Unravelling hierarchical levels of structure in lipid membranes

Alexandre Blanco-González
University of Santiago de Compostela

Rebeca García Fandiño
USC

Angel Píñeiro (✉️ Angel.Pineiro@usc.es)
Dept. Física Aplicada, Universidad de Santiago de Compostela

Article

Keywords: Lipid bilayer, cell membrane, hierarchical levels of structure, molecular 29 dynamics simulations

Posted Date: February 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1287323/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Unravelling hierarchical levels of structure in lipid membranes

Alexandre Blanco-González1,2,3, Rebeca García-Fandiño2,* Ángel Piñeiro1,*

1Departamento de Física Aplicada, Facultade de Física, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain. 2Departamento de Química Orgánica, Center for Research in Biological Chemistry and Molecular Materials, Universidade de Santiago de Compostela, Campus Vida s/n, E-15782 Santiago de Compostela, Spain. 3MD.USE Innovations SL, Edificio Emprendia, 15782 Santiago de Compostela, Spain.

*Corresponding authors: rebeca.garcia.fandino@usc.es, Angel.Pineiro@usc.es

Abstract:

In analogy with the hierarchical levels typically used to describe the structure of nucleic acids or proteins, and keeping in mind that lipid bilayers are not just mere envelopers for biological material but directly responsible for many important functions of life, we propose and discuss here how membrane models can also be interpreted in terms of different hierarchies in their structure. Namely, lipid composition, interaction between leaflets, existence and interaction of domains arising from the coordinate behavior of lipids and their properties, plus the manifest and specific perturbation of the lipid organization around macromolecules embedded in a membrane are hereby used to define the primary, secondary, tertiary and quaternary structures, respectively. Molecular Dynamics simulations are used to illustrate this proposal. The new paradigm arising from such description is expected to have significant consequences on
deciphering the underlying factors governing membranes and their interactions with other molecules.

Keywords: Lipid bilayer, cell membrane, hierarchical levels of structure, molecular dynamics simulations.
1.- Introduction

What Rosalind Franklin, James Watson and Francis Crick experienced when they first got a glimpse of the DNA 3D structure\textsuperscript{1,2}, or when John Kendrew captured what was coming out of myoglobin RX into his sausage model\textsuperscript{3,4}, was surely comparable to what a 20 diopters-myopic person feels when putting on glasses for the first time. Those events changed the logic of biology because the way in which we currently view DNA or proteins affects how we think about their function. This has a significant impact on hypotheses development, experimental design, and data interpretation.

The three-dimensional structure of large biological molecules such as proteins and nucleic acids is complex and confusing to the naked eye. It is not easy to describe them accurately using just words. Because of this, it is convenient to introduce new concepts. A series of hierarchical levels –typically referred as primary, secondary, tertiary, and quaternary structures– are commonly employed to describe biological macromolecules. These different “organization levels” exhibit regularity and predictability for natural phenomena represented in diverse ways, thus facilitating scientific inquiry and hierarchical causal explanations of different biological processes.

In proteins and nucleic acids, the primary structure is associated to the covalent chemical structure. It usually consists of linear arrays of covalently bound building blocks with a common scaffold and a variety of chemical substitutions. The number of different building blocks is restricted to five in typical nucleic acids (including RNA), and to twenty in proteins. Thus, the primary structure can be specified by the linear order, or sequence, of these building blocks. The secondary structure refers to regular patterns of interaction between adjacent or facing residues. The most illustrative examples are alpha-helices or beta-sheets observed in many proteins and double helices found in virtually all
nucleic acids. The *tertiary structure* refers to the location of the atoms in the three-
dimensional space, taking into consideration geometrical and steric constraints. It usually
provides the starting point for studies that attempt to correlate structure and function. In
proteins it comes from the folding of the secondary structure into distinct arrangements
known as domains, whereas the tertiary arrangement of DNA's double helix in space
includes B-DNA, A-DNA, and Z-DNA. *Quaternary structure* describes the assembly of
individual molecular units into more complex arrays. The simplest example of quaternary
structure is a protein that consists of multiple subunits, which are not connected to each
other by covalent bonds and may be identical or different. In nucleic acids, it refers to
interactions with other molecules, for example in the form of chromatin which leads to
its interactions with the histone small proteins.

This hierarchical representation in four structural layers of organization is extremely
robust and well established empirically for proteins and nucleic acids but it quickly loses
its significance when applied throughout the whole spectrum of biological entities.
Different macromolecules and supramolecular structures, such as polysaccharides and
lipid membranes, exhibit clear three-dimensional shapes and are built out of smaller
subunits. However, different levels of structure have not been defined in these cases.
Polysaccharides are quite similar to proteins in that they are chains of repeated building
blocks (monosaccharides) that acquire well-defined three-dimensional topologies. The
case of cell membranes requires a more detailed explanation. According to the fluid
mosaic model proposed by Singer and Nicolson in 1972\(^5\), the cellular plasma membrane
is a two-dimensional liquid, in which lipid and protein molecules are mixed like a mosaic
and undergo thermal diffusion. This model is still widely accepted by the research
community. Under this approach all the molecules are expected to be homogeneously
distributed throughout the plasma membrane, undergoing simple Brownian diffusion.
However, an increasing number of evidence showing that this is not the case is being found. Like proteins, nucleic acids, and even sugars, lipid membranes can be also described using a hierarchical organization in different levels of structure. The new paradigm arising from such description is expected to help to understand the underlying factors governing membranes and their interactions with other molecules. This is especially important since lipid membranes do constitute one of the most important parts of any living cell, acting as the main barrier between the cytoplasm and its surrounding environment. Their function is not limited to a mere division between the inside and the outside of cells, but it is extended to a wide range of processes, such as serving as anchor points and modulators for signaling and transport proteins, or playing a key role in cell motility. Lipid membranes are also known to be subject to substantial attacks and changes upon disease related processes. For instance, some families of pathogens such as enveloped viruses –HIV and SARS-CoV-2, among many others– are known to enter a cell by fusing their membranes with the host’s. Once internalized and just before the viral reproduction cycle is completed, they leave the cell by selectively ripping apart the infected cell’s membrane, making it their own and leaving the host’s cell heavily modified in its lipidic composition. This modification is hypothesized to sometimes be able to trigger an autoimmune response in the diseased host leading to chronic inflammation, which might have fatal consequences. On the other hand, cancerous processes are also known to display significant alterations on the membranes of the affected tissues.

While healthy mammalian somatic cells present an asymmetry in their composition depending on the leaflet, leaving most of the negatively charged lipids towards the cytoplasm; cancer cells tend to lose this property providing them with negative charges on their exterior, a characteristic that share with bacterial cells. All of this illustrates the important role that lipid membranes play as targets both to invasive pathogens and to...
potential therapies for diseases derived from such processes. It is also reasonable to think that they play a role in some chronic autoimmune responses\textsuperscript{20,21}. Therefore, a proper understanding of the underlying factors governing membranes and their interactions with different species is mandatory to obtain rational criteria from which then derive effective and efficient solutions to the problems that they are involved in. Characterizing living membranes is not easy, as they are densely packed fluids with literally thousands of different components, and their compositions vary greatly depending on the type or source of the cell. Even with much simpler models, empirical study can be cumbersome and there are great difficulties in reaching the resolution needed to get to the atomic or quasi-atomic levels at which some important processes happen. Therefore, computational Molecular Dynamics (MD) simulations are an excellent choice to develop and analyze models of such systems\textsuperscript{22}.

In the present work, a set of hierarchical levels of structure are proposed for lipid membranes, in line with those widely recognized for other biomolecular structures such as proteins and DNA. Original MD simulations results of a large POPC lipid bilayer at coarse grained (CG) resolution and of several macromolecules embedded in an atomistic POPC membrane from a previous work are employed to illustrate the proposal. We aim to introduce an original perspective for the description of lipid membranes, inspired by what Franklin, Watson and Crick did with DNA and Kendrew did with proteins. Hopefully, this new approach will contribute to understand better the behavior of membrane models as well as of actual biological envelopes based on lipids.
2.- Results & Discussion

The different hierarchical levels of structure proposed to reinterpret membrane models are described here in detail. As explained in the introduction, MD simulations of a large POPC membrane at CG resolution as well as atomistic simulations of different macromolecules embedded in a DOPC membrane taken from a previous work are used to illustrate the new concepts. The large homogeneous membrane was chosen to show the impact of the system size in some of the analyzed properties and to prove that even a membrane with a single component exhibits the proposed levels of structure. The atomistic simulations of more complex systems were employed to show the highest level of structure, as explained later.

Figure 1.- Schematic representation of primary (A), secondary (B), tertiary (C) and quaternary (D) levels of structure of a lipid bilayer. Each color in A represents a different lipid type (just some examples, POPC, POPE, DLPC, DOPC are represented). See text for details.
2.1.- The **primary structure** of the lipid bilayer: a “sequence” of thousands of non-covalent basic units

In our analogy with the hierarchical representation of four structural layers of organization in proteins and nucleic acids, the **primary structure** (Fig. 1A) of lipid membranes is identified with the composition of each leaflet, that is different in asymmetric bilayers. Variations in headgroups and aliphatic chains allows the existence of more than 1,000 different lipid species just in a single eukaryotic cell. As explained in the introduction section, a number of infections and diseases have a serious impact on the lipid composition of the host cells. This lipid composition can be employed as a target for the actuation of the immune system and also for artificial therapeutic treatments. Actually, all living organisms have a series of short, cationic and amphipathic peptides known as Host Defense Peptides (HDPs) which are able to recognize and disrupt pathological and pathogenic membranes based on their composition. These HDPs are an important part of the innate immune system. This illustrates the huge importance of the lipid composition of biological membranes, and it justifies that it is identified with the primary structure of the lipid bilayer. However, there are clear discrepancies in this analogy with respect to the primary structure of nucleic acids and proteins. The most obvious difference is that the lipids in a membrane are not covalently bound, in contrast to the nucleotides in DNA or RNA and to amino acids in proteins. Another evident difference is that lipid membranes are locally flat, even though they can form closed quasi-spherical compartments with significant defects or protuberances. As a consequence of the previous two features, membranes have lipids diffusing throughout them, incorporate other molecules into their structure, fuse with other membranes and exhibit a high ability to adapt to environmental changes by altering their composition. Lipid membrane composition alterations take place as a
response to almost any environmental change: pH, temperature, diet, etc.\textsuperscript{26} This would be equivalent to a protein or a nucleic acid altering its sequence as a response to an external perturbation, which is not possible because their residues are joined by covalent bonds. The high plasticity and flexibility of biological membranes allow them to maintain their bioactivity despite the effect of external perturbations. Thus, within our proposal, the primary sequence of membrane is simply the list of lipids in each leaflet.

2.2.- The \textbf{secondary structure} of the lipid bilayer: a Velcro strap more resistant than helices or beta sheets

The secondary structure of DNA arises from the interaction between two polynucleotide chains forming a double helix while in proteins it is defined by local and specific interactions between amino acids leading to alpha-helices or beta-sheets, among other patterns. The case of lipid bilayers is similar to DNA in the sense that they consist of two monolayers facing each other as a Velcro strap whereas DNA has two complementary chains interacting as a zipper (Fig 1B). This analogy is not only structural but also with respect to the interactions. The local interactions (by H-bonds) between two nucleotides in a DNA fragment is significantly stronger than the hydrophobic (mainly van der Waals type) interaction between two lipids of different leaflets in a membrane patch, however these latter interactions are highly cooperative-exactly as in a Velcro strap- and altogether makes the interaction between monolayers extremely strong.

This secondary structure is critical for the stability of lipid bilayers as lipid monolayers do not exist in solution (although they are stable in a hydrophobic/hydrophilic interphase). Again, there are differences between the two systems for this level of structure, \textit{e.g.} the interaction between leaflets taking place in two dimensions instead of
the one-dimensional interaction between polynucleotide chains, and the fluidity of lipids in a membrane compared to the fixed position of nucleotides in each DNA strand. As a result of this level of structure, the two leaflets in a lipid bilayer are highly correlated to each other. This is clearly illustrated in the topography of both monolayers as well as in the field lines arising from the curvature, isocurvature and anisocurvature (see Methods) vector fields (Fig. 2).

**Figure 2.** Topographic heatmap and LRS vector fields for both leaflets (see labels on top of each plot). Image obtained from the average over the last ns of the trajectory. The topography maps exhibit clear hills and wells which projection in the XY plane spans a diameter of ~15 nm with an amplitude of ~0.8 nm in Z. The negative correlation between the two leaflets is evident. The field lines obtained from the curvature, isocurvature and anisocurvature vector fields are also clearly correlated with the topography of the membrane. The correlation/connection between different regions and the anti-complementarity between the two leaflets are also obvious in these vector field maps. The curvature shows field lines falling from the top of the hills and rising from the bottom of the wells, as expected. The isocurvature exhibits strong curls and long fields lines contouring regions of same elevation. The anisocurvature displays long and straighter field lines, with no curls at all, and with vector line sources and wells in the regions where the isocurvature presents curls.
2.3.- Determination of local quantitative descriptors for the lipid bilayer. The **tertiary** structure of the lipid bilayer.

Higher levels of structure can be inferred from the coordinate behaviour of lipids, leading to local domains that interact with each other (Fig. 1C). The definition of the domains depends on the analyzed property. For instance, the topography of a membrane patch may exhibit protuberances and valleys (Fig. 2), the same structure could also have regions of different thickness, and the lipid tails could be orientated in a coordinate way forming patterns. Each of these properties would lead to different criteria to define domains. The analysis proposed in the methodology section identifies different kind of lipid domains as well as of interactions or couplings between them. The results corresponding to the large POPC monolayer simulated using coarse grained resolution are presented and discussed in what follows.

2.3.1.- Identification of domains

Since we are interested in the structure of the bilayer, only the results for the last part of the trajectory, which is assumed to be well equilibrated, will be shown here. As shown above (Fig. 2), the diameter and amplitude of the membrane protuberances conditions the vector fields arising from the LRS. Both the topography and these vector fields from well-defined domains. The distribution of the angles between the surface normal vectors and the Z axis of the simulation box exhibits a maximum at ~0.06 rad (3.4°) and, as expected, they are indistinguishable for both leaflets (Fig. S1). The heatmaps obtained from those angles show that larger angles correspond to the regions of maximum slope, i.e., halfway between hills and wells (Fig. S1).

The vector fields arising from the lipid tail projection in the XY plane were determined from the vectors connecting the PO4 group with the last bead of each tail (Fig. S2).
Both tails behave in a similar way, with their corresponding vector fields emanating from the highest points and converging in the lowest regions. This pattern was already expected since the curvature itself conditions the orientation of the lipid tails in the observed directions. The correlation between the topography and these vectors seems to be slightly more marked for SN1 than for SN2, probably due to the perturbation induced by the presence of the double bond in SN2. The lines arising from the vector field of the lipid tail projection in the XY plane are antiparallel to those obtained from the curvature ($Z'$) vector field. It is also useful to measure the angle between the lipid tail vectors and the $Z'$ axis of the LRS. These distributions are equivalent to those shown by the surface normal vectors (Fig. S1) but with the vectors pointing to the opposite direction (Fig. S3). The corresponding heatmaps (Fig. S4) are less clear than those obtained from the surface normal vectors, likely because the lipid bilayer is in fluid phase, but slighter lower angles (i.e., with lower projection over the $Z'$ axis of the LRS) can be observed for the regions where the topographic slope is maximum.

The membrane thickness calculated considering the LRS exhibits a narrow normal-like distribution centered around 3.9 nm (Fig. S5). The associated heatmap shows domains of different thickness. The correlation between the thickness and the topography is not evident, although the order of the lipid tails is expected to simultaneously depend on both properties and so they are expected to have some connection.

2.3.2.- Filtering non-redundant properties

In general, it is clear that all the analyzed properties arise from the interaction between groups of lipids and also that they define structural regions or domains in the lipid bilayer that span several cells. These domains, as well as the interaction between them, will be identified in the present work with the tertiary structure of the lipid bilayer, in
analogy with that in nucleic acids and proteins arising from the relative spatial location of nucleic acids and amino acids.

However, as stated above, the definition of the domains is not clear because those defined by the topography may be different from those defined by the thickness or by any other property. It is convenient to identify which properties are coupled to each other, thus providing identical or cooperating domains, and which properties are independent from the rest thus providing non redundant information. Next, a minimal set of quantitative descriptors obtained from the properties calculated in the previous section will be identified. This set of variables should be able to describe the structure of the lipid bilayer without redundancy. The SVD factorization of the matrix containing their z-score normalized values over all the grid cells and frames for the last 100 ns of our trajectory will be employed for this aim (see methods section for details). As both leaflets behave equivalently, the SVD analysis was separately carried out for each leaflet and lipid tail, giving equivalent results (Figs S6-S9).

The explained variance per eigenvector is a step function with several groups of properties contributing equally to the total variance (Figs S6-S9, panel A). This is a consequence of the close connection between variables -such as the topography of the membrane and the curvature, isocurvature and anisocurvature vectormaps- and their cooperative behavior defining the membrane domains, which in this case would not be defined by just one property but by the combination of several ones. Complementarily, just 11 variables are required to explain the 95% of the total variance (Figs S6-S9, panel A). The $V^T$ matrix containing the eigenvector compounds in the space generated by the properties used for the analysis, indicates too that several variables are coupled, since they can be explained by almost identical eigenvectors.
This coupling was quantified through the dot product between the column vectors consisting of the normalized relevance matrix (Figs S6-S9, panel C). The results of these dot products were then represented as a coupling matrix (Figs S6-S9, panel D). In these matrices, it is clearly seen that the X and Y components of the isocurvature and anisocurvature vectors are fully coupled to each other and the X and Y components of the surface normal are strongly coupled to the \( X' \) and \( Y' \) components of the lipid tails, respectively. These correlations were already expected after visually inspecting Fig. 2 and Fig S2, but finding them as a result can be considered as a validation of the method. The previous information suggests that the minimum number of properties used to efficiently characterize the lipid membrane can be reduced. Such set of properties could be considered as an orthogonal basis of quantitative descriptors for the lipid bilayer. It is interesting to observe the absolute lack of correlation between different properties, indicating that they are independent and complementary to each other. For instance, the topography of the membrane does not display any correlation with the X or Y component of either the iso or anisocurvature but does show it with the Z component of the latter vector field. Some coupling is also observed with some of the other variables, including the tail components in the LRS and the thickness. Interestingly, the topography does not couple at all with the Z component of the normal. It is clearly seen in Fig. 2 that the X and Y components of the normal vector are negligible in the maxima and minima of the membrane protuberances, thus a clear correlation between the Z component of said vector and the relative Z position of the lipid heads was to be expected. The reason for the observed lack of correlation is that it is only present at such maxima and minima but, not in general, such as at intermediate positions which weigh much more in the ensemble. For instance, the normal vectors at flat regions are also
parallel to the Z axis of the global reference system, although the relative Z position of
the corresponding head groups is null in these cases.

The redundant properties of the previous analyses were detected based on their shape
similarities in the Coupling, Relevance and $V^T$ matrices. After removing them, the SVD
analysis was repeated with the following descriptors: the X and Z component of the
vectors consisting the LRS, the SN1 tail tilt, the membrane thickness and the $Z'$
component of the SN1 tail vectors in the LRS. Note that by not explicitly considering
some variables we are not actually eliminating them, but assuming that they can be well
represented by some other property. Thus, the labels of Fig. 3 will show which variables
are equivalent to each other in the proposed orthogonal model.

No redundant eigenvectors were observed in the cumulative variance plot (Fig. 3A), as
expected when all the properties considered for the analysis are independent from each
other. The final $V^T$, normalized relevance and coupling matrices (Fig. 3B-D) shows that
there are strong couplings between different variables, but no redundant pair of
properties are present.
Figure 3.- Results for the SVD analysis performed on the nonredundant set of variables. Variance (red) and cumulative variance (blue), 95% explained marked with a red line (A); $V^T$ matrix (B); relevance matrix (C); coupling matrix (D). The Z component of the anisocurvature, and by extension the membrane topography, are coupled to tail-related parameters (tilt and $Z'$ component) and the membrane’s thickness, and completely uncoupled from the surface-defining vectors ($X'$ and $Y'$) of the LRS. On the other hand, the XY components of the $Z'$ LRS vector, and by extension the surface components of the lipid tails, are coupled entirely to the $X'$ and $Y'$ LRS vectors. Additionally, the Z component of the $Z'$ LRS vector, and by extension the membrane’s curvature, are almost exclusively coupled to the membrane’s thickness, which in turn couples to the $Z'$ component of the lipid tails and the topography/anisocurvature vector field. Hence, we can infer that the membrane’s topography is governed by how stretched the lipid tails are, influencing at the same time the membrane’s thickness. That will create a concrete 3D shape in the membrane, with higher and lower regions which in turn will rule the orientation of the lipid tails, thus creating a feedback of influence. From the initial set of 39 properties, considering both leaflets, around 9 quantitative descriptors resulted to be not superfluous. In analogy with the tertiary structure of proteins, the different domains interacting with each other –described by the set of non-redundant properties as shown in Fig. 3D– and their couplings can be associated to a tertiary structure in lipid bilayers. Thus, the coupling matrices (Fig. 3D) are expected to
be a fingerprint of the tertiary structure of a lipid bilayer. In the present work a simple homogenous POPC bilayer is analyzed. The size and shape of the domains defined by the properties employed here, or by a different set of quantitative descriptors, as well as the crossed interactions between such domains, could be totally different in a membrane of different composition. Going further, the projection of the MD trajectory of the membrane model on the different eigenvectors obtained from our analysis (Fig. S10) could provide an alternative group of complementary domains defined by the collective behavior of groups of lipids considering specific linear combinations of the quantitative descriptors employed in our analysis.

2.4.- The quaternary structure of the lipid bilayer

The levels of structure described above are expected to be present in all membranes, regardless of their composition, however the highest-order hierarchy remains to be established. In nucleotides, the quaternary structure is related to the interaction with different biomolecules which fold the double stranded DNA in different patterns. Following the same philosophy, it is sensible to test the response of membranes to the presence of macromolecules in their environment (Fig. 1D).
Figure 4 - Schematic representation of the quaternary structure in a DOPC lipid bilayer containing a carbon nanotube (A,B) and a $\beta$-barrel protein model (C,D). The vector fields obtained from the orientation of the lipid SN1 tail (A,C) and SN2 tail (B,D) in the upper leaflet are plotted together with the corresponding divergence represented in red-blue color gradient for divergence-convergence, respectively.

Taking advantage of previous work\textsuperscript{27} we test and compare the behaviour of an atomistic DOPC membrane model when in the presence of a carbon nanotube (CNT) and a simple $\beta$-barrel model (Fig. 4 and Fig. S11). Just by visual inspection, the influence of the macromolecules on the lipid organization is remarkable and distinguishable. In both cases the lipid tails become highly ordered around the macromolecule and they distort the nominal kind of organization seen in previous sections. The CNT creates a series of concentrical circular ripples around itself and generates a continuous pattern in the lipid tails. The divergence heatmap also displays the aforementioned rings, indicating that...
there will be circular regions defined by coordinate orientation of lipids. On the other hand, the beta-barrel generates local clumps of lipids with the same orientation, but there is no overall cohesion, contrary to the former case. Similarly, the divergence heatmap also displays protein-induced lumps with no clear global organizational pattern.

Thus, the case of the quaternary structure is defined straightforwardly from that of proteins of nucleic acids: it is the response and interaction of lipid membranes to other macromolecules, which is distinctive depending on the case.

**Conclusions**

A hierarchical set of structural levels is defined for lipid bilayers in analogy to the primary, secondary, tertiary and quaternary structural levels of nucleic acids and proteins. The primary structure is identified with the lipid composition of each leaflet, the secondary structure is defined by the interaction between the two leaflets conforming the membrane, the tertiary structure arises from the lipid domains obtained from the analysis of the properties calculated throughout the membrane surface, as well as from the interaction or coupling between such domains, and the quaternary structure comes from the impact of macromolecules embedded in the lipid bilayers. The coupling matrix proposed to identify the tertiary structure is expected to be a fingerprint associated to each membrane composition since it describes how the orientation of the lipid tails depends on the membrane topography, or how the membrane curvature affects the bilayer thickness, among other connections. The impact of the presence of macromolecules embedded in the membrane, corresponding to the quaternary structure, is huge. The pattern of the lipid tails orientation seems to be highly specific around rigid...
structures such as carbon nanotubes and less ordered around more flexible structures.

This is a first approach to the definition of hierarchical structural levels for membrane models. Part of our treatment is based on a set of local quantitative descriptors, but different sets of properties could be employed with the same objective. Our proposal is illustrated with a simple POPC bilayer as well as by the simulation of several macromolecules embedded in a DOPC membrane. The connection of the domains obtained from our analysis with biological activity is still unexplored. This new approach and analytical methods have a lot of room for improvement, but we hope it will impact in the understanding of membrane models as well as in the design of compartmentalized structures for new biotechnological applications.

Methods

1.- Simulation parameters

A flat-square membrane consisting of 5000 Martini 2.2 POPC lipids\textsuperscript{28} per leaflet was built using the CHARMM-GUI Martini Maker\textsuperscript{29–31}. The resulting bilayer, spanning about 57 nm in the X and Y dimensions, was solvated using 115429 non-polarizable Martini water particles, thus leading to a 57x57x8 nm\textsuperscript{3} box. No ions were added to the system since the employed lipid is zwitterionic and so the membrane model is neutral in the absence of additional charges. Periodic boundary conditions were applied to the three spatial dimensions and the energy of the whole system was minimized for 5000 steps using the GROMACS 2020.4 molecular dynamics engine\textsuperscript{32,33} with a steepest descent algorithm, a tolerance of 10 kJ·mol\textsuperscript{−1}·nm\textsuperscript{−1} and a step size of 0.01 nm. The simulation box was then equilibrated for 100 ns of molecular dynamics simulation in five consecutive stages, where the time step was sequentially increased from 2 fs to 20
fs using the leap-frog integrator. During this process the temperature was maintained at 310 K with a velocity rescale algorithm and a coupling constant of 1 ps, while the pressure was maintained at 1 atm with a Berendsen semi-isotropic barostat, a coupling constant of 5 ps and a compressibility modulus of $3 \cdot 10^{-4}$ bar$^{-1}$. The electrostatic interactions were calculated with the reaction-field method with a cutoff radius of 1.1 nm and a dielectric constant of 15 beyond the cutoff. The van der Waals interactions were calculated with a cutoff of 1.1 nm.

After equilibration, an unrestrained 10-μs-long MD production trajectory using a leap-frog integrator with a timestep of 25 fs was obtained. Removal of center of mass motion was performed each 1000 steps, and position coordinates were written with the same period (25 ps). Neighbor lists were obtained with the Verlet method and updated every 20 steps. Temperature was maintained at 310 K with a velocity-rescale algorithm and a coupling constant of 1 ps; while pressure was kept constant at 1 atm with a semi-isotropic Parrinello-Rahman barostat, a coupling parameter of 12 ps and a compressibility modulus of $3 \cdot 10^{-4}$ bar$^{-1}$. Electrostatic interactions were calculated with the reaction-field scheme, a cutoff radius of 1.1 nm and a dielectric constant of 15 beyond the given radius. van der Waals interactions were calculated with a cutoff of 1.1 nm.

2.- Analysis

Most of the analyses were carried with homemade scripts written in Python. The recurrently used libraries were MDAnalysis to directly read trajectory files in GROMACS format, Numba to facilitate high-performance parallelization schemes, NumPy for general data handling and mathematical operations, and Matplotlib for data representation. Visual depictions of the membrane were made using VMD.
2.1.- Determination of local properties

The analysis performed in this work relies on the discretization of the lipid membrane in a square grid spanning the whole XY plane, with an average of $N_L$ lipid heads (PO4 groups) per cell. The coordinates of the grid cells are common for the two leaflets and were kept constant over the entire segment of trajectory analyzed. A battery of different properties was then determined for each cell and independently analyzed for each leaflet.

In order to describe the dynamic behavior of the membrane, as well as local domains and the connections between them, a local reference system (LRS) per grid cell as a function of time was determined. The orthonormal vectors describing these LRS are: (i) the normal to each cell, $Z'_i(t)$; (ii) the normalized cross product between $Z'_i(t)$ and the gradient of its projection on the Z axis of the global reference system, $X'_i(t)$; and (iii) the cross product between the previous two vectors, $Y'_i(t)$. The protocol used to define these vectors is described in detail in the Supplementary Material (SM). The vector fields arising from $Z'_i(t)$, $X'_i(t)$ and $Y'_i(t)$ will be identified with the curvature, the isocurvature and the anisocurvature, respectively, of the lipid bilayer. By using these LRS it is possible to determine structural properties such as membrane thickness or tail orientation while accounting for local deformations of the lipid bilayer. Such deformations might be ignored in small membranes, but they are significant in patches of several tens of nm long per edge, as it will be shown later.

The set of properties analyzed in the present paper can be computed with different time and size resolution. The size resolution is given by the number of lipids per grid cell ($N_L$) while the time resolution ($\Delta t$) depends on the frequency of the fluctuations observed for each property. Different values of these two parameters were tested in our
code and after several tests we decided to present the results obtained with $N_L = 10$ and a moving average window of $\Delta t = 1$ ns. These values allow observing significant changes in the quantitative properties computed for the lipid bilayer, with negligible noise and optimizing the computational resources employed. Using these parameters, the topography of the membrane was determined as the average $Z$ coordinate of the PO4 groups per grid cell. The local membrane curvature was quantified as the projection of $Z'_i(t)$ on the $Z$ axis. The thickness of the membrane and the orientation of the lipid tails were determined using the LRS. The thickness is given by the average difference between the position of PO4 groups along the local $Z'_i(t)$ axes, while the orientation of the tails is the difference between the coordinates of the PO4 and last beads of each tail expressed in the LRS.

2.2. - Global analysis of the lipid bilayer

Once the previously described properties are computed for each grid cell, time dependent local domains could be directly observed by plotting them in suitable representations (see results section and SM). A Singular Value Decomposition (SVD) analysis is then performed to identify a set of independent properties associated to the membrane structure as well as to quantify the coupling between the different studied properties (see SM for a detailed description). Each of the three components of the previously defined vectorial magnitudes were independently treated as scalar properties for this analysis.
References

1. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).

2. Franklin, R. E. & Gosling, R. G. Evidence for 2-chain Helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172**, 156–157 (1953).

3. Kendrew, J. C. The three-dimensional structure of a protein molecule. *Sci. Am.* **205**, 96–110 (1961).

4. Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666 (1958).

5. Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science (80-. ).* **175**, 720–731 (1972).

6. Kusumi, A., Suzuki, K. G. N., Kasai, R. S., Ritchie, K. & Fujiwara, T. K. Hierarchical mesoscale domain organization of the plasma membrane. *Trends Biochem. Sci.* **36**, 604–615 (2011).

7. Alberts, B. *et al.* *Molecular Biology of the Cell.* (Garland Science, 2014).

8. Nussinov, R., Tsai, C. J. & Jang, H. Ras assemblies and signaling at the membrane. *Curr. Opin. Struct. Biol.* **62**, 140–148 (2020).

9. Corradi, V. *et al.* Emerging Diversity in Lipid-Protein Interactions. *Chem. Rev.* **119**, 5775–5848 (2019).

10. Stieger, B., Steiger, J. & Locher, K. P. Membrane lipids and transporter function. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1867**, 166079 (2021).

11. Sens, P. & Plastino, J. Membrane tension and cytoskeleton organization in cell
motility. *J. Phys. Condens. Matter* **27**, 273103 (2015).

12. Kielian, M. & Jungerwirth, S. Mechanisms of Enveloped virus entry into cells. *Mol. Biol. Med.* **7**, 17–31 (1990).

13. Dimitrov, D. S. Virus entry: Molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* **2**, 109–122 (2004).

14. Münz, C. The autophagic machinery in viral exocytosis. *Front. Microbiol.* **8**, 269 (2017).

15. Garcia-Fandino, R. & Piñeiro, Á. Delving Into the Origin of Destructive Inflammation in COVID-19: A Betrayal of Natural Host Defense Peptides? *Front. Immunol.* **11**, 3532 (2021).

16. Szlasa, W., Zendran, I., Zalesińska, A., Tarek, M. & Kulbacka, J. Lipid composition of the cancer cell membrane. *J. Bioenerg. Biomembr.* **52**, 321–342 (2020).

17. Alves, A. C., Ribeiro, D., Nunes, C. & Reis, S. Biophysics in cancer: The relevance of drug-membrane interaction studies. *Biochim. Biophys. Acta - Biomembr.* **1858**, 2231–2244 (2016).

18. Glukhov, E., Stark, M., Burrows, L. L. & Deber, C. M. Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes. *J. Biol. Chem.* **280**, 33960–33967 (2005).

19. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**, 133–159 (2015).

20. Cas, M. D., Roda, G., Li, F. & Secundo, F. Functional lipids in autoimmune inflammatory diseases. *Int. J. Mol. Sci.* **21**, 3074 (2020).
21. Leitinger, N. The role of phospholipid oxidation products in inflammatory and autoimmune diseases: Evidence from animal models and in humans. *Subcell. Biochem.* **49**, 325–350 (2008).

22. Moradi, S., Nowroozi, A. & Shahlaei, M. Shedding light on the structural properties of lipid bilayers using molecular dynamics simulation: a review study. *RSC Advances* vol. 9 4644–4658 (2019).

23. Sud, M. *et al.* LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **35**, D527–D532 (2007).

24. Magana, M. *et al.* The value of antimicrobial peptides in the age of resistance. *Lancet Infect. Dis.* **20**, e216–e230 (2020).

25. Huan, Y., Kong, Q., Mou, H. & Yi, H. Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields. *Frontiers in Microbiology* vol. 11 582779 (2020).

26. Ernst, R., Ballweg, S. & Levental, I. Cellular mechanisms of physicochemical membrane homeostasis. *Current Opinion in Cell Biology* vol. 53 44–51 (2018).

27. Garcia-Fandiño, R., Piñeiro, Á., Trick, J. L. & Sansom, M. S. P. Lipid Bilayer Membrane Perturbation by Embedded Nanopores: A Simulation Study. *ACS Nano* **10**, 3693–3701 (2016).

28. De Jong, D. H. *et al.* Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–697 (2013).

29. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

30. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS,
AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

31. Hsu, P. C. *et al.* Charmm-gui martini maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *J. Comput. Chem.* **38**, 2354–2363 (2017).

32. Lindahl, Abraham, Hess & Spoel, van der. GROMACS 2020.4 Manual. (2020) doi:10.5281/ZENODO.4054996.

33. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

34. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

35. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

36. Verlet, L. Computer ‘experiments’ on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).

37. Chialvo, A. A. & Debenedetti, P. G. An automated Verlet neighbor list algorithm with a multiple time-step approach for the simulation of large systems. *Comput. Phys. Commun.* **70**, 467–477 (1992).

38. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
39. Gowers, R. et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. in *Proceedings of the 15th Python in Science Conference* 98–105 (2016). doi:10.25080/majora-629e541a-00e.

40. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).

41. Lam, S. K., Pitrou, A. & Seibert, S. Numba: A LLVM-based Python JIT Compiler. in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC - LLVM ’15* (ACM Press).

42. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).

43. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

44. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

**Acknowledgements**

This work was supported by the Spanish Agencia Estatal de Investigación (AEI) and the ERDF (RTI2018-098795-A-I00, and PID2019-111327GB-I00), by Xunta de Galicia and the ERDF (ED431F 2020/05 and Centro singular de investigación de Galicia accreditation 2016-2019, ED431G/09), and by Fundação para a Ciência e Tecnologia (FCT) funds and FEDER European funds (POCI-01-0145-30579). R.G.-F. thanks Ministerio de Ciencia, Innovación y Universidades for a “Ramón y Cajal” contract (RYC-
2016-20335). A. B.-G thanks Xunta de Galicia for a “Doutoramento Industrial” contract (06_IN606D_2021_ 2600276). All calculations were carried out at the Centro de Supercomputación de Galicia (CESGA).

Contributions

Á.P., and R. G-F. conceived the study. A.B. performed the simulations and analysis. All authors discussed the results, contributed to write the manuscript.

Corresponding authors

Correspondence to Rebeca Garcia-Fandino or Ángel Piñeiro.

Competing interests

The authors declare no competing interests.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SM20220122.pdf