson, ACS Infect. Dis. **5**, 2096 (2019).
[28] H. Yang, M. Zhou, H. Li, T. Wei, C. Tang, Y. Zhou, and X. Long, ACS Omega **5**, 4798 (2020).
[29] A. Krämer, A. Ghysels, E. Wang, R. M. Venable, J. B. Klauda, B. R. Brooks, and R. W. Pastor, J. Chem. Phys. **153**, 124107 (2020).
[30] R. J. Dotson, C. R. Smith, K. Bueche, G. Angles, and S. C. Pias, Biophys. J. **112**, 2336 (2017).
[31] O. De Vos, T. Van Hecke, and A. Ghysels, Adv. Exp. Med. Biol. **1072**, 399 (2018).
[32] A. Ghysels, A. Krämer, R. M. Venable, W. E. Teague, E. Lyman, K. Gawrisch, and R. W. Pastor, Nat. Commun. **10**, 5616 (2019).
[33] R. J. Dotson and S. C. Pias, Adv. Exp. Med. Biol. **1072**, 405 (2018).
[34] M. Ghorbani, E. Wang, A. Krämer, and J. B. Klauda, J. Chem. Phys. **153**, 125101 (2020).
[35] Q. Al-Awqati, Nat. Cell Biol. **1**, E201 (1999).
[36] M. N. Möller, Q. Li, M. Chinnaraj, H. C. Cheung, J. R. Lancaster, Jr., and A. Denicola, Biochim. Biophys. Acta, Biomembr. **1858**, 2923 (2016).

[37]A. Ghysels, S. Roet, S. Davoudi, and T. van Erp, "Exact non-Markovian permeability from rare event simulations," Phys. Rev. X (unpublished) (2021).
[38]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **1-2**, 19 (2015).
[39]S. J. Marrink, A. H. De Vries, and A. E. Mark, J. Phys. Chem. B **108**, 750 (2004).
[40]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. De Vries, J. Phys. Chem. B **111**, 7812 (2007).
[41]P. Mark and L. Nilsson, J. Phys. Chem. A **105**, 9954 (2001).
[42]T. S. van Erp, Phys. Rev. Lett. **98**, 268301 (2007).
[43]A. Lervik, E. Riccardi, and T. S. van Erp, J. Comput. Chem. **38**, 2439 (2017).
[44]E. Riccardi, A. Lervik, S. Roet, O. Aarøen, and T. S. Erp, J. Comput. Chem. **41**, 370 (2020).
[45]L. W. Votapka, C. T. Lee, and R. E. Amaro, J. Phys. Chem. B **120**, 8606 (2016).
[46]E. Riccardi, A. Krämer, T. S. van Erp, and A. Ghysels, J. Phys. Chem. B **125**(1), 193–201 (2021).
[47]E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[48]J. Comer, K. Schulten, and C. Chipot, J. Chem. Theory Comput. **10**, 554 (2014).