Clusters of proteins in bio-membranes: insights into the roles of interaction potential shapes and of protein diversity

Nicolas Meilhac†‡, Nicolas Destainville††
† Université de Toulouse; UPS; Laboratoire de Physique Théorique (IRSAMC); F-31062 Toulouse, France
‡ CNRS; LPT (IRSAMC); F-31062 Toulouse, France

(Dated: January 27, 2013)

It has recently been proposed that proteins embedded in lipidic bio-membranes can spontaneously self-organize into stable small clusters, or membrane nano-domains, due to the competition between short-range attractive and longer-range repulsive forces between proteins, specific to these systems. In this paper, we carry on our investigation, by Monte Carlo simulations, of different aspects of cluster phases of proteins in bio-membranes. First, we compare different long-range potentials (including notably three-body terms) to demonstrate that the existence of cluster phases should be quite generic. Furthermore, a real membrane contains hundreds of different protein species that are far from being randomly distributed in these nano-domains. We take this protein diversity into account by modulating protein-protein interaction potentials both at short and longer range. We confirm theoretical predictions in terms of biological cluster specialization by deciphering how clusters recruit only a few protein species. In this respect, we highlight that cluster phases can turn out to be an advantage at the biological level, for example by enhancing the cell response to external stimuli.

1. INTRODUCTION

Live cells contain a large amount of membranes playing a great variety of biological roles. Membranes are an essential component of Life. Their first function is to separate the diverse cell organelles (e.g. the Golgi apparatus, mitochondria, the endoplasmic reticulum, the nucleus in animal cells) from each other. The plasma membrane, on which this paper primarily focuses, separates the interior of the cell from the environment. However, membranes are not passive barriers but they control in an active way the movement of molecules or information in and out the organelles or the cell, thus maintaining substantial differences of composition between the cell and its environment. Plasma membranes also play an important role in cell adhesion and motility, thus ensuring a large amount of biological functions. The principal constituents of plasma membranes are lipids and proteins, accounting each for typically 50% of the membrane mass, and constituting a two-dimensional complex fluid where interactions are weak, typically on the order of the thermal energy $k_B T$ [1, 2]. Adopting a coarse-grained point of view, lipids can be considered as an underlying fluid (or solvent) in which included proteins evolve. Effective protein-protein interactions mediated by this fluid emerge, because an inclusion creates a perturbation of the membrane that in turn influences neighbor inclusions.

Membrane organization is a topical issue in cell biology. Understanding how hundreds of different protein and lipid species organize in a same membrane to perform as many biological functions [1, 3] remains a challenge to which physicists have recently paid attention, by applying concepts of soft matter to the biophysics of bio-membranes [2, 4]. Far from the pioneer mosaic model of Singer and Nicolson [5], it is now understood that lipids and proteins are non-homogeneously distributed in membranes. They are subtly organized in small domains, or compartments, or clusters [6], of highly variable composition, the length-scale of which ranges from a few nanometers to microns. Heterogeneity is even thought to be a key ingredient for biological functions by concentrating a few protein species in a same nano-domain. Indeed, it facilitates the encounter of different proteins species that must interact to perform together a given biological task [7–10]. When in a same membrane domain, they encounter much more easily than if they were randomly diffusing on the membrane surface [11]. Concentrating identical receptors in a same domain also facilitates the response to small inputs by reducing noise [12]. Grouping together structure proteins is also essential in the context of cell adhesion [13, 14]. A wide range of experimental techniques demonstrate that co-localization of a few different protein species in the same membrane nanodomains is a common feature (see, e.g. [5, 12, 15, 25]) and as stated by T. Lang and S.O. Rozzoli, clusters are “more than a pretty picture” [6].

However, the precise way proteins organize and the physicochemical mechanisms responsible for this organization are currently a matter of controversy [12, 23, 26–30]. In particular, a consensus has not been reached yet on the reason why domains remain sub-micrometric. The ideas developed in this paper propose elements of answers to the two following questions: 1) What causes compartmentalization of membrane proteins in sub-micrometric domains in live cells? 2) How can compartmentalization mechanisms account for segregation of membrane proteins in specialized, heterogeneous sub-micrometric structures where they more easily encounter to perform biological functions?

Motivated by this biophysical context, statistical mechanics arguments have recently been proposed to account for the existence of protein nano- or micro-domains in a membrane at equilibrium [12, 24, 27, 31, 33]. They rely on the following mechanism: attraction at short
range (∼ 1 \text{ nm}) between proteins favors condensation of membrane proteins in a dense phase, but some weaker repulsion at longer range (∼ 10 \text{ nm}) prevents a complete phase separation because once clusters have nucleated, sufficiently large clusters repel each other due to additivity of the repulsion between their proteins. Alternatively, too large clusters are unstable because of the long-range repulsion. The resulting phase at equilibrium is called a “cluster phase” \cite{32}, by analogy with similar phases in soft condensed matter \cite{34,36}. However, we will show that real protein-protein interaction potentials are very complex and potentially very diverse. The present work addresses the question of the robustness of the cluster phases with respect to the potential shape and demonstrates that cluster phases should be generic in bio-membranes in spite of potential diversity.

In addition, previous studies did not take into account protein diversity. Recent theoretical investigations have shown that modulating the short-range attraction in order to account for diversity could lead to composition heterogeneity of clusters because the ensuing energy gain is larger than the corresponding entropic cost \cite{37}. The present work also addresses this question on numerical grounds and confirms these theoretical investigations (see also \cite{38}, Supporting Material).

The paper is thus organized as follows. We recall the biophysical context and present our model in the two first sections. Then Section 4 is mainly devoted to the question of the robustness of cluster phases with respect to the potential shape, more precisely to its long-range repulsion term. In Section 5, we focus on the role of protein diversity through its effects on the short-range attraction. The last section is devoted to discussions and conclusions.

2. BIOPHYSICAL CONTEXT

Interactions between proteins embedded in bio-membranes are manifold and the literature on the topic is extensive. In addition to hard-core and electrostatic interactions, they feel several effective forces specific to these systems, because they are mediated by the membrane. As we shall see, the free energies involved in these interactions are on the order of the thermal energy $k_B T$ and thus play a role at equilibrium.

If the membrane is seen as an elastic sheet of curvature elastic modulus $K$, it is locally deformed by the presence of a protein, whether it be embedded or peripheral, and the response to this constraint implies effective long-range forces. First, thermal undulations of the elastic membrane in the dimension normal to the membrane plane are perturbed by the presence of proteins, in a way that depends on their separation $r$. The entropy, depending on the number of effectively accessible undulation modes, thus depends on separation, and an attractive Casimir-like force ensues, which is pairwise additive and decays as $r^{-4}$ at large $r$ \cite{39,40}. In addition, when proteins are (up-down) asymmetric with respect to the membrane plane, which we assume to be the major- ity case for both peripheral and transmembrane proteins, they can be modeled as conical inclusions with a half-aperture angle (or contact angle) denoted by $\alpha$. This angle models not only the crystallographic shape of the protein \cite{41,42}, but also its local interactions, in particular of electrostatic nature, with the surrounding lipids.

The two lipid monolayers of cell membranes are very different in composition and in charge \cite{1}, which also breaks the up-down symmetry, even for cylindrical inclusions. Such asymmetric inclusions feel an additional repulsion, also decaying as $K r^{-4}$ at leading order. This repulsion, proportional to the squares of the effective contact angles $\alpha$, is due to the elastic deformation imposed to the elastic membrane \cite{39,40,43,44}. Note the important following point: even for differently oriented conical inclusions with contact angles $\alpha_1$ and $\alpha_2$ of opposite sign, this elastic contribution remains repulsive because it is proportional to $\alpha_1^2 + \alpha_2^2$ at leading order. This repulsion compensates the Casimir-like attraction as soon as $\alpha$ is typically larger than $5^\circ$. It is not pairwise additive anymore, since three-body terms, which can be attractive, exist at the leading order $r^{-4}$ \cite{44,45}. But we will demonstrate that they are not strong enough to counter-balance repulsive two-body forces at long range (see also the Supporting Information). As for $n$-body interactions with $n > 3$, they decay faster than $r^{-4}$ \cite{44} and will not be considered in the present work \cite{45}. Furthermore, proteins need not be isotropic inclusions, in this sense that their horizontal section can be better modeled by an ellipse (instead of a disk). In this case, the potential depends on the relative orientation of the ellipses \cite{41,44}, and decays as $r^{-3}$. But except for strong anisotropies, it is averaged over orientations because of the rapid protein rotational diffusion \cite{41}, which significantly lowers the strength of the interaction and makes it decaying as $r^{-3}$ as well. In case of strong anisotropy, the ensuing elastic interaction can even become attractive \cite{41}. In the present work, we only consider weakly or moderately anisotropic inclusions, such that the net elastic interaction is repulsive at long range. We discuss strong anisotropy in the Supporting Information. Beyond anisotropy, protein shape can make interaction potentials even more complex \cite{47}.

Even though these interactions mediated by the membrane should decay algebraically at long range, they are likely to be partially screened for at least two reasons: 1) Long wavelength membrane excitations are damped when the membrane is coupled to the much more rigid cytoskeleton \cite{48}; 2) When weak, but non-vanishing membrane tension is taken into account, the repulsion is screened beyond distances of a few tens of nanometers and thus decays exponentially at long range \cite{49} (see also \cite{31,50} for related calculations). However, estimating the values of membrane tensions in live cells, and therefore of the screening length, is a difficult experimental task. It is thus relevant to consider both exponentially
and algebraically decaying potentials. At shorter range, proteins first experience a hard-core repulsion at contact. In this respect, protein radii \( a \) are variable, ranging from a fraction of a nanometer for single membrane-spanning \( \alpha \)-helices, to a couple of nanometers for large proteins (such as receptors, pumps or channels) \[1\]. As a first step, we consider mono-disperse distributions of protein radii \( a \) in this work, with the typical radius \( a = 2 \) nm. The general case is briefly discussed in the Supporting Information. At contact or close to contact, interactions are of hydrophobic, or (screened) electrostatic, or hydrogen-bond origins, arising from apolar, or polar, or charged amino-acids near the protein surface. Because of the rotational averaging discussed above, we do not take into account the anisotropy of these interactions either. Van der Waals forces (in \( r^{-6} \)) should also play a role at short distances.

Two additional forces mediated by the lipidic membrane play a role when separations between protein surfaces are typically less than one nanometer. The hydrophobic core of transmembrane proteins is likely not to match the equilibrium membrane thickness, what is usually called hydrophobic mismatch. The mismatch is said to be positive when the core is too thick, and negative in the converse case. Two proteins with mismatches of identical (resp. different) sign feel an attractive (resp. repulsive) short-range force due to the induced elastic variations of the bilayer thickness \[15, 51, 54\]. Mismatch has been experimentally demonstrated to promote the aggregation of transmembrane proteins \[55\]. There also exists many-body effects associated with hydrophobic mismatch \[59\]. They are not included in the present study because no simple effective expression is known for these forces.

In addition, attractive depletion forces, due to the two-dimensional osmotic pressure of lipids on proteins, tend to bring them closer when they are about a nanometer apart \[67\]. Hydrophobic and depletion forces have been studied numerically by Molecular Dynamics \[38, 54, 58--60\]. Their range is nanometric and the involved energies are equal to a few \( k_B T \). Attractive or repulsive hydrophobic mismatch modulates this strength by a couple of \( k_B T \) and can consequently change the degree of aggregation (or oligomerization) of proteins (e.g. rhodopsins in \[61\]).

Finally, cell membranes are constituted of a large variety of different lipid species \[62\] and proteins are known to recruit in their neighborhood lipids for which they have a better affinity \[8\]. For two identical proteins, the closer they are, the more energetically favorable the configuration because there is an interface energy associated with this “wetting” phenomenon. Typical energies are also of order \( k_B T \) \[51, 62\]. These references propose a protein-driven mechanism for domain formation invoking such forces, but the limited domain size due to additional repulsive forces has not been discussed in this context.

To conclude, protein interaction potentials have diverse contributions, and writing a generic potential shape is a tedious task appealing to intensive Molecular Dynamics simulations, out of the scope of the present work. For this reason, it is useful to consider different realistic, typical shapes, and to analyze the similarities and the differences between the ensuing cluster phases, as already discussed in reference \[52\]. We focus on two types of interaction potentials: potentials decaying algebraically as \( r^{-4} \) at long range, corresponding to unscreened or partially screened repulsion; and potentials decaying exponentially at long range, corresponding to complete screening. We thus span a large variety of experimental contexts. We demonstrate that conclusions are qualitatively similar in both cases, thus proving that cluster phases of membrane proteins should be generic.

3. MODEL

Here and in the following, all energies are implicitly in units of \( k_B T \), because it is the relevant scale of energies in this context.

The pairwise potentials \( U_2(r) \) considered in this work comprise a hard-core repulsion at very short distances, i.e. \( U_2(r) = \infty \) if \( r < 2a \), an attractive part at short range and a weak repulsive one at larger distances. The attractive part is chosen to be exponential in all cases, because the short-range forces considered above decay rapidly by nature.

![FIG. 1: Examples of pair potentials \( U_2 \) used throughout this paper. From left to right at small \( r \): two algebraic potentials \[equation (3.2)\] with \( K = 150 \) (black) and \( 50k_B T \) (red); and two bi-exponential potentials \[equation (3.1)\] with \( E_{\text{max}} = 0.4 \) (blue) and \( 0.1k_B T \) (green) and \( 1/\gamma_r = 4 \) (blue) and 16 nm (green). \( E_{\text{min}} = -4k_B T \), and \( 1/\gamma_r = 2 \) nm in all cases. In all cases, cluster co-exist with monomers at equilibrium, as illustrated by the four random snapshots (at density \( \phi = 0.1 \), same colors as the corresponding curves).](image)

As motivated above, we will study on the one hand
bi-exponential potentials of the form [32] [64] [65]:

\[ U_2^{\exp}(r) = -\varepsilon_a \exp(-\gamma_a r) + \varepsilon_r \exp(-\gamma_r r), \]  

(3.1)

where all parameters are positive. They contain an attractive and a repulsive term and \( \gamma_r < \gamma_a \) (see examples in figure 1). They correspond to complete screening. Parameter values will be specialized in our numerical studies, in agreement with the biophysical context of the previous Section. As in reference [32], we set \( 1/\gamma_a = 2 \text{ nm}, \) \( 1/\gamma_r = 4 \text{ or } 16 \text{ nm}, \) and \( \varepsilon_a \) and \( \varepsilon_r \) are adapted to fit the desired energy at contact \( E_{\text{min}} \equiv U_2^{\exp}(2a) \) and the desired energy barrier, \( E_{\text{max}}. \) For example, \( 1/\gamma_r = 4 \text{ nm}, \) \( E_{\text{min}} = -4k_B T \) and \( E_{\text{max}} = 0.4k_B T \) lead to \( \varepsilon_a = 55.1k_B T \) and \( \varepsilon_r = 9.4k_B T. \) The maximum value \( E_{\text{max}} \) is then reached at \( r = 9.8 \text{ nm}. \)

On the other hand, we will consider unscreened potentials, which still display an exponential short-range attraction but the repulsive part of which is slowly, algebraically decaying:

\[ U_2^{\text{pow}}(r) = \begin{cases} -\varepsilon_a \exp(-\gamma_a r) + E_a & \text{if } r \leq r_1 \\ C/r^p & \text{if } r \geq r_1 \end{cases} \]  

(3.2)

where the energy shift \( E_a \) enforces the continuity of \( U \) at \( r_1. \) The existence of cluster phases for \( p = 2 \) has been explored in reference [32]. They also exist for \( p = 6 \) (data not shown). Here we focus on \( p = 4 \) because the repulsion between isotropic inclusions in an elastic membrane displays such a long-range behavior, with \( C = S^2\frac{4K}{\pi}(a_1^2 + a_2^2) - \frac{9}{2p^2}, \) where \( S = \pi a^2 \) is the inclusion area, \( K \) is the membrane elastic modulus and \( a_{1,2} \) are contact angles [39] [40]. The first term in the squared brackets corresponds to pure elastic deformations of the membrane and the second one to the Casimir-like forces. The constant \( C \) is positive for sufficiently large values of \( K \) and \( a_{1,2}. \) The values of the contact angles \( \alpha \) are difficult to estimate experimentally. Here, we shall consider two typical values, \( \alpha = 0 \text{ or } 10^\circ, \) so that, with the values of \( K \) considered below, \( C \) is always positive except if \( a_1 = a_2 = 0. \) As above, we set \( 1/\gamma_a = 2 \text{ nm} \) and we choose \( r_1 = 2a + 1/\gamma_a, \) 1/\gamma_r \) being the typical extend of the attraction beyond the hard core distance, 2a. Finally, \( \varepsilon_a \) is set so that \( E_{\text{min}} \) has the desired value (see below). Examples are given in figure 1.

As far as the elastic modulus \( K \) is concerned, the experimental values found in the literature are very variable, because they depend strongly on the lipidic composition, in particular the cholesterol concentration. The typical values of \( K \) range from 10 to 200k_B T [66] [68] and references therein). In this work, we consider the values \( K = 30, 50 \text{ and } 150k_B T, \) thus spanning the experimental interval. Accordingly, if \( \alpha_1 = \alpha_2 = 10^\circ, \) then \( E_{\text{max}} = 0.21, 0.40 \text{ and } 1.34k_B T, \) respectively.

In Section 4 we will also explore the role of long-range three-body forces mediated by the elastic membrane, also decaying as \( r^{-4}: \)

\[ U_3^{\text{pow}} = B \sum_{p < q < r} \left[ \cos(2\gamma_p) + \cos(2\gamma_q) + \cos(2\gamma_r) \right], \]  

(3.3)

where \( \gamma_p \) is the angle \( (r_{pq} - r_{pr}), \) \( r_{pq} = r_q - r_p, \) and so forth, and \( B = S^2\frac{12K}{\pi}(\alpha_1^2 + \alpha_2^2) \) [43] [45]. When wished, \( U_3^{\text{pow}} \) is added to the two-body contributions of equation (3.2). In this work, we only consider two- and three-body forces because they dominate at large separations. The more general case of many-body forces will be addressed in the Discussion.

In order to take protein diversity into account, we also consider in section 5 the case were several protein species are present in the membrane. Forces then depend on the nature of the interacting proteins. On the one hand, we can play on the long-range terms by making \( \varepsilon_r \) as in (3.1) or \( C \) as in (3.2) depend on interacting protein species, through their contact angles \( \alpha. \) On the other hand, we can modulate the short-range attraction by using species-dependent values of \( \varepsilon_a \) and \( \varepsilon_r \) to modulate \( E_{\text{min}} \) at fixed \( \gamma_a, \gamma_r \) and \( E_{\text{max}}. \)

Note that so far, we have used the terminology “potentials or forces at long range”. However, our potentials are not long-ranged in the strict statistical mechanics sense of the word because, even though their range extends beyond several tens of nanometers, the integral \( \int rU(r)dr \) is convergent at large \( r. \) However, we shall keep the terminology “long-range” in the following, by opposition to “short-range”.

We carry out Monte Carlo simulations in the canonical ensemble as prescribed in reference [32]. More details can be found in the Supporting Information. Our systems contain \( N \geq 100 \) particles that interact via the potentials discussed above, and evolve in a two-dimensional continuous medium of area \( A \) representing the lipidic “sea” (the solvent). Boundary conditions are periodic. The hard-core diameter is chosen as \( d_0 = 2a = 4 \text{ nm}, \) the typical diameter of a protein of average molecular weight [1]. Figures 1, 6, 9 and S2-3 display several simulation snapshots for various parameter sets. The distributions of cluster sizes \( P(k) \) displayed below are defined as

\[ P(k) = \langle N_k \rangle / \sum_{k=1}^{\infty} \langle N_k \rangle, \]  

(3.4)

where \( \langle N_k \rangle \) is the measured average number of clusters of size \( k. \) The mean cluster size is defined as \( \langle k \rangle = \sum kP(k). \) It is also equal to the number of particles, \( N, \) divided by the number of clusters (including monomers).

4. Long-range Potentials

In this section, we focus on the role of long-range repulsive potentials on cluster phases. As motivated above, we first compare two kinds of pairwise potentials: exponentially [equation (3.1)] and algebraically [equation (3.2)]
decaying ones. In a second time, we study the role of three-body forces at long range [equation (3.3)]. Finally, we focus on the implication of protein diversity as far as long-range forces are concerned. More precisely, we take into account the fact that a macroscopic fraction of up-down symmetric proteins might not be concerned by the elastic long-range repulsion.

4.1. Comparison between two kinds of long-range pair potentials

We first compare exponentially decaying forces and algebraically decaying ones ($U(r) \sim r^{-\lambda}$), associated with the repulsion due to the elastic deformations of the membrane. The $N$ proteins are identical. Bi-exponential potentials have already been numerically studied in detail in references [32] [64] [65]. The main feature of the ensuing cluster phases is that cluster-size distributions $P(k)$ appear to be bimodal for broad ranges of parameters and concentrations: large, dense clusters co-exist with a gas of monomers. Figure 8 below will provide related distributions in figure 2 (we have nevertheless observed that a bi-modal regime is restored for large bending moduli $K$ in figure 2 that more than 25% (resp. 75%) of the proteins dwell in clusters containing more than $k = 10$ particles as soon as $\langle k \rangle \geq 3$ (resp. $\langle k \rangle \geq 7$). Independently of a precise definition of a cluster phase as in [37] [69], this indicates that a macroscopic fraction of proteins live in assemblies, which is the biological mechanism we are primarily interested in.

The proportionality regime is still being observed, even though on shorter concentration ranges (see figure 2).
At the highest density studied, $\phi = 0.2$, average cluster sizes $\langle k \rangle$ range from 4 to 8 for the parameter sets studied and clusters of size up to $k = 20$ are commonly observed. Note that lowering $E_{\text{min}}$ at fixed $K$ (i.e., increasing the attraction) should increase $\langle k \rangle$ \cite{33}. Unfortunately, due to the long-tail distribution, increasing $\langle k \rangle$ requires increasing $N$, which is rapidly limiting in terms of computational cost. However, we have studied the effect of varying $|E_{\text{min}}|$ in the $K = 50k_B T$ and $\phi = 0.05$ case (figure 2, bottom). The power-law-like regime extends as $|E_{\text{min}}|$ grows and $\langle k \rangle$ increases concomitantly (see Inset).

To conclude, this numerical study indicates that replacing and exponentially decaying repulsion at long range by an algebraic one does not alter the existence of cluster phases for wide ranges of parameters, thus demonstrating that they should be robust with respect to the precise shape of interaction potential, provided that it displays both a short-range ($\sim 1 \text{ nm}$) attraction and a longer-range ($\sim 10 \text{ nm}$), weaker repulsion.

To finish with, we compare our data to available experimental ones. Clusters of bacteriorhodopsin (BR) in proteoliposomes can be observed by freeze-fracture electron microscopy. This technique makes possible the determination of cluster-sizes and of their distributions with a correct accuracy \cite{70} (see Table I and figure 4 in this reference). We assume a BR radius of 1.7 nm \cite{70} to convert numbers of BR per $\mu \text{m}^2$ to densities $\phi$. The algebraically decaying potential \cite{32} has been chosen for simulations, with $K = 150k_B T$, $1/\gamma_a = 2 \text{ nm}$ and $E_{\text{min}} = -4k_B T$. With these parameter values, a good qualitative agreement is obtained, as displayed in figure 3. Cluster-size distributions are bimodal in both cases, with a monomer peak and clusters containing a few proteins (compare both our green curve in figure 7 and figure 4 of \cite{70}). This good qualitative agreement supports our numerical approach. Similar cluster sizes ($k \approx 10$) have been observed at lower protein density but after photoactivation of BR \cite{71}.

### 4.2. Role of three-body elastic interactions

It has been proved \cite{44,45} that the elastic energy for isotropic inclusions is dominated at long range by the sum of the two-body terms of equation (3.2) considered so far and the three-body ones of equation (3.3). Here we take these three-body terms into account in simulations. Contrary to two-body forces, we did not include any cut-off at short distances, the hard-core repulsion playing this role. Note that in this paragraph, $E_{\text{min}}$ still refers to the two-body energy at contact of $U_2^{\text{low}}(2a)$ [equation (3.2)].

Simulating three-body forces consists of an important increase of the computational complexity, which grows from $O(N^2)$ to $O(N^3)$ operations per Monte Carlo step. For this reason, systems with $N > 100$ can hardly be simulated and we could perform a few simulations only, in order to check that cluster phases still exist in this case and thus that two-body repulsion dominates three-body forces, which can be attractive or repulsive (see also the Supporting Information).

The principal features observed in the previous subsection are conserved. Cluster phases appear to be stable at equilibrium and their distributions display a power-law-like regime until a maximum size. Examples are provided in figure 4. When increasing the density, the mean cluster-size $\langle k \rangle$ grows, even if we do not have enough points at our disposal to conclude that a proportionality regime still exists. Lowering $E_{\text{min}}$ at fixed $K$ also increases $\langle k \rangle$.

However, some differences are also noticeable. First, higher values of the energy at contact $E_{\text{min}}$ are required...
to obtain distributions comparable to those of the previous paragraph. Their are close to $-1.5k_BT$ instead of $-4k_BT$. This indicates that three-body forces have an attractive contribution, on average, as already noticed in reference [3] and in the Supporting Information. In addition, at high $\phi$, simulations show that a big cluster, containing a large fraction of the $N$ particles, can transiently appear and subsequently disintegrate into smaller clusters. As a consequence, one distribution displays a secondary peak at large cluster sizes in the figure, for $\phi = 0.2$, $E_{\text{min}} = -1.5k_BT$ and $N = 150$ (whereas it is absent if $N = 100$). This might be the signature of a transition between a cluster phase and a liquid-gas coexistence phase when increasing $\phi$. To clarify this point, simulating larger systems will be required. But again, simulating three-body forces is expensive and our statistics are poorer than in the previous sections.

4.3. Modulation of long-range potentials

We now explore the consequences of protein diversity in terms of long-range forces, for both kinds of pair potentials (bi-exponential, equation (3.1), and algebraic, equation (3.2)). Examples of modulated potentials are displayed in figure 5.

![FIG. 5: Algebraic pair potentials $U_2^{\text{pow}}$ (three upper curves) and bi-exponential ones $U_2^{\text{exp}}$ (lower curves) modulated at long range. This modulation takes into account the fact that cylindrical inclusions (in red) are up-down symmetric ($\alpha_R = 0$) while conical ones (in blue) are asymmetric ($\alpha_R = 10^\circ$), as in the previous paragraphs. Here $E_{\text{min}} = -4k_BT$ and $1/\gamma_c = 2$ nm for all curves, $K = 50k_BT$ for the algebraic repulsion, $E_{\text{max}} = 0.1k_BT$, 0.05$k_BT$ and 0 for the three bi-exponential potentials, respectively (and $1/\gamma_v = 16$ nm).

In the algebraic case, the modulation is ensured by taking $\alpha = 0$ (cylindrical inclusions thereafter) instead of $\alpha = 10^\circ$ (conical inclusions) when calculating the constant $C$. This constant is essentially divided by 2 for large $K$ values when a conical and a cylindrical inclusion interact. It becomes negative for two cylindrical inclusions, which feel a mutual attraction because of Casimir forces. As for bi-exponential potentials, we adapt both $\varepsilon_a$ and $\varepsilon_r$ so that $E_{\text{min}}$ remains unchanged whereas $E_{\text{max}}$ is divided by 2 for cylindrical-conical interactions and $E_{\text{max}} = 0$ for cylindrical-cylindrical ones. In the latter case, we simply set $\varepsilon_r = 0$ and adapt $\varepsilon_a$ to set the value of $E_{\text{min}}$. The interaction is everywhere attractive. We denote by $N_{\text{cyl}}$ and $N_{\text{con}}$ the numbers of cylindrical and conical particles and $N = N_{\text{cyl}} + N_{\text{con}}$. We compare systems with a finite cylindrical-inclusion fraction $x_{\text{cyl}} \equiv N_{\text{cyl}}/N$ to systems of $N$ identical conical particles, as above.

Figure 6 provides simulation snapshots. For small cylindrical-inclusion fractions $x_{\text{cyl}}$, they are very similar to the mono-particle case as studied above. Figure 7 shows examples of cluster-size distributions $P(k)$ that are weakly affected by the presence of a small amount of cylindrical inclusions, at fixed $\phi$. Consequently, proportionality regimes are also preserved (figure 8 bottom).

To interpret these numerical observations, we discuss them in the framework of the liquid-gas transition of binary mixtures, even though in the present context of cluster phases, transitions are certainly not true thermodynamic transitions but crossovers [37]. We see a cluster as a liquid droplet co-existing with gas. In the liquid phase, conical and cylindrical inclusions are miscible because they essentially feel the same short-range attraction. This miscibility is visible on simulation snapshots (figure 6). However, the attraction is slightly stronger for cylindrical inclusions. Their second virial coefficient $B_2 = \pi \int_0^\infty r^2[1 - \exp(-U(r))]dr$ is thus slightly weaker, and the transition occurs at a lower density. Consequently, clusters are bigger and enriched in cylindrical inclusions, as observed. The average fraction $x_{\text{cyl}}$ of cylindrical inclusions within clusters is $> x_{\text{cyl}}$. However, conical inclusions remain the majority in clusters and too large clusters are unstable, in spite of the cylindrical-inclusion fraction. The presence of a minority of such particles does not destabilize the cluster phase.

However, we anticipate that there exists a value of $x_{\text{cyl}}$ above which this mechanism will cease being valid. To test this hypothesis, we increased the fraction of cylindrical inclusions up to $x_{\text{cyl}} = 0.4$. Figure 8 shows our results for bi-exponential potentials, for both $N = 100$ and 200 particles. As $x_{\text{cyl}}$ grows, the cluster peak drifts to larger sizes in the bimodal distributions for $N = 100$. But the $N = 200$ simulations indicate that this large peak is in fact the superimposition of two peaks for $x_{\text{cyl}} > 0.3$. And the inspection of simulation snapshots indicates that there are indeed two kinds of clusters: mixed clusters as in the low $x_{\text{cyl}}$ regime, the size distribution of which matches the low $x_{\text{cyl}}$ one; and one larger assembly that is highly enriched in cylindrical inclusions. This indicates the existence of a saturation phenomenon: beyond $x_{\text{cyl}} \approx 0.3$, additional cylindrical inclusions “precipitate” in a single large assembly with rare conical proteins. Note however that this large assembly is not stable: one observes in simulations that this cluster can disintegrate and be reformed somewhere else. This instability seems...
FIG. 6: Top: Snapshot of a cluster phase with $N_{\text{cyl}} = 20$ cylindrical inclusions (in red) and $N_{\text{con}} = 80$ conical ones (in blue), in the case of bi-exponential pair potentials as in figure 5. The density is $\phi = 1/50$. Middle and bottom: Snapshots of cluster phases with $N_{\text{cyl}} = 10$ cylindrical inclusions and $N_{\text{con}} = 90$ conical ones, in the case of the pairwise algebraic repulsion with $K = 150k_B T$, $E_{\text{min}} = -4k_B T$ and $1/\gamma_a = 2$ nm, for two different densities $\phi$ and two different cylindrical inclusion numbers $N_{\text{cyl}}$ ($N = 100$ particles in all cases). The green distribution is (weakly) bimodal.

5. MODULATION OF SHORT-RANGE POTENTIALS AND SPECIALIZATION OF CLUSTERS

We now focus on the role of short-range potential modulations in order to explore further the role of protein diversity. One of our major goals here is to validate on numerical grounds the analytical predictions of reference [37] concerning protein cluster specialization: for a sufficiently large affinity difference between same- and distinct-family proteins, clusters become heterogeneous and contain essentially one protein family. We indeed consider systems with $q$ families of proteins. Each family represents proteins, not necessarily identical, that have a preferential affinity $E_{\text{min}}$ at close range. There are typically $q = 10^3$ families in a real membrane. The energy at contact of two particles in a same family $i$, called same-family particles hereafter, is denoted by $E_{i,i,\text{min}} < 0$. The energy at contact of particles of two distinct families $i$ and $j$ (distinct-family particles) is denoted by $E_{i,j,\text{min}} \geq E_{i,i,\text{min}}$.

In this section, we focus on bi-exponential potentials [equation (3.1)] because large clusters are more abundant in this case due to their bimodal distribution. This provides better statistics when computing Binder cumulants below. We consider a single density, $\phi = 1/50$, and the most symmetric case where the $q$ families contain the same number $M = N/q$ of particles. All same-
family energies at contact, \(E_{i,i,\text{min}}\), are also identical, as well as distinct-family ones, \(E_{i,j,\text{min}}\). The energy barrier, \(E_{\text{max}}\), is unchanged and energies at contact are set by tuning \(\varepsilon_a\) and \(\varepsilon_r\) in (3.1), as explained in Section 2. Following [37], we introduce a Flory parameter \(\chi \approx 6(E_{i,j,\text{min}} - E_{i,i,\text{min}}) \equiv 6\Delta E\) [78], illustrative of the difference of affinity between same- and distinct-family proteins. In this reference [37], it was demonstrated that there exist a threshold value \(\chi^c\) such that if \(\chi < \chi^c\), particles of different families are mixed within clusters, whereas if \(\chi > \chi^c\), they are demixed and clusters contain predominantly one family. Demixing is due to a competition between entropy of mixing and contact energy. Note that this transition (as well as the one at \(\phi^c\)) is certainly a simple crossover and not a true thermodynamical transition because of the finite cluster size [37].

The theoretical prediction is \(\chi^c = 2\) for \(q = 2\) and

\[
\chi^c = 2 \frac{q - 1}{q - 2} \ln(q - 1) \approx 2 \ln q \quad (5.5)
\]

for \(q \geq 3\) (see also [72]). The great interest of equation (5.5) lies in the fact that \(\chi^c\) grows like \(\ln k\) and \(\Delta E\) remains of the order of \(k_B T\) even if \(q\) is large, making this scenario physically reasonable to account for cluster specialization [37]. This is the reason why we wish to validate this result on numerical grounds.

![Graph](image)

**FIG. 8:** Top: Distributions of cluster sizes in the case of bi-exponential pair interaction potentials as in figure 5, with density \(\phi = 1/50\), and various cylindrical inclusion fractions \(x_{\text{cyl}}\). Comparison between \(N = 100\) (solid lines) and \(N = 200\) (dotted lines) for each value of \(x_{\text{cyl}}\). Bottom: resulting mean cluster-size as a function of the density, exhibiting a proportionality regime.

![Graph](image)

**FIG. 9:** Snapshots of cluster phases with \(q = 2\) families of 50 particles each (blue and green; \(N = 100\)), in the case of bi-exponential pair potentials with distinct-family energy at contact \(E_{i,j,\text{min}} = -3.9k_B T\) (top, mixed clusters) and \(-3.2k_B T\) (bottom, demixed clusters). In both cases, the same-family energy at contact is \(E_{i,i,\text{min}} = -4k_B T\), and the density is \(\phi = 1/50\).

Figure 9 shows two simulation snapshots in qualitative agreement with this prediction. We also simulated systems with up to \(q = 5\) families, showing the same agreement. To quantify further these qualitative observations, we determine the threshold value \(\chi^c\) by use of (modified) Binder cumulants, an efficient tool to detect phase transitions numerically [73]. Our system being equivalent to a mean-field \(q\)-Potts model [37], we naturally use...
the following order parameter:

\[ m = \left| \sum_{j=1}^{q} x_j \omega^{j-1} \right|, \]

(5.6)

where \( x_j \) is the fraction of family \( j \) proteins in a cluster and \( \omega = \exp(2i\pi/q) \) is the \( q \)-th root of unity. The (modified) Binder cumulant is then defined as

\[ B_k = \frac{(m^4)_k}{(m^2)_k^2}, \]

(5.7)

where averages are computed on clusters of size \( k \). Figure 10 displays the numerical values of \( B_k \) for \( q = 2 \) as a function of \( k \) and \( |E_{i,j,\min}| \). All curves cross at \( E_{c,\min} = -3.56k_B T \), which is the signature of a phase transition for finite-size systems (the clusters in the present case) \[73\]. The corresponding value of the Flory parameter is \( \chi_c = 2.64 \), which compares well to the theoretical value, \( \chi_c = 2 \).

![Figure 10: Modified Binder cumulants \( B_k \) for \( q = 2 \) families. Same parameters as in figure 9.](image)

We followed the same procedure up to \( q = 5 \), by simulating systems containing \( N = 50q \) particles \((M = 50)\). The measured threshold values are listed in Table I and compared to theoretical predictions. The agreement improves as \( q \) grows, thus supporting the theoretical approach of reference \[37\] (in principle valid in the large \( q \) limit). To ascertain that the measured thresholds were not corrupted by finite size effects, we also simulated a system of \( N = 210 \) particles for \( q = 3 (M = 70) \). We find \( E_{c,\min} = -3.47k_B T \), in satisfactory agreement with the \( M = 50 \) result above. In addition, cluster size distributions for \( N = 150 \) or \( N = 210 \) are undistinguishable.

![Cluster size distributions \( P(k) \). Same parameters as in figure 9.](image)

Cluster size distributions \( P(k) \) are qualitatively unchanged when tuning \( |E_{i,j,\min}| \), as illustrated in figure 11 (we get similar results for \( q \geq 3 \)). However, when decreasing this parameter or the number of families \( q \) at fixed \( |E_{i,j,\min}| \), the typical cluster size decreases because the average short-range attraction decreases. In a similar way, the fraction of monomers increases slowly when decreasing \( |E_{i,j,\min}| \) (with no apparent singularity at \( |E_{i,j,\min}| \)). Note that cluster sizes used to compute Binder cumulants above (figure 10) were chosen near the maximum of \( P(k) \) to ensure significant statistical sampling. All these numerical observations confirm our previous theoretical predictions \[37\].

6. DISCUSSION AND CONCLUSION

Our first major goal was to show that cluster phases are robust with respect to the precise shape of interaction potentials, provided that they display both a short-range \((\sim 1 \text{ nm})\) attraction with a contact energy of a few \( k_B T \) and a longer-range repulsion \((\sim 10 \text{ nm})\), with an energy barrier on the order of \( k_B T \) or lower. We have focussed on two limiting potential shapes: a full, unscreened \( r^{-4} \) repulsion mediated by the elastic membrane; and a totally screened repulsion, decaying exponentially at long range. In the regime of parameters studied, clusters of size ranging from a few to several tens of particles co-exist with monomers above a critical protein density. We have also demonstrated that taking into account three-body forces propagated by the membrane, which have the same order of magnitude as two-body ones, does not destabilize clus-
ter phases. This suggests that cluster phases are generic in the cell membrane context.

Our second core objective was to explore how protein diversity affects cluster phases. In this respect, we have modeled diversity by playing on interaction potentials, both at long and short range. At long range, we have proved that the introduction in the system of a minority (\(\lesssim 30\%\)) of up-down symmetric inclusions, experiencing mutual attraction instead of repulsion, does not destabilize cluster phases. At short range, we have confirmed previous analytical calculations [37]: modulating interactions favors the segregation of proteins of different families in distinct clusters, provided that proteins of a same family have a slightly higher affinity at contact than proteins of distinct families. Accordingly, the high protein concentration of a membrane (\(\sim 50\%\) in mass) favors the co-localization of proteins of a same family in same clusters. Were the concentration be too low, proteins would essentially behave as a gas of monomers diffusing freely on the membrane. It would take a long time for them to encounter their partners in the membrane by diffusion. Clusterization and crowding due to high concentration thus favors interactions between them, which should be of biological relevance in terms, e.g., of faithful response to external stimuli [12] or of cell adhesion [14].

The present work intends to add some original contributions to a more ambitious program: deciphering cell membrane organization on physical grounds. However, even though we have clarified several points, we have been led to make some assumptions, some of which we intend to discuss now (see also the Supporting Information). The objective of the present work is not to be a definitive answer to the question of membrane organization, but rather to lay the foundations of an emerging scenario.

We have explained why the effective forces propagated by the elastic membrane are dominated at large separations by two- and three-body forces. But sub-dominant, higher-order forces exist that might play a role at intermediate distances and thus influence cluster stability. However, we conjecture that cluster phases will remain stable in this case, because we have shown that the long-range elastic interaction remain repulsive at large separation. Accordingly, distant clusters will feel a repulsive force that will prevent their further coalescence in larger clusters. To definitely clarify this point, it would be necessary to fully take into account many-body contributions, which requires inverting a \(3N \times 3N\) matrix and to calculate its determinant at each Monte Carlo step [44] [46]. For moderate system sizes, intensive simulations on parallel computers might be able to address this question in a near future.

Even though a live cell is out-of-equilibrium due to energy consumption, experiments done \textit{in vitro} after membrane removal from cells still reveal clusters [12]. Since in this context, as in the one of artificial vesicles [15], no active, energy-consuming processes can limit cluster size, the existence of finite clusters at equilibrium has to be understood. The mechanism explored in this study partakes of this approach. However, cluster phases properties, notably cluster-size distributions, are likely to be modified by active processes. Endocytosis/exocytosis and membrane recycling are likely to enhance clusterization because they fragment large assemblies of proteins. Furthermore, protein conformations change when they are activated [74] [75], thus modifying their coupling with the membrane and consequently their interaction parameters. For example, photo-activation of BR in reconstituted proteoliposomes has been shown to promote the formation of clusters [71]. Such ingredients could be incorporated in our simulations, e.g. by modeling proteins as systems with two internal states, the interaction parameters of which depend on their internal state. The perturbation of equilibrium properties could then be quantified.

To clarify these points, our predictions need to be validated at the experimental level. As already pointed out [37], Förster Resonance Energy Transfer (FRET) [76] is adapted to quantify the typical distance between fluorescently labeled proteins, and thus their degree of aggregation [15] [51]. By modulating physical parameters such as the inclusion asymmetry, the membrane rigidity or the hydrophobic mismatch [38] [55], it would in principle be possible to probe this degree of aggregation in model membranes such as vesicles [15] or stacked supported membranes [77]. Fluorescence Correlation Spectroscopy (FCS) [76] should have similar capabilities, since it quantifies the density of entities diffusing in a membrane, a cluster essentially counting as a single entity.

### Supporting Information available

I. Monte Carlo procedure; II. Poly-disperisty of protein diameters and contact angles; III. On large cluster stability; IV. Comparing two- and three-body interactions for asymmetric inclusions; V. Rotational averaging of elastic interactions for anisotropic inclusions. This material is available free of charge via the Internet at [http://pubs.acs.org](http://pubs.acs.org).

---

[1] Alberts, B.; \textit{et al.} Molecular Biology of The Cell., 5th edition \textbf{2008}, Garland Science, New York.

[2] Phillips, R.; Kondev, J.; Theriot, J. Physical Biology of The Cell. \textbf{2009}, Garland Science, New York.

[3] Dowhan, W.; Bogdanov, M. in Biochemistry of lipids, lipoproteins, and membranes, eds. Vance, D.E.; Vance J.E. (4th Edn.) \textbf{2002}, Elsevier, Amsterdam.

[4] Lenne, P.F.; Nicolas, A. \textit{Soft Matter} \textbf{2009}, 5, 2841.
FIG. 12: FIG. FOR TOC.