Ultrafast molecular dynamics of biofuel extraction for microalgae and bacteria milking: blocking membrane folding pathways to damaged lipid-bilayer conformations with nanomicelles

Jean-Numa Gillet*

Jawaharlal Nehru University (JNU), School of Computational and Integrative Sciences, Centre for Computational Biology and Bioinformatics, New Delhi 110067, India

Communicated by Ramaswamy H. Sarma

(Received 22 January 2014; accepted 19 March 2014)

Cell milking is a 100% renewable green energy for CO$_2$ by extraction of biofuels inside the cytosol of photosynthetic micro-organisms as microalgae and bacteria. The cells are exposed to a hydrophobic solvent forming holes and cracks through their membranes from which the biofuels can leak out. In protein folding, the goal would be to find pathways to the unique functional protein conformer. However, in the lipid-bilayer interaction with the solvent for milking, the objective is to block the pathways for damaged membrane conformations of low free energy with undesired nanostructures, using the solvent properties, as shown with an ab initio structural bioinformatic model. Statistical thermodynamics is used to compute the free energy (including entropy) from the molecular dynamics trajectory of the biomolecular system with many conformational changes. This model can be extended to the general problem of biomolecules folding as for proteins and nucleic acids. Using an adaptation of the Einstein diffusion law, the conformational change dynamics of the lipid bilayer depends on the two diffusion coefficients of the solvent: $D_1$ before the irreversible folding transition time and the much smaller $D_2$ thereafter. In contrast to the n-hexane and n-heptane hydrocarbons of smaller size, the residual $D_2 = 4.7 \times 10^{-7} \text{cm}^2/\text{s}$ of the n-decane solvent, with the highest partition coefficient among the three extractors, is the only to present a $D_2$ value that is significantly below the critical threshold of $10^{-6} \text{cm}^2/\text{s}$. Therefore, the membrane would resist to long hydrocarbons and the exposed cells would remain viable for milking.

Keywords: algal and bacterial biology; cell membrane folding; hydrophobic solvents; molecular dynamics simulation; statistical thermodynamics; entropy; Einstein diffusion law; conformational changes

Introduction

The generalized utilization of agricultural stocks and land-based crops to produce biodiesel presents limitations to sustainability due to negative impacts on the environment and economy (Dismukes, Carrieri, Bennette, Ananyev, & Posewitz, 2008). Biofuel extraction from algae, diatoms and photosynthetic bacterial micro-organisms is, however, one of the most promising alternative bioenergy sources (Bradley, Willis, & Seefeldt, 2011; Cheng, Gao, & Wu, 2009; Chisti, 2008; Gouveia & Oliveira, 2009; Hejazi & Wijffels, 2004; Hu et al., 2008; Misra, Patra, Panda, Sukla, & Mishra, 2012). A technical difficulty of this green nanobiotechnology is that the viability of the cells is required to maintain sustainability, which has been an obstacle to its large-scale development. Indeed, after cell growth and harvesting in cultures, intracellular cytosol constituents, which are convertible to either biodiesel or biomass for energy generation, such as lipids, esters, and alcohols (for instance, triacylglycerol and bioethanol), can be extracted from strains of microalgae, diatoms, and cyanobacteria (also known as blue-green algae) by a bio-organic process called cell milking (Cheng et al., 2009; Dismukes et al., 2008, Hejazi & Wijffels, 2004; Peng, Yuan, Zhao, & Lercher, 2012). As shown in Figure 1, a promising technology to enable milking is to expose the cells to an extracellular hydrophobic solvent to create pores and openings through their membranes. In this theoretical study, we address the problem of the interaction of the plasma membrane, constituted by a phospholipid bilayer, with the external solvent. Thereafter, the biofuels inside the cytosol could leak out of the cells from the defects generated in their membranes (e.g.in combination with an electrochemical voltage difference). Alkanes from C$_6$ (n-hexane) to C$_{12}$ (n-dodecane) backbones, as those used in gasoline mixtures, were preferred for the solvent by experimentalists due to their extracting properties and possible long-term biocompatibility with the cell membrane (Moheimani, Matsuura, & Borowitzka, 2013; Playne & Smith, 1983; Yadugiri, 2009). In fact, microalgae, bacteria, and extremophiles were already known for their resistance to the damages caused by hydrocarbons...
and possible survival inside their mixtures. Experimental studies also showed that the solvent needs to have a high partition coefficient, log $P$, which enables to obtain better extraction properties through the membrane and maintain separation of the extracellular chemicals with the aqueous cytosol otherwise causing contamination of the cell cultures (Leo, Hansch, & Elkins, 1971; Moojat, Foucault, Pruvost, & Legrand, 2008). The chemical function descriptor, log $P$, is usually taken as the logarithm of the partition coefficient $P$ of a solvent between 1, n-octanol and water to quantify solvent hydrophobicity and miscibility between the lipid bilayer and cytosol phases in biological systems (http://zinc.docking.org/).

An issue is that current wet-lab experiments cannot assess with the fine details of the essential molecular mechanisms of extraction that could lead to milking. Therefore, an objective of this paper is to propose an ab initio structural bioinformatic model, which is as accurate as possible, to give directions and answers to the experimentalists on the cell membrane interaction with hydrocarbons for cell milking purposes but, more generally, on the fundamental biophysics and biochemistry problem of the phospholipid membrane folding by extracellular organic solvents in living systems. This study could also give indications to researchers involved in the possibility of prebiotic or extraterrestrial life in less hydrophilic – more hydrophobic environments than the today’s Earth (Luisi, Walde, & Oberholzer, 1999; Trevors, 2002). Last but not least, the molecular models developed in this article can be adapted to the incorporation of other solvents than hydrocarbons and also extended to the study of the general biomolecules folding problem for proteins and nucleic acids as well (Anfinsen, 1973; Ben-Naim, 2012; DePaul, Thompson, Patel, Haldeman, & Sorin, 2010; Dill & Chan, 1997; Gillet & Ghosh, 2013).

We use Molecular dynamics (MD) simulation (Adcock & McCammon, 2006; Allen & Tildesley, 1991; Karplus & McCammon, 2002) with 1 fs time step for fine tuning and statistical thermochemistry models based on the heat capacity $C_v$ and entropy $S$. Indeed, in the cell membrane interaction problem, complete calculation of the free energy $G = E - TS$ is mandatory at room temperature $T \approx 298$ K (where $E$ is the total energy) due to the significant hydrophobic phenomena and conformational changes of the phospholipids exposed to the solvent. It is Highlighted that the entropic effects were unfortunately neglected in a number of previous computational approaches on the more general problem of biomolecules folding (Ben-Naim, 2012; Gillet & Ghosh, 2013; Killian, Kravitz, & Gilson, 2007; Mosgaard, Jackson, & Heimburg, 2013; Numata & Knapp, 2012). On the other hand, we update the Einstein diffusion law (Allen & Tildesley, 1991; Mark & Nilsson, 2002), which is necessary to explain the ultrafast dynamics of the lipid-bilayers folding, conformational changes and micellization, and connect it to the free energy (see the Methods).

We found that longer hydrocarbon molecules with log $P \geq 5$, but with a number of backbone carbons that can be taken as small as about 10 (as n-decane with log $P = 5.68$), are better to delay irreversible membrane damages as uncontrolled cracks and holes and extend cell life for milking, due to the faster reduction of the non-constant (non-Fick type) diffusion coefficient through the lipid bilayer. After initial fast free energy funneling with a quite significant diffusion coefficient $D_1$ (for $t \leq t^*$), longer hydrocarbons enable to reduce the residual diffusion coefficient $D_2$ (for $t \geq t^*$), obtained after a nanosecond-order transition time $t^*$, to an almost 0 value, which significantly slows down the transition from a flat phospholipid membrane (microstate {1}) to a damaged bilayers structure surrounded by inverse micelles forming lipid nanoparticles or nanowires in its periphery (microstate {2}). After the folding transition time to microstate {2}, the diffusion blocking effect of longer chains leads to a free energy landscape that becomes more stochastic, instead of a funnel (Dill & Chan, 1997), with a significant number of ultrafast $G$ oscillations of moderate amplitude occurring during sub-Angstrom changes of the root-mean square deviation (RMSD) of the backbone carbon positions of the alkanes. The result is that the average $G$ decrease tendency and conformational change speed to the final state of the irreversibly damaged membrane are significantly slowed down. Less work quantity corresponding to the negative of the free energy change $\Delta G = \Delta E - T \Delta S$ (which often depends only

Figure 1. Exposition of the plasma membrane (phospholipid bilayer) of an algal cell to a hydrocarbon medium (orange color) for extraction of biofuel constituents (burgundy color) of the cytosol through holes and defects created by the extracellular solvent, possibly leading to milking of the cell.
marginally on the total energy difference $\Delta E$ due to significant entropic energy fluctuations $-T\Delta S$ is available per time unit for undesired damage self-assembly as lipid nanostructures. Therefore, the bilayer membrane is better conserved over a longer time period with hydrocarbons of larger size, possibly allowing the DNA-proteins machinery (for instance, from a genetically engineered cell strain (Doshi, Nguyen, & Chang, 2013; Kung, Runguphan, & Keasling, 2012)) to repair damages before they become lethal, which can lead to the cell viability and milking.

In contrast, when n-hexane with only six backbone carbons and lower log $P=3.66$ is utilized for the solvent, the residual diffusivity $D_2$ remains too high to block the solvent diffusion. In the model of Figure 2, after the folding transition time $t^*$ (Figure 3), the membrane macromolecular configuration rapidly breaks in contact with n-hexane and irreversibly collapses to a vesicular state (Figure 4). Therefore, by extension of this prototype model to the scale of the complete membrane, the significant damages caused to the bilayers structure by n-hexane quickly lead to cellular death and avoid milking. This membrane-to-vesicle transition also gives a beautiful illustration that the lipid-bilayer structure has been the natural choice of evolution to synthesize cell membranes of different shapes as well as lysosomes, endosomes, and miscellaneous microvesicles and nanovesicles (Luisi, Walde, & Oberholzer, 1999) due to their fast self-assembly mechanism (see the Results and Discussion).

**Methods**

**Slim periodic box for MD**

To model the mechanical and thermodynamic effects of three hydrocarbon solvents n-hexane (C$_6$H$_{14}$), n-heptane (C$_7$H$_{16}$), and n-decane (C$_{10}$H$_{22}$) on the cell membrane for the milking process, we use fine-tuning MD at the atomistic scale (including hydrogen) with a 1 fs time step. As depicted in Figure 2, the initial phospholipid membrane conformation is designed with a bilayer of 1-palmytoil-2-oleoyl-sn-glycero-3-phosphatidylcholine (POPC) molecules (Feller & MacKerell, 2000; Feller, Gawrisch, & MacKerell, 2002). POPC (C$_{42}$H$_{82}$NO$_8$P) is chosen as a benchmark due to its wide natural presence in eukaryotic cell membranes (as in those of algae) and its commercial availability by industrial synthesis. Owing to the large magnitude difference between the nanometric characteristic thickness of the phospholipid bilayer ($d_{bob}$) and micrometric average cell diameter ($2R_{cell}$), the molecular system can be simplified to a nanoscale periodic box containing only a part of the membrane, solvent and cytosol, as delimited by the green-contour box in Figure 2(a).

![Figure 2](image_url)

**Figure 2.** Initial macromolecular-system conformation and periodic box design for molecular dynamics. (a) Sketch of a (x, z) section of the initial cell membrane composed of a phospholipid (POPC) bilayer in interaction with an extracting hydrocarbon solvent (orange color). (b) Periodic box design in the (x, z) plane. The H atoms are not depicted for the n-heptane solvent and POPC bilayer. (c) Top view: Due to the high symmetry degree between the x and y axes, the periodic box is reduced to a slim slice in the y direction. The periodic box contains $N_{at}=39,032$ atoms for the mixture with n-heptane. An ensemble of $N_{lip}=117$ POPC macromolecules are modeled for the three solvents (n-hexane, n-heptane, and n-decane).
Moreover, due to the high $R_{cell}/d_{bl}$ ratio, the membrane can be initially assembled as a flat POPC-bilayer structure in the $(x, y)$ plane of the periodic box with the $z$ axis to measure its thickness (Figure 2(b)). In MD, periodic boundary conditions have to be applied on the external surface of a periodic box for accurate and faster calculation of the electrostatic contribution to the force field using an Ewald particle mesh (Allen & Tildesley, 1991; Leach, 2001). In theory, at low temperature, the charged heads of a pure phospholipid bilayer inside water present a crystallographic ordering in the $(x, y)$ plane at equilibrium. Therefore, due to the resulting high correlation between the $x$ and $y$ axes in Figure 2, the length of the periodic box can be shortened by several folds in the direction $y$ compared with that $x$ to form a three-dimensional (3D) slice (Figure 2(c)), which enables us to better observe the folding phenomenon. This satisfactory symmetry simplification reduces the number of atoms in the periodic box by a large factor in order to significantly fasten the MD calculations. A similar approach of reduction of a 3D molecular system to an almost two-dimensional (2D) problem was already used in inorganic membrane nanomaterials (Gillet, 2010). In addition, due to the high electrostatic stability of the charged phospholipid heads in contact with dipolar water, only a slim inner layer of water is necessary to simulate the aqueous cytosol in the interacting cell membrane problem (Figure 2(b) and (c) bottom). In contrast, the outer solvent layer is defined as much thicker to model a sufficient number of diffusing hydrocarbon molecules (Figure 2(b) and (c) top). The periodic box dimensions are set to $270 \text{ Å} \times 27 \text{ Å} \times 140 \text{ Å}$ in $x \times y \times z$, filling a volume of $1020.6 \text{ nm}^3$ containing $N_{lip} = 117$ POPC molecules (for the three solvents). The number of atoms $N_{at}$ in the periodic box is $38\,541$, $39\,032$ or $39\,921$ for the mixtures with n-hexane, n-heptane and n-decane, respectively. The force-field is defined in the CHARMM27 framework (Domanski, Stansfeld, Sansom, & Beckstein, 2010; Feller & MacKerell, 2000; Feller et al., 2002; Yin & MacKerell, 1998).

First, the POPC-bilayer structure with water layers at both its bottom and top is pre-equilibrated to the target temperature $T_{targ} = 298$ K during a $70$ ps MD run. Thereafter, the top water layer is removed and replaced by a solvent box with the same width in $y$ than that of the below POPC bilayer but a length in $x$ that is a little bit shorter (Figure 2(b) and (c)). The solvent layer is created using PACKMOL (Martinez et al., 2009) in agreement with the room temperature solvent density. The molecular system configuration schematized in Figure 2(b) and (c) is used as initial non-equilibrated conformer to start a long MD run with NAMD (Humphrey, Dalke, & Schulten, 1996; Phillips et al., 2005) to determine the dynamics of the cell membrane in interaction with the three hydrophobic solvents. Several vacancies can be seen in the solvent layer at the border with the outer phospholipid layer in Figure 2(b). Indeed, in this figure, the solvent layer has not been equilibrated yet with the other compounds of the mixture. After MD is started from...
this non-equilibrated configuration, the vacancies are quickly filled by solvent molecules during the first tens of picoseconds after $T_{\text{targ}} = 298$ K is reached. Both NVT and NVE ensembles (Allen & Tildesley, 1991) are used to generate the MD trajectories for the three hydrocarbon solvents.

**Heat capacity: quantitative descriptor of the main conformational changes**

To compute the entropic contribution to $G$, we develop a semi-analytical model based on the heat capacity. During a MD simulation in the canonical NVT ensemble, the dynamic heat capacity $C_v$ at constant volume $v$ can be computed from the variance of the total energy $\langle \delta E^2 \rangle$ as (Allen & Tildesley, 1991; Lebowitz, Percus, & Verlet, 1967; Zuckerman, 2010):

$$C_v^{(\text{NVT})}\{\text{microsets}\} = \frac{\langle \delta E^2 \rangle_{\text{NVT}}}{RT_{\text{targ}}^2}, \quad (1)$$

where $C_v$ is calculated in kcal/(mol K) in this paper, $T_{\text{targ}}$ is set to the target room temperature of 298 K of the Langevin MD thermostat (Phillips et al., 2005), and the perfect gas constant $R = 1.9872041 \times 10^{-3}$ kcal/(mol K).

The microcanonical NVE ensemble is also used in this study. However, Equation (1) cannot be used for $C_v$ calculation in NVE. The problem is that $E$ is constant in NVE so that Equation (1) would result in zero heat capacity. Nevertheless, the potential $U$ and kinetic $K$ energies fluctuate in NVE (in contrast to their summation $E = U + K = \text{const}$). In fact, as proposed by Lebowitz, Percus, and Verlet (1967), $C_v$ can be calculated in NVE from either the variance of the kinetic energy $\langle \delta K^2 \rangle$ or that of the potential energy $\langle \delta U^2 \rangle$ that are equal since $\langle \delta E^2 \rangle = 0$ in NVE. We propose a generalization of their original equation to compute $C_v$ of a molecular system that is not restricted to an ideal gas from a MD trajectory in NVE as

$$C_v^{(\text{NVE})}\{\text{microsets}\} = C_v^{(\text{NVE})} = \frac{C_v^{(\text{NVE})}}{1 - \zeta_{\text{NVE}}} \quad \text{where} \quad \zeta_{\text{NVE}} = \frac{\langle \delta K^2 \rangle_{\text{NVE}}}{RT_{\text{targ}} \langle K \rangle_{\text{NVE}}}, \quad (2)$$

In Equation (2), $C_v^{(\text{NVE})} = (3/2)N_a R$ corresponds to the heat capacity of an ideal monoatomic gas, where $N_a$ is the number of atoms in the simulation periodic box. For a perfect gas, there is no potential energy ($U = 0$) so that $\langle \delta U^2 \rangle_{\text{NVE}} = \langle \delta K^2 \rangle_{\text{NVE}} = 0$, leading to $\zeta_{\text{NVE}} = 0$ in Equation (2). Hence, only for a perfect gas, $C_v^{(\text{NVE})} = C_v^{(\text{NVT})}$ is the same constant quantity than that in the NVT ensemble. Also, note that the average kinetic energy $\langle K \rangle_{\text{NVE}}$ in Equation (2) cannot be set to the $(3/2)N_a R$ value, because the number of practical degrees of freedom of the covalently and non-covalently linked macromolecular system in the periodic box is obviously not equal to $3N_a$, in contrast to the ideal gases tested by Lebowitz et al. (1967).

The average operator $\langle \rangle$ used to compute $C_v$ in Equation (1) for NVT or Equation (2) for NVE represents the mean of the thermodynamic quantity between the brackets for a given collection (“microset”) of equilibrate macromolecular configurations or states obtained between the times $t$ and $t + \Delta t$ of the MD trajectory. The actual $C_v\{\text{microset}\} = C_v(t + \Delta t/2)$ is interpolated at the middle of the time interval $\Delta t$ for the concerned state microset $\{t, t + \Delta t\}$. In this way, we obtain a discrete ensemble of heat capacity samples $C_v\{\text{microsets}\} = C_v[j\Delta t + \Delta t/2]$ for the integer $j = 0, 1, 2, 3, \ldots$ along the complete MD trajectory. The heat capacity $C_v$ at constant pressure $p$ in the isobaric-isothermal NPT ensemble is nearly equal to $C_v$ computed at constant $v$ with Equation (1). Indeed, $C_v^{(\text{NPT})} \approx C_v^{(\text{NVT})}$ in the dense liquid phases contained in the analyzed molecular system, due to (i) the quasi water incompressibility ($C_p^\text{water}/C_v^\text{water} = k_{\text{water}} = 1.01$) and (ii) the high hydrophobic attractions between the hydrocarbons leading to their very small compressibility ($k_{\text{hexane}} = 1.053$ and $k_{\text{heptane}} = 1.062$). The values of $C_v$ and $C_p$ are therefore very close to each other, which can be useful for experimentalists who would like to compare the $C_v$ heat capacities computed using Equation (1) with their $C_p$ experimental data (with a possible deviation of only ~1 to less than 6% in favor of $C_p$).

$C_v$ is an important thermodynamic quantity because a peak in $C_v$ indicates a major conformational change in the macromolecules configuration (Supplementary Material Figure S1). Recall that several successful biophysical experimental technologies (Mosgaard, Jackson, & Heimborg, 2013; Zuckerman, 2010) as the Isothermal titration calorimetry and Differential scanning calorimetry are based on the variation of the heat capacity over time (by measuring the change in the heat flux $Q$ or enthalpy $H$ as $C_v = \Delta Q/\Delta T = \Delta H/\Delta T$ over time, corresponding to $C_v = \Delta E/\Delta T$ in the NVT ensemble with a smooth temperature gradient $\Delta T$). Therefore, in our model, the time interval $\Delta t$ (chosen to compute $C_v$ with Equation (1) or (2)) should not be taken too short otherwise the $C_v$ signal over time, averaged with a too low number of equilibrated states, becomes noisy with many undesired oscillations of unphysical too large amplitude, with respect to the Debye-Einstein limit for semi-classical phonon models of $C_v = 3N_a R$ (Allen & Tildesley, 1991; Gillet, 2013). On the other hand, if $\Delta t$ is too long, a number of $C_v$ oscillations are damped out and it is not possible anymore to locate all significant conformational changes over the simulation time. In practice, we use $\Delta t = 375, 275$ or $340$ ps for the cell membrane in interaction with the n-hexane, n-heptane, or n-decane solvents, respectively. These $\Delta t$ values are chosen because the characteristic ultrafast dynamics of the conformational
changes is in an order of time lower than 0.5 ns, as also obtained in experiments related to the cell membrane (Kel et al., 2013).

Numerical results showing heat capacity peaks (related to conformational changes) over time are given for the n-hexane solvent in the Supplementary Material Part A.

From entropy to free energy

A general Clausius-type integral (Allen & Tildesley, 1991; Zuckerman, 2010) to compute the entropy from the heat capacity is

$$ S = \int_{T_{ref}}^{T} C_v \frac{dT}{T} $$

(3)

at constant $v$. $C_v$ can be replaced by $C_p$ at constant $p$ to compute $S$ in the NPT ensemble. In the latter equation, $T_{ref}$ is a reference initial temperature. Using Equation (3), we compute the entropy of an actual microset of states between the times $t$ and $t+\Delta t$ in NVT or NVE, respectively, as

$$ S^{(NVT,NVE)}_{\{\text{microsets}\}} = S^{(NVT,NVE)}_{\{\text{microsets}\}} \langle \ln(T/T_{ref}) \rangle_{\text{NVT,NVE}} $$

(4)

In Equation (4), $C_v$ is computed from the mean over the same current microset $\{t, t+\Delta t\}$, using Equation (1) for NVT and Equation (2) for NVE. The natural logarithm $\ln(T/T_{ref})$ of the current temperature $T$ vs. $t$ (normalized by the constant $T_{ref}$) is also averaged over $\{t, t+\Delta t\}$. Then, a discrete set of entropy samples $S^{\{\text{microsets}\}} = S_{\lfloor T/T_{ref}\rfloor}$ for the integer $j = 0, 1, 2, 3, \ldots$ is interpolated along the MD trajectory at the middle point of each successive time interval with the same width $\Delta t$. $T_{ref}$ in Equation (4) is set to the mean temperature of the initial microset of states between $t = 5$ ps (in practice) and $t = \Delta t$. As a result, the computed entropy is an absolute quantity because the $S^{\{\text{microsets}\}}$ series starts from a zero value at the first interpolation step. We find starting $T_{ref}$ minimal temperatures that are lower than $T_{\text{targ}} = 298$ K by only $\sim 3-4$ K for the three hydrophobic solvents. Hence, the utilized mean-integral separation in Equation (4) is set to the mean temperature of the initial microset of states between $t = 5$ ps (in practice) and $t = \Delta t$.

Computation of the complete free energy can finally be performed in NVT and NVE as

$$ G^{\{\text{microsets}\}} = E^{\{\text{microsets}\}} - T_{\text{targ}}S^{\{\text{microsets}\}} $$

(5)

where the $S^{\{\text{microsets}\}}$ values are derived from Equations (1)–(4). In NVT, the total energy samples $E^{\{\text{microsets}\}} = E_{\lfloor T/T_{ref}\rfloor}$ for the integer $j = 0, 1, 2, 3, \ldots$ are obtained after the same averaging procedure than that explained in the precedent over the microsets $\{t, t+\Delta t\}$ from the MD trajectory. With Equation (5), undesired Brownian motion noise is as well filtered out from the free energy signal of $t$. After these mathematical operations, a clear free energy function is obtained by cubic spline interpolation of the $G^{\{\text{microsets}\}}$ samples due to their close proximity in time (where deeps correspond to local-minimum configurations and peaks to energy barriers). Equation (5) for $G$ could be extended to the NPT ensemble by addition of a $pv$ (pressure \times volume) term that would shift $G$ to higher energy values.

Langevin piston methods were developed to simulate the constant $p$ effects in MD. Unfortunately, the thermodynamic limit with a macroscale average $v$ should be reached to achieve a constant $p$ with the true NPT partition function (Zuckerman, 2010). Therefore, NPT modeling is still hazardous using a nanoscale MD simulation box. Owing to the fixed size of the periodic box, the $pv$ term would be (in principle) an additive constant. Nevertheless, we are interested in the free energy difference $\Delta G$ but not in its absolute value. The $pv$ term cancels out using Equation (5) for $\Delta G$. Consequently, the previous developments would remain the same in NPT, if we could accurately calculate the $\Delta G^{\{\text{microsets}\}}$ samples at equilibrium in NPT.

Adaptation of the Einstein diffusion law

The general Einstein diffusion relation to obtain the diffusion coefficient $D$ can be written as (Allen & Tildesley, 1991; Mark & Nilsson, 2002):

$$ 2tD = \langle n_{\text{dim}} \rangle^{-1} \langle (r_i(t) - r_i(t_0))^2 \rangle $$

(6)

where $n_{\text{dim}} = 1, 2, 3$ is the number of problem spatial dimensions and $r_i$ represents the coordinates vector to the particle indexed by $i$ vs. $t$. In Equation (6), averaging $\langle \cdot \rangle$ is over the $N$ diffusing particles. In MD, a number of practitioners considered $r_i(t_0 = 0)$ as the coordinates of particle $i$ related to the last configuration returned after initial energy minimization of the macromolecular structure (i.e. $t_0 = 0$ in Equation (6)). Unfortunately, this frozen configuration is calculated at $T = 0$ K but is not equilibrated at $T = 298$ K. Hence, $t_0 > 0$ has to be set to a higher time value and chosen when the $N$ particles are already equilibrated at $T = 298$ K. The starting $t_0$ is set to 400 ps for n-heptane and n-decane but to 800 ps for n-hexane (Figure 3).

To determine if milking is possible, we have to calculate $D$ of the hydrophobic solvent through the cell membrane. Averaging $\langle \cdot \rangle$ in Equation (6) is performed over the $B_{\text{sol}} = 5490$ backbone carbon atoms of the n-hexane solvent in interaction with the membrane in the periodic box ($B_{\text{sol}} = 5719$ carbons for n-heptane and $B_{\text{sol}} = 6150$ carbons for n-decane). Then, Equation (6) can be modified to compute the diffusion coefficient from the MD trajectory (Supplementary Material Part B) as

$$ D = \langle 2n_{\text{dim}} \rangle^{-1} \frac{d[\text{MSD}(t)]}{dt} $$

(7)
where MSD is the mean-square deviation (square of the RMSD) between the solvent carbon coordinates at times \( t \) and \( t_0 \).

First, consider the example of the Fick diffusion (Gillet, Degorce, & Meunier, 2005; Karp, 2013) of diluted small molecules of light molecular masses, from one side of the cell membrane with a higher concentration of the diffusible species to the other side with a lower concentration. To respect the Fick law, \( D \) must be a non-negative constant in the left-hand side of Equation (7). Therefore, the slope \( \gamma = \frac{d[\text{MSD}(t)]}{dt} \) in the right-hand side of Equation (7) must also be a non-negative constant and the MSD vs. \( t \) curve should be a straight line. In practice, the MSD computed from a MD trajectory presents statistical fluctuations. The only way to conciliate the discrete MD and continuous Fick approaches using Equation (7) is that the MSD vs. \( t \) curve could be fitted by a straight regression line MSD\(^2\)(\( t \)), possibly with a high correlation coefficient. Then, the diffusion constant \( D \) is calculated from the regression line slope \( \gamma = \frac{\Delta \text{MSD}}{\Delta t} \) divided by \( 2n_{\text{dim}} \).

The interaction of the cell membrane with a pure hydrophobic solvent seems different from a Fick diffusion mechanism. However, during several nanoseconds after the MD simulation is started, we observed a system evolution with a single constant diffusion coefficient \( D_1 \) of the solvent, similarly to the Fick law (Figure 3). The issue is that \( D_1 \) is quite large. Hence, irreversible deformations of the lipid-bilayer structure could quickly form. This fracture dynamics would be too fast, avoiding the trigger of cellular networks leading to the cell membrane repARATION that occur within longer biological time orders. Hopefully, diffusion of the highly hydrophobic solvent is thereafter quickly blocked when the diffusible molecular species begins to reach the boundary between the lipid bilayer and aqueous cytosol, due to their non-miscibility with water (high log \( P \)). Therefore, the diffusion coefficient becomes a non-constant thermodynamic observable of time (Allen & Tildesley, 1991), as \( D = D(t) \), but tends to slow down with \( t \). However, we find that the diffusivity observable can be interpolated as a single step function of \( t \). Our adaptation of the Einstein diffusion law for the cell membrane interaction with hydrocarbons therefore results in a two-microstate kinetic model, where the constant diffusivity of each microstate \{1\} or \{2\} can be computed from two algebraic equations:

\[
D_1 = \frac{\gamma_1^*}{(2n_{\text{dim}})} \quad \text{for} \quad t \leq t^* \tag{8}
\]

and

\[
D_2 = \frac{\gamma_2^*}{(2n_{\text{dim}})} \quad \text{for} \quad t \geq t^* \tag{9}
\]

In Equation (8), \( D_1 = D(t \leq t^*) \) is the high-value diffusion coefficient during the quick mechanical fracturing by the diffusible molecular species, occurring before \( t^* \). In contrast, after the folding transition time \( t^* \) is reached, the diffusion coefficient significantly decreases to the tiny value \( D_2 = D(t \geq t^*) \) expressed by Equation (9). In our adaptation of the Einstein diffusion law, \( t^* \) corresponds to the time abscissa where the two MSD-interpolation straight lines, fitted from the MD trajectory data, intersect together (Figure 3). Before \( t^* \) is reached, the small residual diffusion coefficient \( D_2 \) is computed from the slope \( \gamma_2^* \) of a second regression line (Equation (9)), as depicted in Figure 3.

The dimensionality of our system is between two and three dimensions (Figure 2(b) and (c)) leading to \( 2 \leq n_{\text{dim}} \leq 3 \) in Equations (6)–(9). However, the periodic box defining our system is 10-fold slimmer in \( y \) than in \( x \) (Figure 2(c)). We, therefore, prefer to assign \( n_{\text{dim}} \) to 2 in Equations (6)–(9) as for a 2D system. In addition, the value \( n_{\text{dim}} = 2 \) enables us to compute an upper bound limit to the diffusion coefficients \( D_1 \) and \( D_2 \). Indeed, we are looking for the solvent with the lowest residual \( D_2 \). If \( D_2 \) is very small, but also an upper bound value, we can be confident that a diffusion blocking effect will occur after the folding transition time \( t^* \) between the two kinetic microstates, which delays the bilayers membrane damaging of further important mechanical fracturing. Also, the diffusion blocking effect can be characterized by the ratio \( D_1/D_2 \) that is independent of \( n_{\text{dim}} \).

Table 1 compares the \( D_1 \) and \( D_2 \) diffusion coefficients for the three hydrocarbons as a function of log \( P \) and \( t^* \). N-heptane presents lightly higher \( D_1 = 2.68 \times 10^{-2} \) cm\(^2\)/s and \( D_2 = 1.90 \times 10^{-6} \) cm\(^2\)/s than those of n-hexane due to its higher log \( P = 4.16 \) (leading to stronger hydrophobic interactions with the phospholipid tails) but for a carbon backbone that is larger by one methyl than that of n-hexane. For n-heptane, the ratio \( D_1/D_2 = 14.1 \) is nearly equal to that of n-hexane (\( D_1/D_2 = 13.6 \)). Therefore, the harmful damaging effects of both solvents to the cell membrane lead to the similar result of avoiding milking. The situation is different for longer hydrophobic chains such as

| No. methyls | log \( P \) | \( D_1 \) (cm\(^2\)/s) | \( D_2 \) (cm\(^2\)/s) | \( D_1/D_2 \) | \( t^* \) (ns) |
|------------|---------|-----------------|-----------------|-------------|---------|
| n-hexane   | 6       | 3.66            | 1.68E-05        | 1.24E-06    | 13.55   |
| n-heptane  | 7       | 4.16            | 2.68E-05        | 1.90E-06    | 14.11   |
| n-decane   | 10      | 5.68            | 1.78E-05        | 4.7E-07     | 37.87   |

Table 1. Comparison between \( D_1 \) and \( D_2 \) for the three solvents.
n-decane. Before the transition time \((t \leq t^*)\), the n-decane diffusion coefficient \(D_1 = 1.78 \times 10^{-5} \text{ cm}^2/\text{s}\) has a similar value than that of n-hexane. This is the result of a balance between the higher n-decane log \(P = 5.68\) (increasing the hydrophobic interaction with the phospholipid tails) and the longer n-decane backbone with 4 additional methyls (making more difficult the penetration through the lipid bilayers). In contrast, for \(t \geq t^*\), the n-decane diffusivity \((D_2 = 4.7 \times 10^{-7} \text{ cm}^2/\text{s})\) sharply decreases by an impressing factor of 37.9-fold.

**Results and discussion**

**Membrane MD with the n-hexane solvent**

We first present computational results for the mixture with n-hexane. The periodic box has the same design than the 3D slim slice depicted in Figure 2 (but the solvent is replaced by n-hexane). A MD trajectory was run for a simulation time of 10.52 ns. Figure 4 presents five transient molecular conformations (listed as ‘0’, ‘3’, ‘4’, ‘5’*, and ‘6’) obtained during this MD run. The conformers ‘0’, ‘3’, ‘4’, ‘5’*, and ‘6’ have to be matched with those denoted by the same indices on the \(G\) vs. \(t\) curve depicted in Figure 5. Used as initial reference, configuration ‘0’ is obtained at time \(t = 10\) ps of the MD trajectory (Figure 4(a)). After only \(<0.5\) ns, a break already begins to open on the top of the phospholipid bilayers, when the first local minimum of the \(G\) vs. \(t\) curve is obtained, corresponding to conformer ‘1’ in Figure 5(a) and (b). At \(t = 4.34\) ns, conformer ‘3’ is reached by the random walker (Figure 4(b)). The breaking point is shown by the top red arrow, where a hole forms. The membrane fracturing is already significant for this configuration so that the initial lipid-bilayers structure is separated in two arms forming a tweezers shape. Phospholipid nanoparticles are also being formed as inverse micelles at the top of the two lipid-bilayers arms by attraction of their charged heads. Due to the hydrophobic interaction between n-hexane and phospholipid tails, these nano inverse micelles can increase their stability over time. A number of n-hexane molecules also diffuses at the left- and right-hand sides of the two phospholipid-arms termini (Figure 4(b) bottom), because closing of the tweezers molecular pattern gives additional free space for the solvent diffusion. Therefore, two other nanoparticles are being self-assembled on the top of the two arms termini. As shown in Figure 4(c), at \(t = 5.44\) ns, the tweezers molecular pattern is already significantly closed (conformer ‘4’). An important hole opens between the two bilayers arms (red arrow). However, due to the non-miscibility, the n-hexane molecules diffusing through the hole do not mix with the inner water (representing the aqueous cytosol). Other diffusing phospholipids (Figure 4(c) bottom) begin to capture almost all water in the periodic box inside a nanocluster to form a nanovesicle (by electrostatic interaction of their charged heads with dipolar water and hydrophobic attraction of their hydrophobic tails with n-hexane). As depicted in Figure 4(d), at the transition time \(t = t^* = 7.31\) ns, the newly assembled nanovesicle has just begun to be completely closed by a surrounding phospholipid layer (conformer ‘5’*). After the transition time \((t \geq t^*)\), the nanovesicle shows minor shape deformations. At \(t = 10.52\) ns, the four phospholipid nanoparticles, circled by dashed black lines in Figure 4(e) depicting conformer ‘6’, have grown in size by self-assembly. Some diffused water molecules leak to join the charged nanoparticle cores. Most of the water nanocluster is now surrounded by single-layer phospholipids instead of natural bilayers (Figure 4(e)). The lipid-bilayer structure is disappearing using n-hexane. For \(t \geq t^*\), the RMSD of the solvent carbon positions augments at a much slower pace with time, compared with the initial ultrafast dynamics of the conformational changes occurring when \(t \leq t^*\). The diffusivity of the hydrocarbon molecules sharply diminishes after the transition time \((t \geq t^*)\). For the n-hexane solvent (see the Methods), the diffusion coefficient reduces by 13.6-folds from \(D_1 = 1.68 \times 10^{-7} \text{ cm}^2/\text{s}\) for \(t \leq t^*\) down to \(D_2 = 1.24 \times 10^{-6} \text{ cm}^2/\text{s}\) for \(t \geq t^*\) (Table 1). The goal is to obtain \(D_2\) as small as possible to slow down the ultrafast dynamics towards even more extended and irreversible damages of the cell membrane. In fact, the residual diffusion coefficient \(D_2 > 10^{-6} \text{ cm}^2/\text{s}\) for both n-hexane and n-heptane solvents presents a too large value. As a result, it would not be possible to limit the cell membrane damages over a sufficiently long period with n-hexane or n-heptane for sustainability. A hydrocarbon with higher log \(P\) and longer carbon backbone has to be used.

**Free energy evolution and E–S compensations**

The free energy \(G = E - TS\) is the balance of the total energy \(E\) and entropic energy \(-TS\). We develop a statistical model based on the heat capacity \(C_v\) to compute \(S\) (see the Methods). Figure 5 displays the \(G\) vs. \(t\) curve calculated from the MD trajectory for the mixture with n-hexane as a thick black line. All the compounds of the conformers listed as ‘0’, ‘3’, ‘4’, ‘5’*, and ‘6’ in Figure 5 were sketched in Figure 4. The ordinate axis in Figure 5 is in kcal/(mol of POPCs). Indeed, the displayed energies are extensive thermodynamic quantities and have to be normalized by \(N_{lip} = 117\). In Figure 5(a), the total and entropic energies are, respectively, depicted with blue dashed and red dashed-dotted lines. The entropic energy \(-TS\) is shifted upwards by the value \(E(0)\) of the initial-conformer total energy for graphical comparison. For \(t \leq t^*\), an ultrafast \(G\) decrease is first observed. Six transient marked conformational changes (related to the six first local minima of the \(G\) vs. \(t\) curve) rapidly occur due to the high diffusion coefficient \(D_1\) until the transition conformer ‘5’* is obtained.
Indeed, the configurations collected during the first part of the MD trajectory (\(t \leq t^*\)) form a large-state ensemble (microstate \(\{1\}\)) with the same kinetics governed by the significant diffusivity \(D_1\), while those appearing after the transition time (\(t \geq t^*\)) constitute a second-state ensemble (microstate \(\{2\}\)) with the much smaller diffusion coefficient \(D_2\) (see the Methods). For the mixture with n-hexane, microstate \(\{1\}\) contains all individual configurations dominated by the bilayers geometry (Figure 5). In contrast, microstate \(\{2\}\) is composed of all states evolving at a reduced kinetics with a low \(D_2\) in the vesicular geometry related to the pink transparency in Figure 5(a) and (c).

\(E\) is computed from three MD sub-trajectories (Figure 5(a)). From \(t = 0\) to 5 ns, the molecular conformations are sampled in the NVT canonical ensemble. During this first sub-trajectory, \(E\) shows a quick decrease trend until the \(G\) local minimum related to conformer ‘3’ is obtained. A total energy–entropy (\(E–S\)) compensation mechanism, also observed in the protein folding problem (Gillet & Ghosh, 2013), already appears during the first
five nanoseconds. Until conformer $'2'$ is reached, both $E$ and $-TS$ terms presents an almost monotonic reduction. However, for $t \geq 1.5$ ns, $E$ continues its descending trend but the $-TS$ contribution augments resulting in a first $E-S$ compensation. At $t = 5$ ns, the MD run is restarted from the last conformer of the first NVT sub-trajectory in the

------

Figure 5. Free energy evolution $G$ vs. $t$ of the mixture with n-hexane by molecular dynamics. (a) Comparison between the free energy $G$ (black thick line) and its components: the total energy $E$ (dashed blue line) and entropic energy $-TS$ (dashed-dotted red line) shifted in $Y$ by the $E(0)$ value. (b) $G$ vs. $t$ trajectory from $t = 0$ to 6 ns. The initial reference conformer $'0'$ is taken at $t = 0.01$ ns and followed by four main local minima of $G$ related to the conformers $'1'$ ($t = 0.56$ ns), $'2'$ ($t = 1.50$ ns), $'3'$ ($t = 4.34$ ns), and $'4'$ ($t = 5.44$ ns), evolving with the high diffusion coefficient $D_1 = 1.68 \times 10^{-5}$ cm$^2$/s. (c) $G$ vs. $t$ trajectory from $t = 5$ to 10.52 ns. The transition conformer $'5'^*$ corresponds to the $G$ local minimum at $t = t^* = 7.31$ ns. Last conformer $'6'$ is taken at $t = 10.52$ ns, but evolves with the reduced diffusion coefficient $D_2 = 1.24 \times 10^{-6}$ cm$^2$/s. In (b) and (c), the POPC structures, in interaction with n-hexane on the periodic box, of the seven conformers listed as $'0'$, $'1'$, ..., and $'6'$ are depicted closed to their indices on the $G$ vs. $t$ curve. The conformers $'0'$, $'3'$, $'4'$, $'5'^*$, and $'6'$ have to be matched with Figure 4(a)–(e), respectively, for their integral representations.
NVE microcanonical ensemble, where $E$ is a thermodynamic constant. The NVE sub-trajectory is from $t = 5$ to 8.71 ns. During this part of the MD run, $\Delta E = 0$ so that the free energy variation $\Delta G = -T\Delta S$ is purely entropic. $E$ now presents a plateau at a higher constant value. The $-TS$ term significantly decreases to compensate the sharp $E$ increase to the plateau, leading to a deeper local minimum of the free energy related to conformer $'4'$ (Figure 5(a)). A zoom on the first six nanoseconds of the free energy evolution curve is displayed in Figure 5(b). The phospholipid structures (excluding the other compounds) of seven chosen conformations are sketched close to their $(t, G)$ locations in Figure 5(b) for the conformers $'0'$, $'1'$, $'2'$, $'3'$, and $'4'$, and in Figure 5(c) for the conformers $'4'$, $'5'$, and $'6'$ (for $t = 5$–10.52 ns). The NVE ensemble is kept until $t = 8.76$ ns (as shown by the $E$ plateau end in Figure 5(a)). The free-energy global minimum (conformer $'5'$) of microstate $\{1\}$ is reached at $t = t^* = 7.31$ ns. At this instant, the $\Delta G$ variation, however, remains of purely entropic nature. This is another example of the requirement to accurately compute $S$ (Figure 5(a)). Conformer $'5'$ marks the transition between the membrane $\{1\}$ and nanovesicle $\{2\}$ microstates. From $t \geq t^*$, the diffusion coefficient is reduced to the lower $D_2$ quantity. However, for n-hexane, the residual $D_2$ keeps a non-negligible value. Consequently, the overall $G$ decrease trend continues after the transition time (Figure 5(a) and (c)). Therefore, the nanovesicle geometry with peripheral nanostructures becomes more stable with time ($t \geq t^*$). At $t = 8.76$ ns, the NVT ensemble is reinitialized, resulting in a single-simulated annealing step (Gillet & Sheng, 2000; Kirkpatrick, Gelatt, & Vecchi, 1983). Indeed, a short minimization-equilibration procedure has to be performed from the last NVE conformer to reinitialize the equilibrium temperature to $T_{\text{target}} = 298$ K in the NVT ensemble. At the same time, a strong $E$–$S$ compensation naturally appears with a jump of the $-TS$ term to compensate the sharp $E$ decrease due to the simulated annealing step.

Cell milking and n-decane
Figure 6 shows the $G$ vs. $t$ curve as a black thick line from the MD trajectory and heat capacity framework (see the Methods) for the mixture with n-decane. The initial reference conformer ‘0’ is taken at $t = 10$ ps. Thereafter, three local minima indexed as ‘1’, ‘2’*, and ‘3’ correspond to main conformational changes. The last conformer ‘4’ of the MD trajectory is at $t = 10.95$ ns. Snapshots of the phospholipid structures of the five conformers ‘0’, ‘1’, ... and ‘4’ are given close to their related $(t, G)$ tuple in Figure 6. Already for conformer ‘1’ (at $t = 2.22$ ns), two breaking points (marked by red arrows in Figure 6) appear in the outer lipid layer due to the strong hydrophobic interaction with n-decane. Conformer ‘2’* is the transition configuration, obtained at the transition time $t = t^* = 4.25$ ns, between microstate

Figure 5. (Continued)
{1} containing the states with the diffusion coefficient $D_1 = 1.78 \times 10^{-5}$ cm$^2$/s ($t \leq t^*$) and microstate {2} composed of those with the residual diffusivity $D_2 = 4.7 \times 10^{-7}$ cm$^2$/s ($t \geq t^*$) (Table 1). The quantities $D_1$ and $D_2$ are calculated from the MSD data of the mixture with n-decane (see the Methods). We observed at the transition time that the membrane is only marginally damaged. In addition, the membrane does not completely close to form a nanovesicle at $t = t^*$, in contrast to the mixture with n-hexane (Figures 4 and 5).

For $t > t^*$, the conformational-change dynamics is almost blocked around the same state due to the near-zero $D_2$. Indeed, for a period of ~5 ns, from conformers ‘3’ ($t = 6.32$ ns) to ‘4’ ($t = 10.95$ ns), $G$ oscillates around 308 kcal/(mol of POPCs) with a non-significant deviation of only +/- 1 kcal/(mol of POPCs), as depicted inside the pink transparency in Figure 6. As shown from the lipid structures of conformers ‘3’ and ‘4’ (Figure 6), a few phospholipids start to diffuse from the left and right membrane termini due to the periodic box definition (Figure 2). Indeed, the POPC atoms located at the extreme left and right membrane termini are not constrained to fixed positions. These phospholipids finally join together by self-assembly to form a small cavity around a water nanocluster (conformer ‘4’). However, this cavity is an artifact of the boundary conditions and remains marginal compared with the vesicle obtained from the mixture with n-hexane. No significant movement of entire parts of bilayers structures occurs between the fractures. The top lipid-bilayers part of the self-assembled structure (conformer ‘4’ in Figure 6) resists quite well to the damages caused by n-decane over time. At the same time, the second hole, on the right-hand side, in the lipid-bilayers top structure of conformer ‘4’ becomes larger in diameter (Figure 6), possibly allowing inner energetic compounds to be extracted out of the cell cytosol by a sustainable process of milking.

**Free energy landscape**

For the mixture with n-decane, the free energy landscape is presented vs. the RMSD* reaction coordinate of the n-decane molecules in Figure 7. The RMSD* coordinate is obtained from the transform $\text{RMSD}^* = \sqrt{\text{MSD}^*(t)}$. As explained in the Methods, the MSD* function of $t$ is calculated from two interpolation fits of the MSD with the
Figure 7. Free energy landscape vs. reaction coordinate RMSD*. (a) Fast G decrease trend until the G local minimum of transition conformer ‘2*’ is reached for RMSD = 61.3 Å. The states collected until conformer ‘2*’ is reached evolve with the diffusivity $D_1 = 1.78 \times 10^{-5}$ cm$^2$/s. (b) For RMSD $\geq 61.55$ Å, the G landscape becomes more stochastic with seven G local minima (including conformer ‘3*’) encountered in a RMSD$^*$ interval of only 10 Å. As highlighted by the pink transparencies in (a) and (b), the diffusivity sharply decreases to $D_2 = 4.7 \times 10^{-7}$ cm$^2$/s, when RMSD$^*$ $\geq 61.3$ Å. The states listed as ‘0’, ‘1’, … and ‘4’ have to be matched to those with the same indices in Figure 5.

Slope $\gamma_1 = 4D_1$ for the first regression line, when $t \leq t'$ and that $\gamma_2 = 4D_2$ for the second regression line when $t \geq t'$. After elimination of the $t$ variable between the RMSD$^*(t)$ and $G(t)$ distributions from the MD trajectory, we obtain a free energy of mean field $G = G$(RMSD$^*$) allowing us to draw the free energy landscape, where the abscissa is the RMSD in Å of the solvent carbon positions and the ordinate is the free energy of their related conformer (for the entire system). The points listed as ‘0’, ‘1’, … and ‘4’ on the G vs. RMSD$^*$ curve in Figure 7 have to be matched with those with the same indices on the G vs. t curve in Figure 6.

Before the RMSD reaches the value of 61.3 Å, the G decrease trend is almost of funnel type (Figure 7). Three first local minima of smooth curvature are encountered by the random walker just after the MD simulation is started from the initial conformer denoted by ‘0’. After occurrence of the local-minimum conformation ‘1’, the G landscape shows a clear funnel shape (Dill & Chan, 1997) until the transition conformer ‘2*’ between the microstates {1} and {2} is obtained. The transition configuration (‘2*’) corresponds to the lowest minimum of the first part of the G landscape, where the interacting solvent molecules have a dynamics governed by the high diffusion coefficient $D_1 = 1.78 \times 10^{-5}$ cm$^2$/s (Table 1). After this intermediate state is passed, the solvent diffusivity sharply decreases by 37.9-fold to the tiny value $D_2 = 4.7 \times 10^{-7}$ cm$^2$/s (see the Methods). We are entering in the second part of the G landscape which is more stochastic (Gillet & Ghosh, 2013, Zuckerman, 2010). Indeed, the G landscape now becomes quite rough with a number of seven local minima within a RMSD$^*$ interval of only ~10 Å, as depicted inside the pink transparency on the right-hand side of Figure 7(a) and (b). After the transition state ‘2*’ is passed, the kinetics (led by $D_2$) for further conformational changes is significantly slowed down for the mixture with n-decane (Figure 7(b)). Hence, the characteristic timescale for their occurrence is significantly enlarged. By increasing the size of our small-dimensionality prototype to the full membrane scale, we envision that the lipid-bilayer damages could be limited for a longer time period. This enhanced diffusion-blocking phenomenon of the solvent molecules would give enough time for cellular processes to be activated and repair the membrane against more significant defects, which would help to maintain sustainability when n-decane or a hydrophobic solvent with a longer carbon backbone as n-dodecane is used, leading to milking.

Conclusion and perspectives
Better milking process sustainability should be obtained with longer hydrocarbons as n-decane, due to its high log $P \geq 5$ and number of methyls $\geq 10$, compared with smaller molecules as the toxic n-hexane and n-heptane. In recent experiments, Xu, Zhang, and Chen (2011) showed that alkane solvents become biocompatible for milking of the microalgae Nannochloropsis sp. only from log $P \geq 5.5$, which also enables to increase the total cytosolic lipid extraction. These experiments confirm our
Figure 8. Free energy evolution $G$ vs. $t$ of the mixture with n-decane by molecular dynamics. (a) Comparison between the free energy $G$ (black thick line) and its components: $E$ (dashed-dotted blue line) and $-TS$ (dashed red line) shifted in $Y$ by the $E(0)$ value. The conformers '0', '1', '2*', '3', and '4' have to be matched with their corresponding indices in Figures 6 and 7. Their related POPC structures are depicted in Figure 6. The same $E$–$S$ compensation mechanism than that observed in Figure 6 for the n-hexane solvent can be observed.

Figure 9. Free energy evolution $G$ vs. $t$ of the mixture with n-heptane by molecular dynamics. (a) Comparison between the free energy $G$ (black thick line) and its components: $E$ (dashed-dotted blue line) and $-TS$ (dashed red line) shifted in $Y$ by the $E(0)$ value. Before the transition time $t^* = 4.25$ ns, the conformers evolved with the large diffusion coefficient $D_1 = 2.68 \times 10^{-5}$ cm$^2$/s. When $t \geq t^*$, the diffusivity reduces by 14.1-fold to $D_2 = 1.90 \times 10^{-6}$ cm$^2$/s. The diffusion coefficient decrease remains insufficient to avoid the quick occurrence of new significant damages to the cell membrane, as seen by three deep $G$ local minima occurring after the transition time. The $G$ local minimum at $t^* = 4.25$ ns, related to the main unfolding event, is indicated by a magenta star.
computational results, while we obtained them independently from these authors.

A perspective for future research would be to use surfactant macromolecules, such as natural detergents, as a component of the extracting solvent. A potential target is saponin that is under investigation for its immunological properties in mammalians and possible integration in vaccine adjuvants (Sun, Xie, & Ye, 2009). Saponin, including molecular species with log $P \geq 5$, is also known to permeabilize cells by leaving holes in their membranes (Jamur & Oliver, 2010), an interesting feature for cell milking. Therefore, due to its large macromolecular size and amphiphatic behavior, we believe that the number of holes created in the membrane by saponin could be limited and better controlled over time with respect to a pure hydrocarbon solvent.

**Supplementary material**

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/073911102.2014.907544. It is composed of Part A: Heat capacity peaks and Part B: Demonstration of Equation (7).

**Acknowledgments**

I thank the Department of Biotechnology – Centre of Excellence (DBT – COE, Indian Government) for this publication as well as JNU’s Vice Chancellor S. K. Sopory and Rector S. Pai. I acknowledge Professor I. Ghosh from JNU as well as Ms Ekta Jain from The Energy and Resources Institute (TERI) in New Delhi for fruitful discussions.

**References**

Adcock, S. A., & McCammon, J. A. (2006). Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chemical Reviews, 106*, 1589–1615.

Allen, M. P., & Tildesley, D. J. (1991). *Computer simulation of liquids*. Oxford: Oxford University Press.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science, 181*, 223–230.

Ben-Naim, A. (2012). Levinthal’s question revisited, and answered. *Journal of Biomolecular Structure and Dynamics, 30*, 113–134.

Bradley, W. D., Willis, R. M., & Seefeldt, L. C. (2011). Biodiesel production by simultaneous extraction and conversion of total lipids from microalgae, cyanobacteria and wild-mixed culture. *Bioresearch Technology, 102*, 2724–2730.

Cheng, Y., Lu, Y., Gao, C., & Wu, Q. (2009). Alga-based biodiesel production and optimization using sugar cane as the feedstock. *Energy and Fuels, 23*, 4166–4173.

Chisti, Y. (2008). Biodiesel from microalgae beats bioethanol. *Trends in Biotechnology, 26*, 126–131.

DePaul, A. J., Thompson, E. J., Patel, S. S., Haldeman, K., & Sorin, E. J. (2010). Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics. *Nucleic Acids Research, 38*, 4856–4867.

Dill, K. A., & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Structural Biology, 4*, 10–19.

Dismukes, G. C., Carrieri, D., Bennette, N., Ananyev, G. M., & Posewitz, M. C. (2008). Aquatic phototrophs: Efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology, 19*, 235–240.

Domanski, J., Stansfeld, P., Sansom, M. S. P., & Beckstein, O. (2010). Lipidbook: A public repository for force field parameters used in membrane simulations. *Journal of Membrane Biology, 236*, 255–258.

Doshi, R., Nguyen, T., & Chang, G. (2013). Transporter-mediated biofuel secretion. *Proceedings of the National Academy of Sciences of the USA, 110*, 7642–7647.

Feller, S. E., Gawrisch, K., & MacKerell Jr, A. D. (2002). Polyunsaturated fatty acids in lipid bilayers: Intrinsic and environmental contributions to their unique physical properties. *Journal of the American Chemical Society, 124*, 318–326.

Feller, S. E., & MacKerell Jr, A. D. (2000). An improved empirical potential energy function for molecular simulations of phospholipids. *Journal of Physical Chemistry B, 104*, 7510–7515.

Gillet, J.-N. (2010). Marked thermal exaltation in hybrid thin membranous nanomaterials covered by stretched nanodots for thermoelectrics and passive cooling. *ACS Applied Materials and Interfaces, 2*, 3486–3492.

Gillet, J.-N. (2013). Ultralow thermal conductivity minimum in silicon supercrystals with weakly and strongly bonded germanium nanodots for thermoelectrics. *ACS Sustainable Chemistry and Engineering, 1*, 152–160.

Gillet, J.-N., Degorce, J.-Y., & Meunier, M. (2005). General model and segregation coefficient measurement for ultra-shallow doping by excimer laser annealing. *Applied Physics Letters, 86*(222104), 1–3.

Gillet, J.-N., & Ghosh, I. (2013). Concepts on the protein folding problem. *Journal of Biomolecular Structure and Dynamics, 31*, 1020–1023.

Gillet, J.-N., & Sheng, Y. (2000). Iterative simulated quenching for designing irregular-spot-array generators. *Applied Optics, 39*, 3456–3465.

Gouveia, L., & Oliveira, A. C. (2009). Microalgae as a raw material for biofuels production. *Journal of Industrial Microbiology and Biotechnology, 36*, 269–274.

Hejazi, M. A., & Wijffels, R. H. (2004). Milking of microalgae. *Trends in Biotechnology, 22*, 189–194.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., & Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: Perspectives and advances. *Plant Journal, 54*, 621–639.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics, 14*, 33–38.

Jamur, M. C., & Oliver, C. (2010). Permeabilization of cell membranes. *Methods in Molecular Biology, 588*, 63–66.

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology, 9*, 646–652.

Kiel, O., Tamimi, A., Thielges, M. C., & Fayer, M. D. (2013). Ultrafast structural dynamics inside planar phospholipid multilayer model cell membranes measured with 2D IR spectroscopy. *Journal of the American Chemical Society, 135*, 11063–11074.

Killian, B. J., Kravitz, J. Y., & Gilson, M. K. (2007). Extraction of configurational entropy from molecular simulations via an expansion approximation. *Journal of Chemical Physics, 127*, 024107.
Ultrafast molecular dynamics of biofuel extraction for microalgae and bacteria milking

Kirckpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*, 671–680.

Kung, Y., Runguphan, W., & Keasling, J. D. (2012). From fields to fuels: Recent advances in the microbial production of biofuels. *ACS Synthetic Biology, 1*, 498–513.

Leach, A. R. (2001). *Molecular modeling. Principles and applications* (2nd ed). Harlow: Prentice Hall, Pearson Education.

Luisi, P. L., Walde, P., & Oberholzer, T. (1999). Lipid vesicles as possible intermediates in the origin of life. *Current Opinion in Colloid and Interface Science, 4*, 33–39.

Man, P., & Nilsson, L. (2002). A molecular dynamics study of tryptophan in water. *Journal of Physical Chemistry B, 106*, 9440–9445.

Martinez, L., Andrade, R., Birgin, E. G., & Martinez, J. M. (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry, 30*, 2157–2164.

Misra, N., Patra, M. C., Panda, P. K., Sukla, L. B., & Mishra, B. K. (2012). Homology modeling and docking studies of FabH (β-ketoacyl-ACP synthase III) enzyme involved in type II fatty acid biosynthesis of *Chlorella variabilis*: A potential algal feedstock for biofuel production. *Journal of Biomolecular Structure and Dynamics, 31*, 241–257.

Moheimani, N. R., Matsura, H., & Borowikzka, M. A. (2013). Non-destructive hydrocarbon extraction from Botryococcus braunii BOT-22 (race B). *Journal of Applied Phycology, 25*, 1653–1661.

Mooijat, M., Foulds, A., Pruvost, J., & Legrand, J. (2008). Optimal selection of organic solvents for biocompatible extraction of beta-carotene from *Dunaliella salina*. *Journal of Biotechnology, 133*, 433–441.

Mosgaard, L. D., Jackson, A. D., & Heimburg, T. (2013). Fluctuations of systems in finite heat reservoirs with applications to phase transitions in lipid membranes. *Journal of Chemical Physics, 139*(125101), 1–8.

Numata, J., & Knapp, E.-W. (2012). Balanced and bias-corrected computation of conformational entropy differences for molecular trajectories. *Journal of Chemical Theory and Computation*, 8, 1235–1245.

Peng, B., Yuan, X., Zhao, C., & Lercher, J. A. (2012). Stabilizing catalytic pathways via redundancy: Selective reduction of microalgae oil to alkanes. *Journal of the American Chemical Society, 134*, 9400–9405.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., ... Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry, 26*, 1781–1802.

Playne, M. J., & Smith, B. R. (1983). Toxicity of organic extraction reagents to anaerobic bacteria. *Biotechnology and Bioengineering, 25*, 1251–1265.

Sun, H. X., Xie, Y., & Ye, Y. P. (2009). Advances in saponins-based adjuvants. *Vaccine, 27*, 1787–1796.

Trevors, J. T. (2002). Hydrophobic medium (HM) water interface, cell division and the self-assembly of life. *Theory in Biosciences, 121*, 163–174.

Xu, X.-H., Zhang, L., & Chen, H.-L. (2011). Screening of biocompatible organic solvents for enhancement of lipid milking from *Nannochloropsis sp.* *Process Biochemistry, 46*, 1934–1941.

Yadugiri, V. T. (2009). Milking diatoms – a new route to sustainable energy. *Current Science, 97*, 748–750.

Yin, D., & MacKerell Jr, A. D. (1998). Combined ab initio/empirical approach for the optimization of Lennard-Jones parameters. *Journal of Computational Chemistry, 19*, 334–338.

Zuckerman, D. M. (2010). *Statistical physics of biomolecules. An introduction*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Appendix

Figures 8 and 9 present the $G$ vs. $t$ evolutions (black thick lines) of the mixtures with n-decane and n-heptane, respectively. In addition, the curves $E$ vs. $t$ and $-TS$ vs. $t$ (shifted by the initial $E$ value) of the $G$ components are depicted for each figure by dashed-dotted blue and dashed red lines, respectively. The thermodynamic behaviors of the three curves in Figures 8 and 9 have to be compared with that of the three curves in Figure 5(a) for the mixture with n-hexane. In Figure 8, the thick black line is the same $G$ vs. $t$ curve than that in Figure 6 for the n-decane solvent. The conformations listed as ‘0’, ‘1’, ‘2’, ‘3’, and ‘4’ in Figure 8 have to be matched with those in Figures 6 and 7. In Figure 9, for the mixture with n-heptane, the $D_1$ to $D_2$ transition time is obtained at $t = t^* = 3.95$ ns. This time (also obtained from our Einstein diffusion law adaptation) corresponds to the time abscissa of the valley middle of the fourth $G$ local minimum counted from $t = 0$. This local minimum is indicated by a star.

The same $E$–$S$ compensation mechanisms that are shown in Figure 5(a) for the mixture with n-hexane occur in Figures 8 and 9 for the two other solvents (after a short initial thermal equilibration). Moreover, we intuitively expected the moderate increase in the $D_1$ and $D_2$ values for the n-heptane solvent compared with those of the n-hexane solvent, which was confirmed by the MD trajectories of the mixtures with n-heptane and n-hexane (Table 1). Consequently, for the three solvents, we expect to obtain similar thermodynamic results if other MD trajectories were calculated instead of those presented in this paper.