

# Inverse conformational selection in lipid-protein binding

Amélie Bacle,<sup>†</sup> Pavel Buslaev,<sup>‡,¶</sup> Rebeca Garcia-Fandino,<sup>§,||</sup> Fernando Favela-Rosales,<sup>⊥</sup> Tiago Mendes Ferreira,<sup>#</sup> Patrick F.J. Fuchs,<sup>@,△</sup> Ivan Gushchin,<sup>¶</sup> Matti Javanainen,<sup>▽</sup> Anne M. Kiirikki,<sup>††</sup> Jesper J. Madsen,<sup>‡‡,¶¶</sup> Josef Melcr,<sup>§§</sup> Paula Milán Rodríguez,<sup>@</sup> Markus S. Miettinen,<sup>|||</sup> O. H. Samuli Ollila,<sup>\*,††</sup> Chris G. Papadopoulos,<sup>⊥⊥</sup> Antonio Peón,<sup>||</sup> Thomas J. Piggot,<sup>##</sup> Ángel Piñeiro,<sup>@@</sup> and Salla I. Virtanen<sup>††</sup>

<sup>†</sup>Laboratoire Coopératif "Lipotoxicity and Channelopathies - ConicMeds", Université de Poitiers, 1 rue Georges Bonnet, Poitiers, 86000, France

<sup>‡</sup>Nanoscience Center and Department of Chemistry, University of Jyväskylä, P.O. Box 35, Jyväskylä, 40014, Finland

<sup>¶</sup>Research Center for Molecular Mechanisms of Aging and Age-related Diseases, Moscow Institute of Physics and Technology, Dolgoprudny, 141701, Russia

<sup>§</sup>Center for Research in Biological Chemistry and Molecular Materials (CiQUS), Universidade de Santiago de Compostela, Santiago de Compostela, E-15782, Spain

<sup>||</sup>CIQUP, Centro de Investigaçãõ em Química, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Porto, 4169-007, Portugal

<sup>⊥</sup>Departamento de Ciencias Básicas, Tecnológico Nacional de México - ITS Zacatecas Occidente, Sombrerete, Zacatecas, 99102, México

<sup>#</sup>NMR group - Institute for Physics, Martin Luther University Halle-Wittenberg, Halle (Saale), 06120, Germany

<sup>@</sup>Sorbonne Université, Ecole Normale Supérieure, PSL University, CNRS, Laboratoire des Biomolécules (LBM), Paris, 75005, France

<sup>△</sup>Université de Paris, UFR Sciences du Vivant, Paris, 75013, France

<sup>▽</sup>Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo nám. 542/2, Prague, CZ-16610, Czech Republic

<sup>††</sup>Institute of Biotechnology, University of Helsinki, Helsinki, 00014, Finland

<sup>‡‡</sup>Department of Chemistry, The University of Chicago, Chicago, Illinois, 60637, United States of America

<sup>¶¶</sup>Global and Planetary Health, College of Public Health, University of South Florida, Tampa, Florida, 33612, United States of America

<sup>§§</sup>Groningen Biomolecular Sciences and Biotechnology Institute and The Zernike Institute for Advanced Materials, University of Groningen, Groningen, 9747 AG, The Netherlands

<sup>|||</sup>Department of Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, 14424, Germany

<sup>⊥⊥</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, 91198, France

<sup>##</sup>Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom

<sup>@@</sup>Departamento de Física Aplicada, Faculdade de Física, Universidade de Santiago de Compostela, Santiago de Compostela, E-15782, Spain

E-mail: samuli.ollila@helsinki.fi

## Abstract

Interest in lipid interactions with proteins and other biomolecules is emerging not only in fundamental biochemistry, but also in the field of nanobiotechnology where lipids are commonly used, for example, in carriers of mRNA vaccines. The outward facing components of cellular membranes and lipid nanoparticles, the lipid headgroups, regulate membrane interactions with approaching substances, such as proteins, drugs, RNA or viruses. Because lipid headgroup conformational ensembles have not been experimentally determined in physiologically relevant conditions, an essential question on their interactions with other biomolecules remains unanswered: Do headgroups exchange between a few rigid structures, or fluctuate freely across a practically continuous spectrum of conformations? Here, we combine solid-state NMR experiments and molecular dynamics simulations from the NMRlipids Project to resolve the conformational ensembles of headgroups of four key lipid types in various biologically relevant conditions. We find that lipid headgroups sample a wide range of overlapping conformations in both neutral and charged cellular membranes, and that differences in the headgroup chemistry manifests only in probability distributions of conformations. Furthermore, the analysis of 894 protein-bound lipid structures from the Protein Data Bank (PDB) suggests that lipids can bind to proteins in wide range of conformations, which are not limited by the headgroup chemistry. We propose that lipids can select a suitable headgroup conformation from the wide range available to them to fit the various binding sites in proteins. The proposed *Inverse Conformational Selection Model* will extend also to lipid binding to other targets than proteins, such as drugs, RNA and viruses.

## Introduction

Lipid interactions with other biomolecules are gaining interest not only in molecular cell biology,<sup>1</sup> but also in nanobiotechnology applications, such as design of antibodies<sup>2</sup> and carriers for mRNA vaccines.<sup>3,4</sup> Interactions of cellular membranes or lipid nanoparticles with

other biomolecules, such as proteins, drugs or RNA, are regulated by lipid headgroups, the outward facing components of lipid bilayers.<sup>1,5</sup> Chemical compositions of lipid headgroups vary between different organelles and organisms, and specific interactions with certain lipid headgroups are known to be essential for the function of several proteins.<sup>1,5</sup> However, the driving forces for specific interactions between lipids and membrane binding substances are not fully understood because conformational ensembles of lipids in physiologically relevant liquid state have not been experimentally determined. Therefore, it is not clear if the specificity arises from the differences in accessible conformations between lipid types or from specific intermolecular lipid-protein interactions.

Structures of protein-bound lipids are available in the Protein Data Bank (PDB, <http://www.rcsb.org/>),<sup>6</sup> and crystal structures of lipids have been determined,<sup>7,8</sup> but their relation to lipids in bulk membranes in the liquid lamellar phase remains unclear.<sup>9</sup> Similarly to disordered proteins,<sup>10</sup> lipid molecules sample conformational ensembles that are composed of sets of individual structures whose occurrence probabilities are determined by the Boltzmann distribution.<sup>11</sup> In liquid bilayer state, individual phosphatidylcholine lipid samples the conformational ensemble on nanosecond timescale.<sup>12,13</sup> The most accurate experimental information on conformational ensembles of lipids in this biologically relevant phase are typically derived from NMR experiments, particularly from the C-H bond order parameters,  $S_{CH} = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle$ , where  $\theta$  is the angle between C-H bond and membrane normal, and the average is taken over the conformational ensemble of lipids.<sup>14-16</sup> Because these order parameters are similar to those in living cells,<sup>17-19</sup> the model membranes can be used to resolve the lipid headgroup conformations in biological conditions. According to these experiments, the glycerol backbone conformations are largely similar irrespective of the headgroup,<sup>17</sup> and the headgroup conformations are similar in the phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylglycerol (PG) lipids, whereas the phosphatidylserine (PS) headgroup is more rigid.<sup>7,20</sup> Notably, however, the signs of  $S_{CH}$  are not accessible from <sup>2</sup>H NMR experiments,<sup>21</sup> and universal models to map order parameters

to structural ensembles are not available.<sup>22,23</sup> Therefore, it is not clear if the lipid headgroups in a liquid lamellar bilayer exchange between a few restricted conformations, or can fluctuate freely across a wide conformational range.

Here, we use natural abundance  $^1\text{H}$ - $^{13}\text{C}$  solid-state NMR experiments and MD simulations from the NMRlipids Project to resolve differences in the conformational ensembles of PC, PE, PG, and PS lipid headgroups. Zwitterionic PC are the most common lipids in eukaryotes and PE in bacteria;<sup>5,24</sup> PE is also the second most abundant glycerophospholipid in eukaryotic cells and has been related to various diseases.<sup>25–27</sup> PC lipids are also used in mRNA COVID-19 vaccines.<sup>4</sup> PS are the most common negatively charged lipids in eukaryotes and PG in bacteria, and both are known to affect membrane protein functionality and signaling.<sup>24,28–30</sup> All the four studied lipid types specifically bind to various proteins;<sup>31</sup> we therefore elucidate also the effect of protein binding to headgroup conformations. The resulting *inverse conformational selection model* for lipid-protein binding can be also applied to understand interactions between membranes and other biomolecules than proteins, such as drugs, RNA, or viruses. On the other hand, the detailed understanding of lipid bilayer interfaces can facilitate the design of lipid nanoparticle carriers for mRNA vaccines with less side effects.

## Results and Discussion

### Differences between lipid headgroups from $^{13}\text{C}$ NMR experiments

To experimentally characterize the differences in headgroup conformational ensembles of lipids that are not bound to proteins, we measured the C–H bond order parameters  $S_{\text{CH}}$  and their signs of POPG and POPE in the liquid lamellar phase, as we did previously for POPC and POPS.<sup>32–34</sup> Determination of headgroup and glycerol backbone  $S_{\text{CH}}$  and their signs was straightforward from the data in Figs. 1, S1 and S2 for all the C–H bonds, except for the  $\beta$  and  $g_2$  carbons in POPG. These carbons have overlapping peaks in the INEPT spectra due to their similar chemical environments, and only the magnitude of the larger

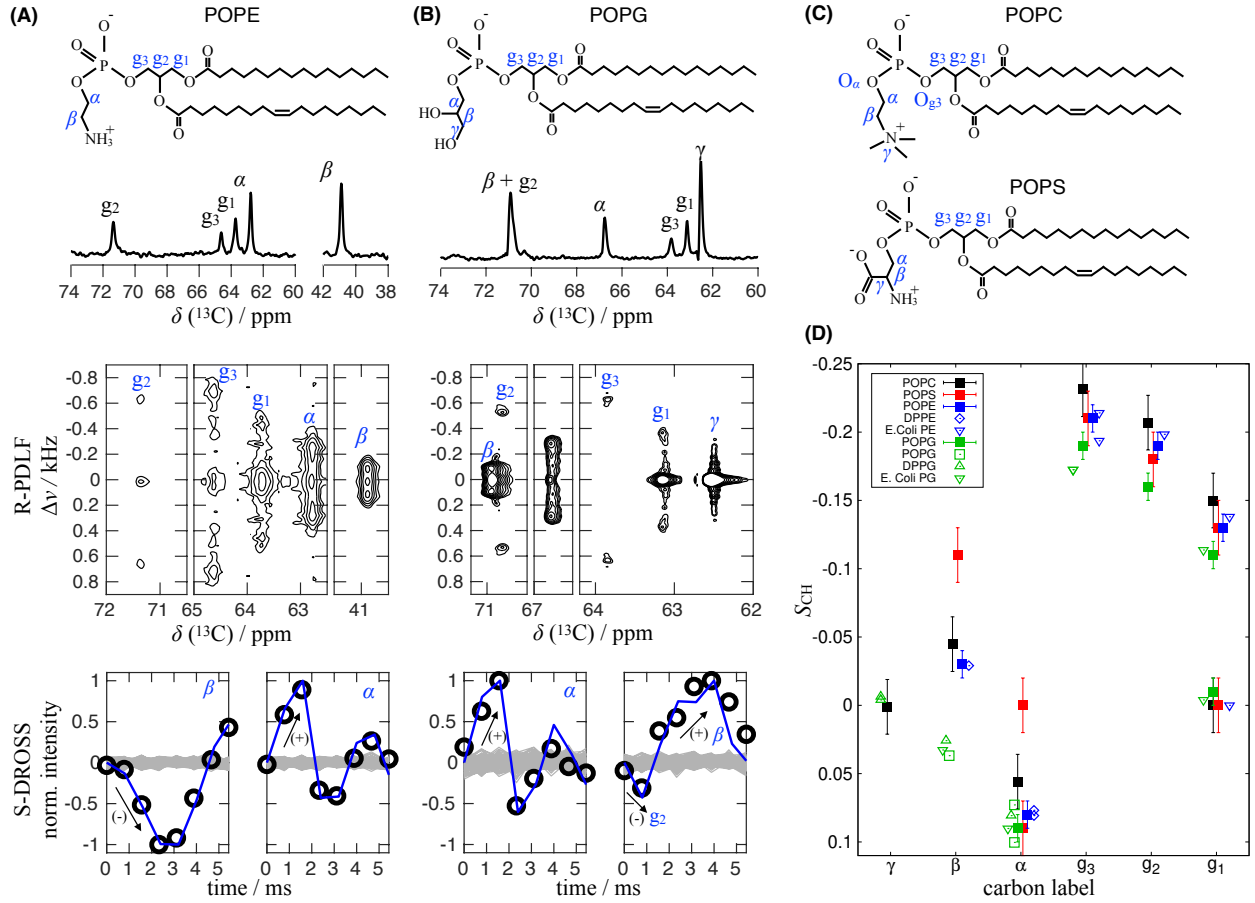


Figure 1: Chemical structure, refocused-INEPT spectrum, 2D R-PDLF spectra, and S-DROSS dipolar slices (from top to bottom) for **A)** POPE and **B)** POPG multi-lamellar vesicles measured by  $^1\text{H}$ - $^{13}\text{C}$  solid-state NMR. The S-DROSS dipolar slices show both experimental data (open circles) and the result of NMR numerical simulations (blue solid lines). Full NMR spectra are shown in Figs. S1 and S2. **C)** Chemical structure of POPC and POPS. **D)** Headgroup and glycerol backbone C-H bond order parameters  $S_{\text{CH}}$  for different phospholipids measured in the lamellar liquid disordered phase. Filled squares show  $S_{\text{CH}}$  magnitudes and signs determined by  $^1\text{H}$ - $^{13}\text{C}$  NMR spectroscopy for POPE (310 K) and POPG (298 K), as measured in this work, and from data previously reported for POPS (298 K)<sup>32</sup> and POPC (300 K).<sup>33,34</sup> Empty symbols show headgroup and glycerol backbone  $S_{\text{CH}}$  magnitudes measured previously with  $^2\text{H}$  NMR and are also plotted as real values using the signs determined in this work for POPG with 10 mM PIPES (298 K),<sup>35</sup> DPPG with 10 mM PIPES and 100 mM NaCl (314 K),<sup>20</sup> DPPE (341 K),<sup>36</sup> *E. coli* PE and *E. coli* PG with 10 mM PIPES and 100 mM NaCl (310 K).<sup>17</sup>

order parameter could be determined from the R-PDLF spectra (Fig. 1B). Based on previous  $^2\text{H}$  NMR measurements,<sup>17,20,35</sup> we assigned the larger  $S_{\text{CH}}$  to the  $g_2$  carbon and used the literature value for the  $\beta$ -carbon in SIMPSON simulations to determine the signs. The decrease in the beginning of the S-DROSS curve suggests that the sign of the larger  $g_2$   $S_{\text{CH}}$  is negative, and the later increase suggests that sign of the smaller  $\beta$   $S_{\text{CH}}$  is positive, as confirmed by the SIMPSON simulations (Fig. 1B).

Experimental order parameters of POPC, POPE, POPG, and POPS glycerol backbones and headgroups from this and previous studies are displayed in Fig. 1D, where the signs from  $^{13}\text{C}$  NMR experiments are used also for the  $^2\text{H}$  NMR data from the literature. The overall agreement of  $S_{\text{CH}}$  determined by different research teams and different techniques for the same lipid headgroup was excellent (within  $\pm 0.02$ ) here and in previous studies.<sup>21,32,37</sup> Therefore, the differences between lipid types are dictated by the headgroup chemistry rather than inaccuracies in experiments, or differences in the acyl chains or in the experimental conditions.

The most distinct order parameters are observed for PS headgroups, for which the  $\alpha$ -carbon  $S_{\text{CH}}$  exhibits significant forking (different values for different hydrogens bound to the same carbon) and the  $\beta$ -carbon has a more negative value than in other studied lipid types. On the other hand, the  $\beta$ -carbon  $S_{\text{CH}}$  of PG headgroup has a positive sign, in contrast to all the other lipid types. The sign has not been previously known from  $^2\text{H}$  NMR experiments which enable detecting only order parameter absolute values.<sup>17,20,35</sup> The glycerol backbone order parameters are similar for all the lipid types, although they move slightly toward positive values (closer to zero) in the order  $\text{PC} < \text{PE} < \text{PS} < \text{PG}$ . Only minor differences between PC and PE headgroups are observed.

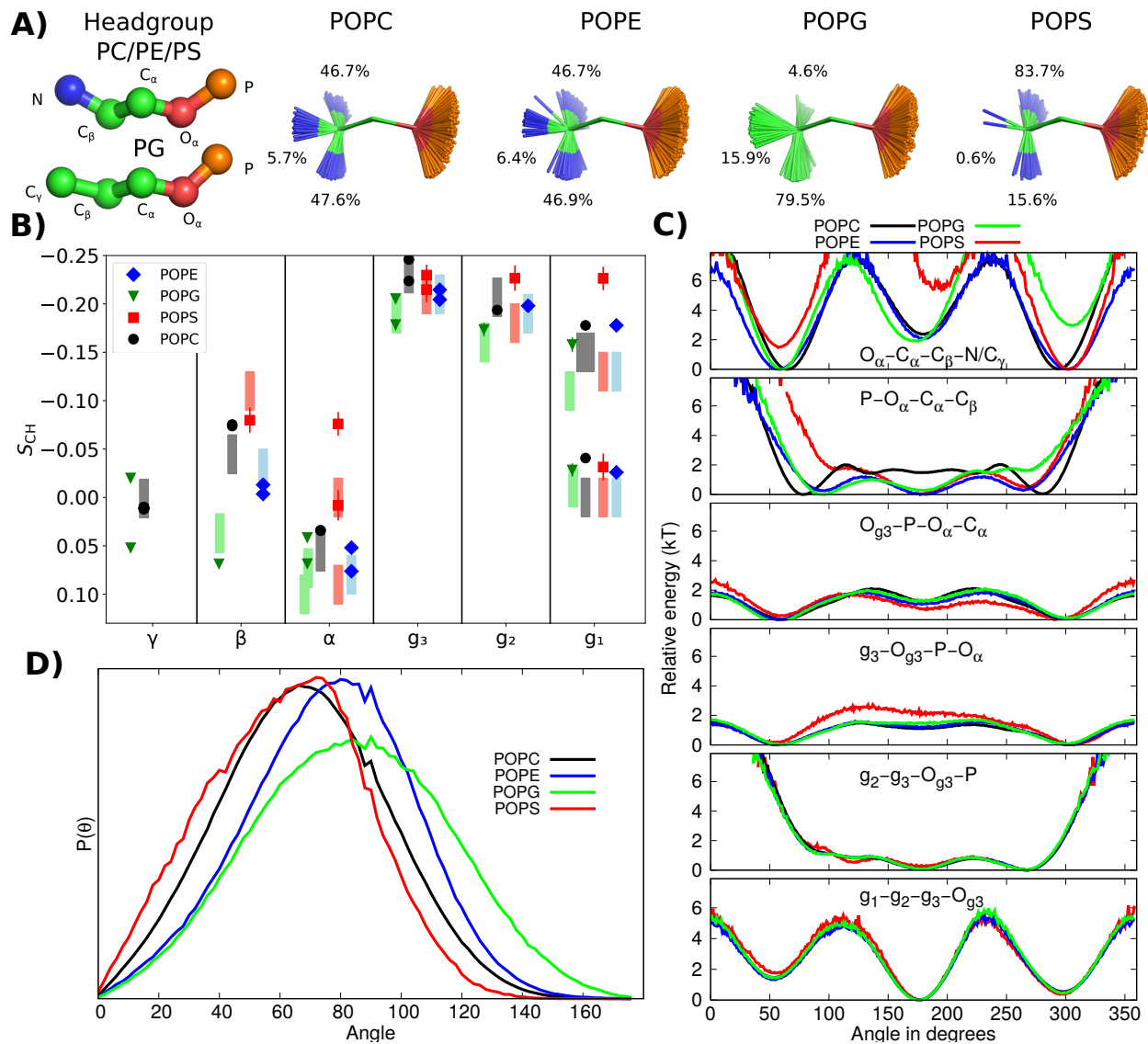


Figure 2: Results from the best simulation model demonstrating the differences in conformational ensembles between different lipid types. The best simulation model (CHARMM36) was selected from the comparison of ten different force fields based on quality evaluation against experimental  $S_{CH}$  values in Figs. S3 and S4. **A)** Snapshots with overlaid  $C_{\beta}$ ,  $C_{\alpha}$  and  $O_{\alpha}$  atoms and occurrence of different conformations. **B)** Headgroup and glycerol backbone region order parameters of the different lipid types. Colored shaded regions show the range of experimental values from Fig. 1. **C)** Relative free energies for individual heavy atom dihedral angles estimated from the inverse Boltzmann formula. Angles corresponding to free energies above  $8 k_B T$  are not shown because they are not observed in simulations. **D)** Distributions of the P-N/ $C_{\gamma}$  vector angle with respect to the membrane normal.

# Conformational ensembles of different lipid headgroups from MD simulations

Ideal MD simulations would correctly predict the conformational ensembles of lipids, thereby giving an interpretation for the experimental C-H bond order parameters,  $S_{CH}$ , due to their direct connection.<sup>21</sup> To find the best MD simulation models (force fields) for this purpose, we took advantage of the power of the NMRlipids Project, an open science collaboration that has gathered a massive number of simulation trajectories, allowing the evaluation of all the available force fields of PC,<sup>37</sup> PS,<sup>32</sup> PE (Fig. S3), and PG (Fig. S4) lipids against NMR data. The comparison of hundreds of simulation trajectories simulated with approximately 20 different force fields revealed that none of them can correctly capture the lipid conformational ensembles and reproduce the experimental  $S_{CH}$  of all segments within experimental error (for detailed discussion on the quality of force fields, see section S3 in the supplementary information and Refs. 13,21,32,37). However, CHARMM36, the force field showing the best agreement with experiments, correctly captures the essential differences between the PC, PE, PG, and PS headgroup order parameters, particularly the positive order parameter of  $\beta$ -carbon in PG and the forking of  $\alpha$ -carbon in PS, see Fig. 2 B.

To understand the structural origin of distinct order parameters for the distinct lipids, we first calculated and visualized the heavy atom dihedral angle distributions from the best performing simulations that reproduced the differences between headgroups (Figs. 2 A and S11). Then, we used the inverse Boltzmann formula to estimate the free energy costs for different dihedral angle orientations. The results in Fig. 2C suggest that all lipid headgroups are very flexible and free energy costs for rotating individual dihedrals to almost any angle is low (below  $8 k_B T$ ). Only *cis* states of P-O $_{\alpha}$ -C $_{\alpha}$ -C $_{\beta}$  and g<sub>2</sub>-g<sub>3</sub>-O<sub>g<sub>3</sub></sub>-P have larger relative free energies and are not observed for any lipids during the simulations.

Major differences between headgroups are observed for the last two dihedrals in the headgroup end, O $_{\alpha}$ -C $_{\alpha}$ -C $_{\beta}$ -N/C $_{\gamma}$  and P-O $_{\alpha}$ -C $_{\alpha}$ -C $_{\beta}$ , which prefer *gauche*<sup>-</sup> conformations for PG and *gauche*<sup>+</sup> for PS, whereas PC and PE exhibit symmetric distributions. Also, the free



energy barriers for  $O_\alpha-C_\alpha-C_\beta-N/C_\gamma$  dihedral rotations between *gauche* and *trans* states are larger for PS and PG lipids than for PC and PE lipids. The rest of the dihedrals are similar between different lipids, with the exception of PS lipids for which slightly larger free energy for *eclipsed anti* conformation in  $g_3-O_{g3}-P-O_\alpha$  dihedral was observed. Therefore we suggest that the main differences between lipid headgroups leading to distinct order parameters occur in the headgroup end, while in the phosphate region they remain similar in different lipids with the exception of PS. The increased barriers for dihedral rotations may explain the more rigid headgroup structures in PS.<sup>7,38</sup> Furthermore, the angle between headgroup dipole and membrane normal decreases in the order of  $PG > PE > PC > PS$  (Fig. 2D). However, the differences between PC and PE in  $P-O_\alpha-C_\alpha-C_\beta$  dihedral and  $P-N$  vector dipole may be artificial as the  $\beta$ -carbon order parameter in PC is too negative even in the best available force field,<sup>37</sup> thereby not being equal to the order parameter in PE as observed in experiments.

In conclusion, the rotation of dihedral angles to almost any position bears relatively low free energy cost and the sampled dihedral angles are within similar ranges for all studied headgroup types. In other words, all studied lipid headgroups sample very broad conformational ensembles in the liquid lamellar phase. Therefore, the lipid headgroups are able to adopt a wide range of conformations when interacting with proteins, ions, or other biomolecules. Furthermore, the structures in lipid crystals<sup>7,8</sup> are not representative of the liquid state, and models aiming to explain NMR data using only a few conformations<sup>14–16,23</sup> are not sufficient to capture the large conformational space of lipids in the liquid lamellar state. However, conformational ensembles of lipids in MD simulation trajectories with accurate force fields that capture the experimental NMR order parameters can be considered realistic, and thereby used to interpret the experimental data.

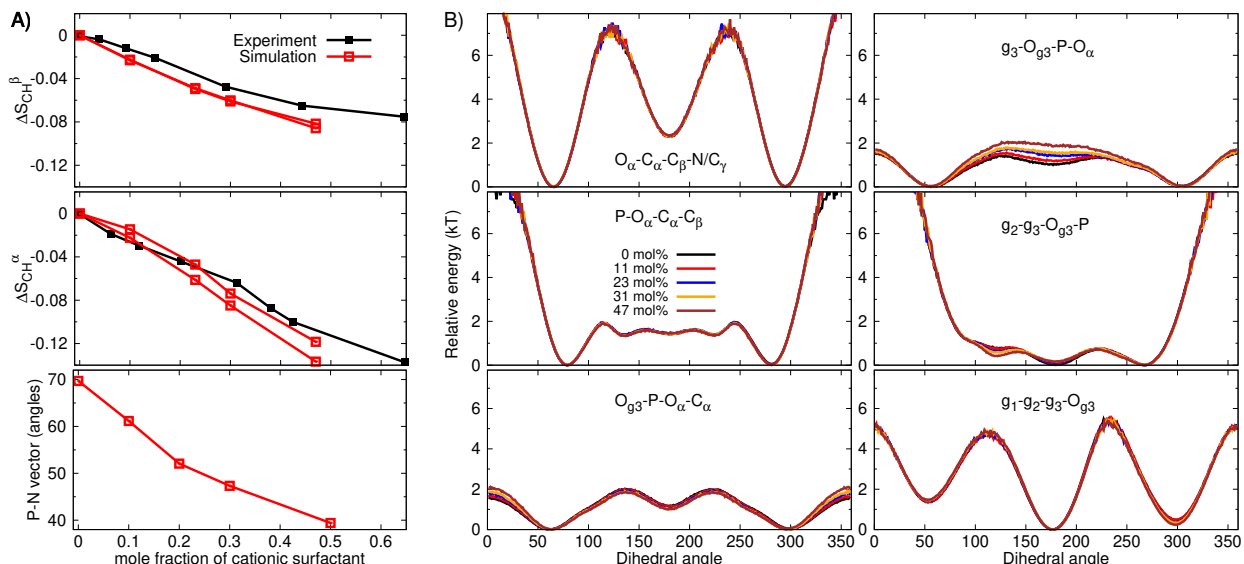


Figure 3: **A)** Modulation of PC headgroup order parameters and P–N vector tilt (average angle between the P–N vector and the bilayer normal) upon addition of cationic surfactant (dihexadecyldimethylammonium) from CHARMM36 simulations compared with the experimental data.<sup>39</sup> **B)** Relative free energies for individual dihedral angles estimated from inverse Boltzmann distributions of heavy atom dihedral angles with different amounts of cationic surfactant from CHARMM36 simulations.

## Lipid conformational ensembles in charged membranes

Charged entities, such as lipids, proteins, surfactants, drugs, and ions incorporated in membranes can change the orientation of the headgroup dipole in PC lipids and affect the order parameters of the lipid headgroups.<sup>40</sup> However, the detailed understanding of the structural and energetic responses of lipids to membrane-bound electric charge is still lacking.<sup>16</sup> For example, it is not clear if the membrane-bound charges restrain the lipids into specific conformations or just redistribute the probabilities of states by keeping the accessible conformations similar.

To resolve lipid headgroup conformational ensembles in cellular membranes bearing positive charge, we calculated the heavy atom dihedral angle distributions from the subset of simulations that correctly captured the experimentally measured decrease in PC headgroup order parameters upon addition of cationic surfactants into a bilayer (figure 3 A). The dihedral angle distributions and relative free energies in figures S12 and 3 B reveal that the addition of

positive charge into a membrane decreases the abundance of *trans* states in  $g_2-g_3-O_{g_3}-P$  and  $g_3-O_{g_3}-P-O_\alpha$ , and *cis* states in  $O_{g_3}-P-O_\alpha-C_\alpha$  dihedrals. The choline region remains essentially unchanged and only minor changes are observed in the other dihedrals even at a surfactant molar fraction of 0.47.

In addition, binding of ions to the membrane may affect the lipid headgroup conformational ensembles under physiological conditions. The bound  $Ca^{2+}$  ion to PC headgroup leads to similar decrease in *trans* state probability for  $g_3-O_{g_3}-P-O_\alpha$  dihedral in the most realistic MD simulation models as observed for cationic surfactants (lipid17ecc and CHARMM36 in Figs. S13, S14 and Fig. S7). The dihedral distributions of the PG headgroup are more sensitive to the bound ions in the most realistic simulations, but upward tilting of the headgroup dipole upon the addition of  $CaCl_2$  is weaker than in PC (Lipid17 and Slipids in Figs. S7, S15 and S16). However, the changes in PG lipid dihedrals upon the addition of  $CaCl_2$  differ between the best models (Figs. S15 and S16), and none of the simulations captures the  $Ca^{2+}$  ion binding affinity and conformational ensemble of PG lipids simultaneously. Moreover, experimental data to evaluate the response of  $\alpha$ -carbon order parameters to the added  $CaCl_2$  in PG is not available. Additionally, the headgroup conformational ensembles in mixtures of PC and charged (PG or PS) or zwitterionic (PE) lipids could not be resolved with the currently available force fields and experimental data (Figs. S5 and S6, and Ref. <sup>32,41</sup>).

In conclusion, the MD simulation results suggest that the structural response of lipid headgroups to membrane-bound charges arises from mild changes in dihedral angle distributions, rather than restriction of lipids to fixed conformations. Therefore, lipid headgroups sample a large space of different conformations also in charged membranes, thereby retaining the capability to adopt multiple conformations when interacting with proteins or other molecules.

## Protein-bound lipid conformations

Interpretation of experimental order parameters using MD simulations in previous sections suggest that PC, PE, PG, and PS lipid headgroups are very flexible, allowing them to

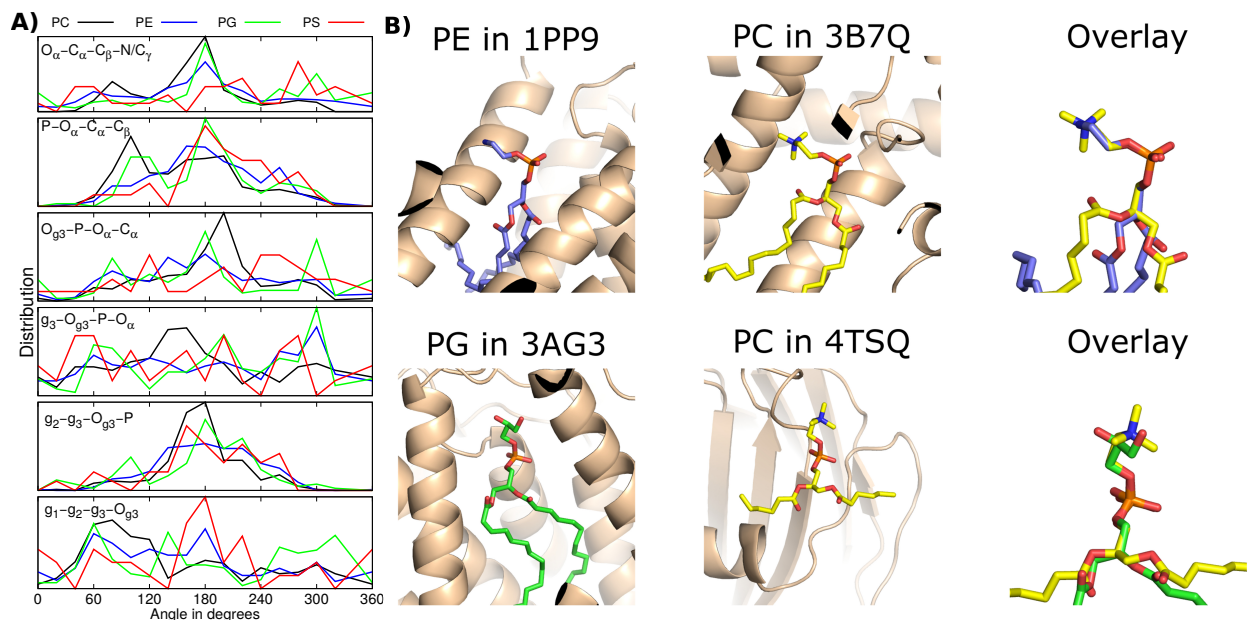


Figure 4: A) Heavy atom dihedral angle distributions calculated from lipid structures in PDB. B) The structure of PE headgroup bound to cytochrome *bc*<sub>1</sub> complex (PDB ID: 1PP9<sup>42</sup>) with identical conformation as the PC headgroup bound to yeast Sec14 (PDB ID: 3B7Q<sup>43</sup>), and the structure of PG headgroup bound to bovine cytochrome *c* oxidase (PDB ID: 3AG3<sup>44</sup>) with identical conformation as PC to FraC, a pore-forming toxin (PDB ID: 4TSQ<sup>45</sup>).

bind proteins in various different conformations. To test this prediction, we analyzed the conformations of all protein-bound lipids from structures deposited in the PDB.<sup>6</sup> We found 311 PC, 394 PE, 154 PG, and 35 PS lipid conformations that were bound to different types of proteins (including both integral and peripheral membrane proteins) and were determined as a part of protein structure using crystallography or cryo-EM. The full list of structures with bound lipids is available from [Data/PDBdistributions/pdbtable.pdf](#) in Ref.<sup>46</sup>

The heavy atom dihedral angle distributions calculated from these conformations (Fig. 4 A) reveal that the protein-bound lipids indeed exhibit a wide range of conformations, independent of the headgroup type. As for bulk lipid bilayers (Figs. 2C, S11), only *cis* conformations of  $P-O_{\alpha}-C_{\alpha}-C_{\beta}$  and  $g_2-g_3-O_{g3}-P$  dihedrals are almost completely absent in all lipids, and significant differences between different headgroups are not observed. Deviation of protein-bound lipid structures from protein-free lipid structures in crystals has been previously proposed to indicate inaccuracies in lipid structures in PDB.<sup>9,22</sup> However, we observe large deviations

from the protein-free lipid crystal structures also in conformational ensembles that reproduce the NMR data in the liquid lamellar phase, indicating that such deviations are realistic also in the protein-bound states.

Our results suggest that flexible lipid headgroups can optimize the intermolecular interactions with proteins by binding in a wide range of conformations. Therefore, the specific binding of lipids to proteins is not driven by differing structural preferences between headgroups. This is demonstrated in Fig. 4 B with two examples where different lipid types bound to different proteins have almost identical headgroup conformations: PE in cytochrome *bc<sub>1</sub>* complex is similar to PC bound to yeast Sec14, and PG bound to bovine cytochrome *c* oxidase is similar to PC bound to pore-forming toxin (FraC). On the other hand, a single lipid headgroup type is capable of accommodating various binding positions and as such would be able to specifically bind to many different kinds of binding sites.

## Conclusions

C–H bond order parameters  $S_{\text{CH}}$  from NMR experiments suggest that lipid headgroup conformational ensembles depend on lipid type (PC, PE, PG, or PS) and membrane charge (cationic lipids, surfactants, accumulated ions or drugs). Our analysis of these data, using MD simulations collected within the NMRlipids Project, revealed that the differences in  $S_{\text{CH}}$  can be explained by small changes in the dihedral angle probability distributions. All four studied headgroup types (PC, PE, PG, and PS) are flexible and access a similarly wide range of conformations, with very few restrictions in dihedral orientations—also when membranes are charged. In conclusion, the observed differences in  $S_{\text{CH}}$  originate from slightly reweighed conformational probabilities rather than changes in accessible structures.

The flexibility and wide conformational space of headgroups suggest that protein-bound lipids can adapt to various binding sites to optimize the lipid–protein interactions. We tested this prediction by analyzing from PDB the conformations of lipids that are tightly bound to

proteins. We found that also the protein-bound lipids exhibit a wide range of conformations, without significant differences between different lipid types: The specificity of lipid binding to proteins is not regulated by accessible lipid structures, but a single lipid type can adapt to various binding pockets.

Our results suggest that lipid headgroup binding to proteins can be described by the *inverse conformational selection model* where unbound lipids sample a wide conformational ensemble offering the large selection of conformations that can bind to various binding sites in substrates, such as protein, drug, RNA, or virus, that remain conformationally fixed. In this model, the binding energy is dominated by intermolecular interactions and changes in lipid conformational entropy, while conformational restrictions inherent to the lipids play a minor role. Here we have focused only on lipid headgroups while potential applications of the proposed model on interactions between hydrophobic acyl chains and proteins are left for future studies. Applications of the proposed model in molecular cell biology include the understanding of lipid-mediated cell signaling and functions of lipid regulated membrane proteins in general. In bionanotechnology, our model facilitates, for example, the design of phospholipid-specific antibodies and lipid nanoparticle carriers with minimal side effects for mRNA vaccines. Furthermore, our work has demonstrated the power of open access MD simulation data to complement the PDB data in elucidating behavior of complex systems of disordered biomolecules.

## Methods

### Experimental C–H bond order parameters

The headgroup and glycerol backbone C–H bond order parameters of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG), purchased from Avanti polar lipids, were measured using natural abundance  $^1\text{H}$ - $^{13}\text{C}$  solid state NMR spectroscopy as described previously.<sup>33,34</sup> The samples

were prepared by simply mixing the lipids with water until a homogeneous preparation of multi-lamellar vesicles (MLVs) was attained and left to equilibrate for approximately 24 hours before measurements. The magnitudes of the order parameters were determined using a R-type Proton Detected Local Field (R-PDLF) experiment,<sup>47</sup> and the order parameter signs from S-DROSS experiments<sup>48</sup> combined with SIMPSON simulations.<sup>49</sup> The set-up of the solid-state NMR experiments were identical as in our previous work.<sup>32</sup> The POPE experiments were recorded at 310 K and POPG experiments at 298 K, where the bilayers are in the liquid disordered phase.<sup>50</sup>

Glycerol backbone peaks from both lipids, and the  $\alpha$ -carbon peak from POPE in the INEPT spectra were assigned based on previously measured POPC spectra.<sup>33</sup> The  $\beta$ -carbon peak from POPE was assigned based on <sup>13</sup>C chemical shift table for amines available at <https://www.chem.wisc.edu/areas/reich/nmr/c13-data/cdata.htm>. The assignment of the  $\alpha$ ,  $\beta$ , and  $\gamma$  peaks of POPG was based on the measured R-PDLF dipolar couplings in comparison to the known order parameters from previous <sup>2</sup>H NMR experiments.<sup>20,35</sup> The  $\beta$ -carbon peak from POPG overlapped with the  $g_2$  peak from glycerol backbone because their chemical environments are similar.

## Molecular dynamics simulations

Molecular dynamics simulation data were collected and analyzed using the methods from the NMRlipids Open Collaboration project ([nmrlipids.blogspot.fi](http://nmrlipids.blogspot.fi)).<sup>21,32,37,51</sup> Simulation details, accessibility information and quality evaluation of approximately 70 combinations of lipid headgroup and force field simulated for this work are in the supplementary information.

Best models for the interpretation of lipid headgroup conformational ensembles from the experimental data were selected using quality evaluation measures defined in the NMRlipids project. Conformational ensembles of headgroup and glycerol backbone in PE and PG simulations were evaluated using the C–H bond order parameters.<sup>37</sup> Interactions between different headgroups were evaluated by monitoring the changes in headgroup order parameters

upon mixing the lipids.<sup>32</sup> The ion binding affinities and response of lipids to bound charge were evaluated by monitoring the changes in lipid headgroup order parameters.<sup>32,51</sup>

The differences in headgroup conformation ensembles were analyzed calculating the distributions of six dihedral angles,  $O_\alpha-C_\alpha-C_\beta-N/C_\gamma$ ,  $P-O_\alpha-C_\alpha-C_\beta$ ,  $O_{g3}-P-O_\alpha-C_\alpha$ ,  $g_3-O_{g3}-P-O_\alpha$ ,  $g_2-g_3-O_{g3}-P$ , and  $g_1-g_2-g_3-O_{g3}$ , labeled in Fig. 1 C. Relative free energy costs for turning dihedral angles with respect to the most probable value (lowest free energy) were estimated from the inverse Boltzmann formula  $\Delta E(\theta) = -kT [\ln [p(\theta)] - \ln [p(\theta_0)]]$ , where  $p(\theta)$  is the dihedral angle distribution and  $\theta_0$  is the most probable angle from MD simulation.

## Analysis of protein-bound lipid conformations

Lipid structures from the Protein Data Bank (PDB,<sup>6</sup>) were searched using PDBe REST API ([www.ebi.ac.uk/pdbe/pdbe-rest-api](http://www.ebi.ac.uk/pdbe/pdbe-rest-api)) using the ligand names listed in the supplementary information. Dihedral distributions from the lipid structures were calculated using the MD-Analysis Python library<sup>52,53</sup> and Jupyter notebook available from *scripts/PDBanalysis.ipynb* folder in Ref 46. Only the structures determined using X-ray crystallography or Cryo-EM with a resolution higher than 3.2 Å were used in the analysis. Some structures of multimeric proteins contained multiple lipids in the same conformation. The conformations were considered identical if all dihedral angles were the same within 3 degrees. In these cases, this conformation was counted only once for the subsequent analyses to avoid its overweighting.

To demonstrate the existence of similar structures of different lipids bound to different proteins, we searched for pairs having the last five heavy atom dihedrals angles in the headgroup end (excluding  $g_1-g_2-g_3-O_{g3}$  from the list in previous section) within 30 degrees of each other among the structures with resolution higher than 2.5 Å. The two most representative examples were handpicked from the results.



## Supporting Information Available

Additional NMR data, ligand names in PDB used to analyse the protein-bound lipid conformations, evaluation of MD simulations against NMR experiments, dihedral angle distributions from simulations, and description of simulated systems with simulation details.

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## Acknowledgement

M.J. thanks CSC – IT Center for Science for computational resources and the Emil Aaltonen foundation for financial support. T.M.F. was supported by the Ministry of Economics, Science and Digitalisation of the State of Saxony-Anhalt, Germany. P.B. was supported by the Academy of Finland (Grant 311031). F.F.-R. acknowledges Tecnológico Nacional de México Proyecto IT16C431, Dirección General de Asuntos del Personal Académico (DGAPA) Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) IG100920, CONACyT Ciencia de Frontera 74884 for financial support and Miztli-Dirección de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) - Universidad Nacional Autónoma de México (UNAM) (Project LANCAD-UNAM-DGTIC-057) facilities for computing-time allocation. I.G. was supported by the Ministry of Science and Higher Education of the Russian Federation (agreement #075-00337-20-03, project FSMG-2020-0003). J.J.M. gratefully acknowledges financial support from the Carlsberg Foundation in the form of a postdoctoral fellowship while at the University of Chicago (grants CF15-0552, CF16-0639, and CF17-0783) and the research framework provided by the Research Computing Center at the University of Chicago. O.H.S.O, A.M.K, and S.I.V acknowledge CSC – IT Center for Science for computational resources and Academy of Finland (grants 315596 and 319902) for financial support. T.J.P. acknowledges use of the Iridis high-performance computing resources at the University of Southampton. J.M. thanks the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high

performance computing cluster. R.G.-F. and A. P. acknowledge the financial support from Fundação para a Ciência e Tecnologia (FCT) and FEDER European funds, through the project PTDC/BIA-BFS/30579/2017 (POCI-01-0145-30579) and UIDB/00081/2020. R.G.-F. thanks Ministerio de Ciencia, Innovación y Universidades for a "Ramón y Cajal" contract (RYC-2016-20335), and also to the Spanish Agencia Estatal de Investigación (AEI) and the ERDF (RTI2018-098795-A-I00) and the Xunta de Galicia and the ERDF (ED431F 2020/05), and Centro singular de investigación de Galicia accreditation 2016-2019, ED431G/09). Á. P. acknowledges to the Spanish Agencia Estatal de Investigación (AEI) and the ERDF (PID2019-111327GB-I00). R.G.-F., A. P. and Á. P. acknowledge to the Centro de Supercomputación de Galicia (CESGA) for technical support and computing time.

## References

- (1) Harayama, T.; Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 281–296.
- (2) Vigant, F.; Santos, N. C.; Lee, B. Broad-spectrum antivirals against viral fusion. *Nat. Rev. Microbiol.* **2015**, *13*, 426–437.
- (3) Pardi, N.; Hogan, M. J.; Porter, F. W.; Weissman, D. mRNA vaccines — a new era in vaccinology. *Nat. Rev. Drug Disc.* **2018**, *17*, 261–279.
- (4) Schoenmaker, L.; Witzigmann, D.; Kulkarni, J. A.; Verbeke, R.; Kersten, G.; Jiskoot, W.; Crommelin, D. J. mRNA-lipid nanoparticle COVID-19 vaccines: Structure and stability. *International Journal of Pharmaceutics* **2021**, *601*, 120586.
- (5) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124.
- (6) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.;

- Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (7) Büldt, G.; Wohlgemuth, R. The headgroup conformation of phospholipids in membranes. *J. Membr. Biol.* **1981**, *58*, 81–100.
- (8) Pascher, I.; Lundmark, M.; Nyholm, P.-G.; Sundell, S. Crystal structures of membrane lipids. *Biochim. Biophys. Acta* **1992**, *1113*, 339 – 373.
- (9) Marsh, D.; Páli, T. Orientation and conformation of lipids in crystals of transmembrane proteins. *Eur. Biophys. J.* **2013**, *42*, 119–146.
- (10) Sormanni, P.; Piovesan, D.; Heller, G. T.; Bonomi, M.; Kukic, P.; Camilloni, C.; Fuxreiter, M.; Dosztanyi, Z.; Pappu, R. V.; Babu, M. M. et al. Simultaneous quantification of protein order and disorder. *Nat. Chem. Biol.* **2017**, *13*, 339–342.
- (11) Buslaev, P.; Gordeliy, V.; Grudinin, S.; Gushchin, I. Principal Component Analysis of Lipid Molecule Conformational Changes in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 1019–1028.
- (12) Ferreira, T. M.; Ollila, O. H. S.; Pigliapochi, R.; Dabkowska, A. P.; Topgaard, D. Model-free estimation of the effective correlation time for C-H bond reorientation in amphiphilic bilayers:  $^1\text{H}$ - $^{13}\text{C}$  solid-state NMR and MD simulations. *J. Chem. Phys.* **2015**, *142*, 044905.
- (13) Antila, H. S.; M. Ferreira, T.; Ollila, O. H. S.; Miettinen, M. S. Using Open Data to Rapidly Benchmark Biomolecular Simulations: Phospholipid Conformational Dynamics. *J. Chem. Inf. Model.* **2021**, *61*, 938–949.
- (14) Seelig, J. Deuterium magnetic resonance: theory and application to lipid membranes. *Q. Rev. Biophys.* **1977**, *10*, 353–418.

- (15) Davis, J. H. The description of membrane lipid conformation, order and dynamics by  $^2\text{H}$ -NMR. *Biochim. Biophys. Acta* **1983**, *737*, 117 – 171.
- (16) Semchyschyn, D. J.; Macdonald, P. M. Conformational response of the phosphatidylcholine headgroup to bilayer surface charge: torsion angle constraints from dipolar and quadrupolar couplings in bicelles. *Magn. Res. Chem.* **2004**, *42*, 89–104.
- (17) Gally, H. U.; Pluschke, G.; Overath, P.; Seelig, J. Structure of Escherichia coli membranes. Glycerol auxotrophs as a tool for the analysis of the phospholipid head-group region by deuterium magnetic resonance. *Biochemistry* **1981**, *20*, 1826–1831.
- (18) Scherer, P.; Seelig, J. Structure and dynamics of the phosphatidylcholine and the phosphatidylethanolamine head group in L-M fibroblasts as studied by deuterium nuclear magnetic resonance. *EMBO J.* **1987**, *6*.
- (19) Seelig, J. Interaction of phospholipids with  $\text{Ca}^{2+}$  ions. On the role of the phospholipid head groups. *Cell Biol. Int. Rep.* **1990**, *14*, 353 – 360.
- (20) Wohlgemuth, R.; Waespe-Sarcevic, N.; Seelig, J. Bilayers of phosphatidylglycerol. A deuterium and phosphorus nuclear magnetic resonance study of the head-group region. *Biochemistry* **1980**, *19*, 3315–3321.
- (21) Ollila, O. S.; Pabst, G. Atomistic resolution structure and dynamics of lipid bilayers in simulations and experiments. *Biochim. Biophys. Acta* **2016**, *1858*, 2512 – 2528.
- (22) Pezeshkian, W.; Khandelia, H.; Marsh, D. Lipid Configurations from Molecular Dynamics Simulations. *Biophys. J.* **2018**, *114*, 1895 – 1907.
- (23) Akutsu, H. Structure and dynamics of phospholipids in membranes elucidated by combined use of NMR and vibrational spectroscopies. *Biochim. Biophys. Acta* **2020**, *1862*, 183352.

- (24) Sohlenkamp, C.; Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev.* **2015**, *40*, 133–159.
- (25) Vance, J. E. Phospholipid Synthesis and Transport in Mammalian Cells. *Traffic* **2015**, *16*, 1–18.
- (26) Calzada, E.; Onguka, O.; Claypool, S. M. Phosphatidylethanolamine Metabolism in Health and Disease. *Int. Rev. Cell Mol. Biol.* **2016**, *321*, 29 – 88.
- (27) Patel, D.; Witt, S. N. Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease. *Oxid. Med. Cell. Longev.* **2017**, *2017*, 4829180.
- (28) Lemmon, M. A. Membrane recognition by phospholipid-binding domains. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 99–111.
- (29) Leventis, P. A.; Grinstein, S. The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys.* **2010**, *39*, 407–427.
- (30) Hariharan, P.; Tikhonova, E.; Medeiros-Silva, J.; Jeucken, A.; Bogdanov, M. V.; Dowhan, W.; Brouwers, J. F.; Weingarth, M.; Guan, L. Structural and functional characterization of protein–lipid interactions of the Salmonella typhimurium melibiose transporter MelB. *BMC Biology* **2018**, *16*, 85.
- (31) Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochim. Biophys. Acta* **2014**, *1838*, 1548 – 1559.
- (32) Antila, H.; Buslaev, P.; Favela-Rosales, F.; Ferreira, T. M.; Gushchin, I.; Javanainen, M.; Kav, B.; Madsen, J. J.; Melcr, J.; Miettinen, M. S. et al. Headgroup Structure and Cation Binding in Phosphatidylserine Lipid Bilayers. *J. Phys. Chem. B* **2019**, *123*, 9066–9079.
- (33) Ferreira, T. M.; Coreta-Gomes, F.; Ollila, O. H. S.; Moreno, M. J.; Vaz, W. L. C.; Topgaard, D. Cholesterol and POPC segmental order parameters in lipid membranes:

- solid state  $^1\text{H}$ - $^{13}\text{C}$  NMR and MD simulation studies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1976–1989.
- (34) Ferreira, T. M.; Sood, R.; Bärenwald, R.; Carlström, G.; Topgaard, D.; Saalwächter, K.; Kinnunen, P. K. J.; Ollila, O. H. S. Acyl Chain Disorder and Azelaoyl Orientation in Lipid Membranes Containing Oxidized Lipids. *Langmuir* **2016**, *32*, 6524–6533.
- (35) Borle, F.; Seelig, J.  $\text{Ca}^{2+}$  binding to phosphatidylglycerol bilayers as studied by differential scanning calorimetry and  $^2\text{H}$ - and  $^{31}\text{P}$ -nuclear magnetic resonance. *Chem. Phys. Lipids* **1985**, *36*, 263 – 283.
- (36) Seelig, J.; Gally, H. U. Investigation of phosphatidylethanolamine bilayers by deuterium and phosphorus-31 nuclear magnetic resonance. *Biochemistry* **1976**, *15*, 5199–5204.
- (37) Botan, A.; Favela-Rosales, F.; Fuchs, P. F. J.; Javanainen, M.; Kanduč, M.; Kulig, W.; Lamberg, A.; Loison, C.; Lyubartsev, A.; Miettinen, M. S. et al. Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *J. Phys. Chem. B* **2015**, *119*, 15075–15088.
- (38) Browning, J. L.; Seelig, J. Bilayers of phosphatidylserine: a deuterium and phosphorus nuclear magnetic resonance study. *Biochemistry* **1980**, *19*, 1262–1270.
- (39) Scherer, P. G.; Seelig, J. Electric charge effects on phospholipid headgroups. Phosphatidylcholine in mixtures with cationic and anionic amphiphiles. *Biochemistry* **1989**, *28*, 7720–7728.
- (40) Seelig, J.; MacDonald, P. M.; Scherer, P. G. Phospholipid head groups as sensors of electric charge in membranes. *Biochemistry* **1987**, *26*, 7535–7541.
- (41) Melcr, J.; Ferreira, T. M.; Jungwirth, P.; Ollila, O. H. S. Improved Cation Binding to Lipid Bilayers with Negatively Charged POPS by Effective Inclusion of Electronic Polarization. *J. Chem. Theo. Comput.* **2020**, *16*, 738–748.

- (42) Huang, L.; Cobessi, D.; Tung, E. Y.; Berry, E. A. Binding of the Respiratory Chain Inhibitor Antimycin to the Mitochondrial bc<sub>1</sub> Complex: A New Crystal Structure Reveals an Altered Intramolecular Hydrogen-bonding Pattern. *J. Mol. Biol.* **2005**, *351*, 573–597.
- (43) Schaaf, G.; Ortlund, E. A.; Tyeryar, K. R.; Mousley, C. J.; Ile, K. E.; Garrett, T. A.; Ren, J.; Woolls, M. J.; Raetz, C. R.; Redinbo, M. R. et al. Functional Anatomy of Phospholipid Binding and Regulation of Phosphoinositide Homeostasis by Proteins of the Sec14 Superfamily. *Mol. Cell* **2008**, *29*, 191–206.
- (44) Muramoto, K.; Ohta, K.; Shinzawa-Itoh, K.; Kanda, K.; Taniguchi, M.; Nabekura, H.; Yamashita, E.; Tsukihara, T.; Yoshikawa, S. Bovine cytochrome c oxidase structures enable O<sub>2</sub> reduction with minimization of reactive oxygens and provide a proton-pumping gate. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 7740–7745.
- (45) Tanaka, K.; Caaveiro, J. M.; Morante, K.; González-Mañas, J. M.; Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* **2015**, *6*, 6337.
- (46) Ollila, O. H. S.; et al., NMRlipidsIVb GitHub repository. <https://doi.org/10.5281/zenodo.5118637>.
- (47) Dvinskikh, S. V.; Zimmermann, H.; Maliniak, A.; Sandstrom, D. Measurements of motionally averaged heteronuclear dipolar couplings in MAS NMR using R-type recoupling. *J. Magn. Reson.* **2004**, *168*, 194–201.
- (48) Gross, J. D.; Warschawski, D. E.; Griffin, R. G. Dipolar Recoupling in MAS NMR: A Probe for Segmental Order in Lipid Bilayers. *J. Am. Chem. Soc.* **1997**, *119*, 796–802.
- (49) Bak, M.; Rasmussen, J. T.; Nielsen, N. C. SIMPSON: A General Simulation Program for Solid-State NMR Spectroscopy. *J. Magn. Reson.* **2000**, *147*, 296 – 330.

- (50) Marsh, D. *Handbook of Lipid Bilayers, Second Edition*; RSC press, 2013.
- (51) Catte, A.; Giryck, M.; Javanainen, M.; Loison, C.; Melcr, J.; Miettinen, M. S.; Monticelli, L.; Maatta, J.; Oganessian, V. S.; Ollila, O. H. S. et al. Molecular electrometer and binding of cations to phospholipid bilayers. *Phys. Chem. Chem. Phys.* **2016**, *18*, 32560–32569.
- (52) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (53) Richard J. Gowers,; Max Linke,; Jonathan Barnoud,; Tyler J. E. Reddy,; Manuel N. Melo,; Sean L. Seyler,; Jan Domański,; David L. Dotson,; Sébastien Buchoux,; Ian M. Kenney, et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the 15th Python in Science Conference. 2016; pp 98 – 105.



# Graphical TOC Entry

