

SUPPLEMENTARY INFORMATION: NMRLipids

Databank makes data-driven analysis of biomembrane properties accessible for all

Anne M. Kiirikki¹, Hanne S. Antila², Lara Bort³, Pavel Buslaev⁴, Fernando Favela⁵, Tiago Mendes Ferreira⁶, Patrick F.J. Fuchs^{7,8}, Rebeca Garcia-Fandino⁹, Ivan Gushchin¹⁰, Batuhan Kav^{11,12}, Patrik Kula¹³, Milla Kurki¹⁴, Alexander Kuzmin¹⁰, Jesper J. Madsen^{15,16}, Markus S. Miettinen^{17,18}, Ricky Nencini¹, Thomas Piggot¹⁹, Ángel Piñeiro²⁰, Suman Samantray¹⁰, Fabián Suarez-Leston^{9,20,21}, and O. H. Samuli Ollila^{1,*}

¹University of Helsinki, Institute of Biotechnology, Helsinki, Finland

²Department of Theory & Bio-Systems and Department of Biomaterials, Max Planck Institute of Colloids and Interfaces, Germany

³University of Potsdam, Institute of Physics and Astronomy, Potsdam-Golm, 14476, Germany

⁴Nanoscience Center and Department of Chemistry, University of Jyväskylä, P.O. Box 35, Jyväskylä, 40014 , Finland

⁵Departamento de Ciencias Básicas, Tecnológico Nacional de México - ITS Zacatecas Occidente, Sombrerete, Zacatecas, 99102, México

⁶NMR group - Institute for Physics, Martin Luther University Halle-Wittenberg, Halle (Saale), 06120, Germany

⁷Sorbonne Université, Ecole Normale Supérieure, PSL University, CNRS, Laboratoire des Biomolécules (LBM), Paris, 75005, France

⁸Université Paris Cité, UFR Sciences du Vivant, Paris, 75013, France

⁹Center for Research in Biological Chemistry and Molecular Materials (CiQUS), Universidade de Santiago de Compostela, Santiago de Compostela, E-15782, Spain

¹⁰no affiliation

¹¹Institute of Biological Information Processing: Structural Biochemistry (IBI-7), Forschungszentrum Jülich, Jülich 52428, Germany

¹²ariadne.ai GmbH (Germany), Häusserstraße 3 Heidelberg 69115, Germany

¹³Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo nám. 542/2, Prague, CZ-16610, Czech Republic

¹⁴School of Pharmacy, University of Eastern Finland, 70211 Kuopio, Finland

¹⁵Global and Planetary Health, College of Public Health, University of South Florida, Tampa, Florida, 33612, United States of America

¹⁶Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, Florida, 33612, United States of America

¹⁷Department of Chemistry, University of Bergen, Norway

¹⁸Computational Biology Unit, Department of Informatics, University of Bergen, Norway

¹⁹Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom

²⁰Department of Applied Physics, Faculty of Physics, University of Santiago de Compostela, Santiago de Compostela, E-15782, Spain

²¹MD.USE Innovations S.L., Edificio Emprendia, 15782 Santiago de Compostela, Spain

*samuli.ollila@helsinki.fi

ABSTRACT

Correlations between area per lipid and thickness with order parameters and form factors

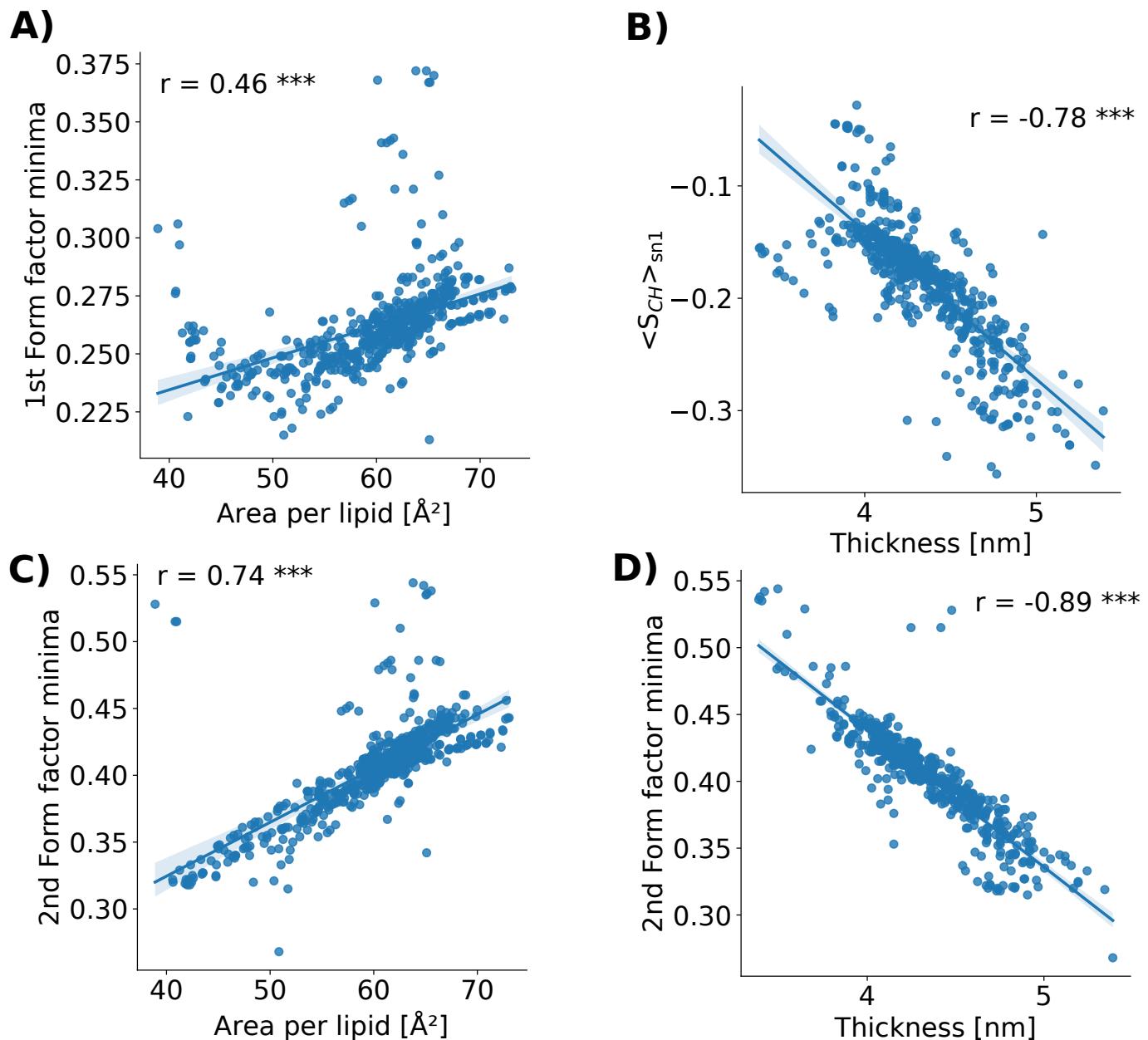


Figure S1. Scatter plots and Pearson correlation coefficients, r , for the membrane area per lipid with X-ray scattering form factor minima (A and B), and for thickness with the average order parameter of the sn-1 acyl chain (B) and with the second minimum from X-ray scattering form factors (D) extracted from the NMRLipids databank. All correlation coefficients have p-value below 0.001 as indicated by ***.

Dependence of form factor and order parameters in the size of simulation

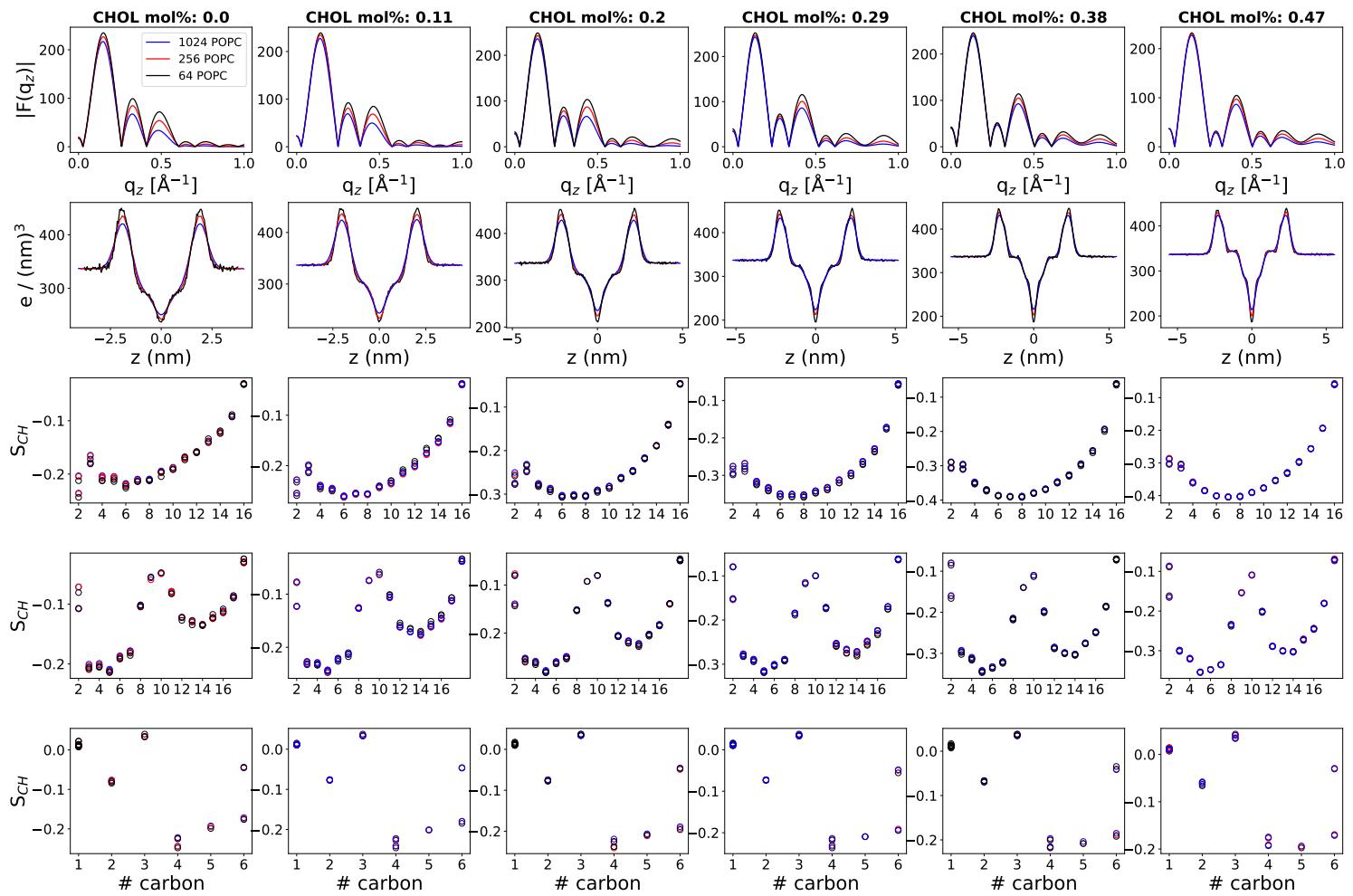


Figure S2. Dependence of form factor, electron density profiles along membrane normal, and C-H bond orderparameters (from top to bottom) on simulation box size with different cholesterol concentration. Simulations with 64, 256 or 1024 POPC lipids taken from Ref. 1.

Finding the best models for PC and PE mixtures

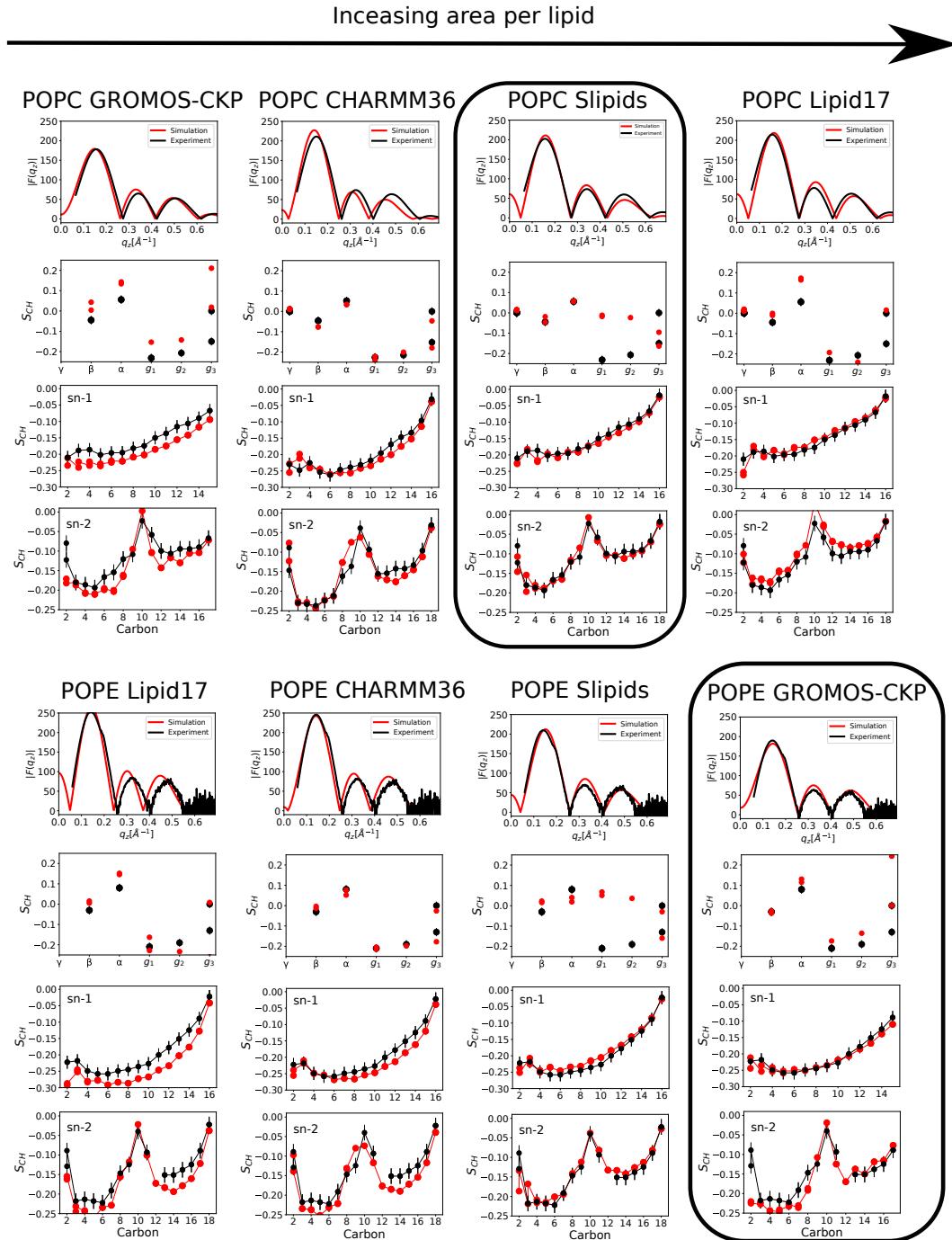


Figure S3. Simulations with the data for both POPC (top) and POPE (bottom) directly compared with the experimental data. The area per lipid in simulations increase from left to right. Simulations with the best overall quality for POPC and POPE order parameters are circled by a square.

Water permeation through membranes

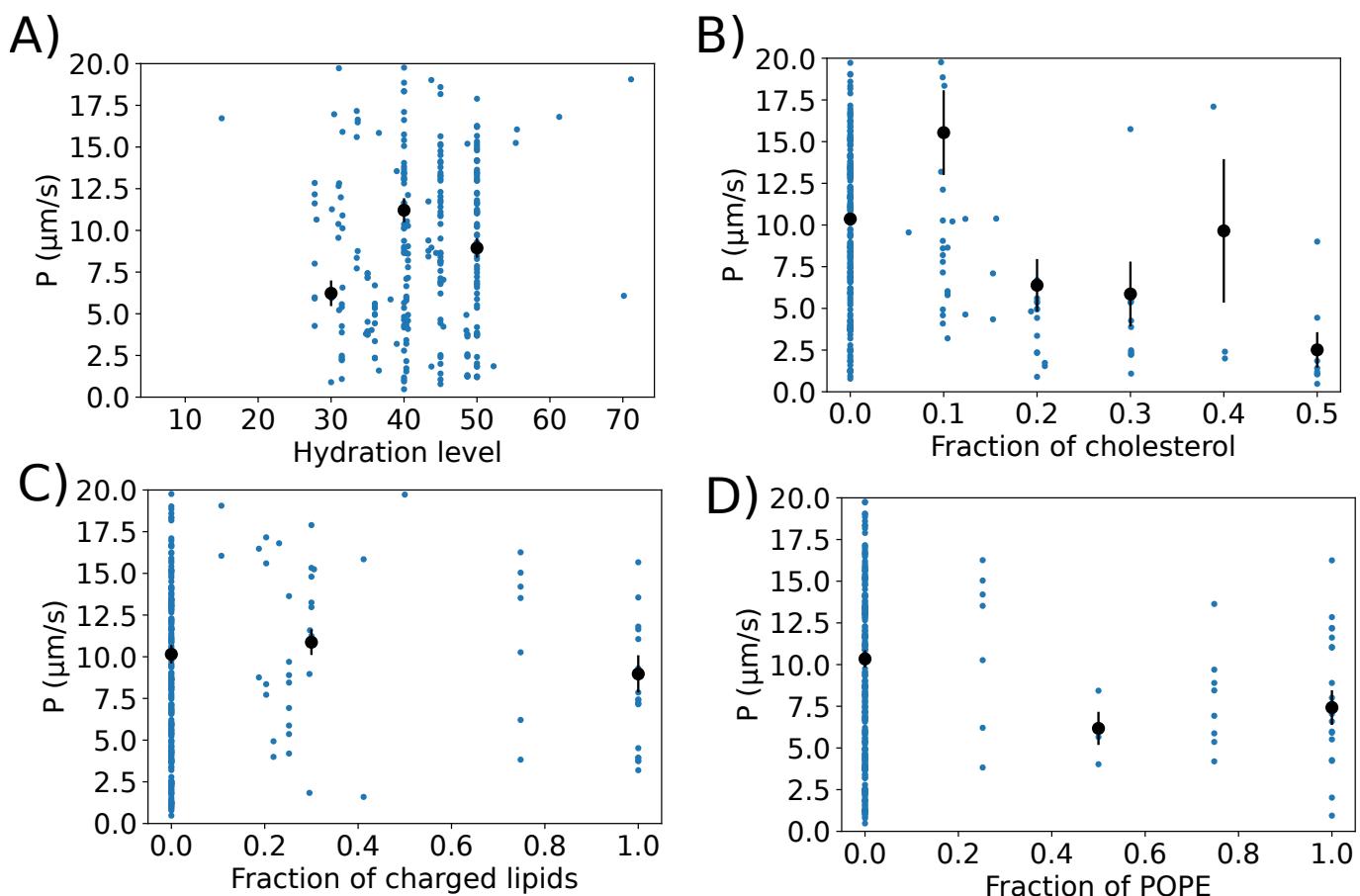


Figure S4. Water permeation through membranes analyzed from the databank as a function of (A) hydration level, (B) fraction of cholesterol, (C) fraction of charged lipids, and (D) fraction of POPE in membrane. Values from simulations with non-zero permeation values are shown with blue dots. Histogrammed values are shown with black dots.

Water diffusion along membranes

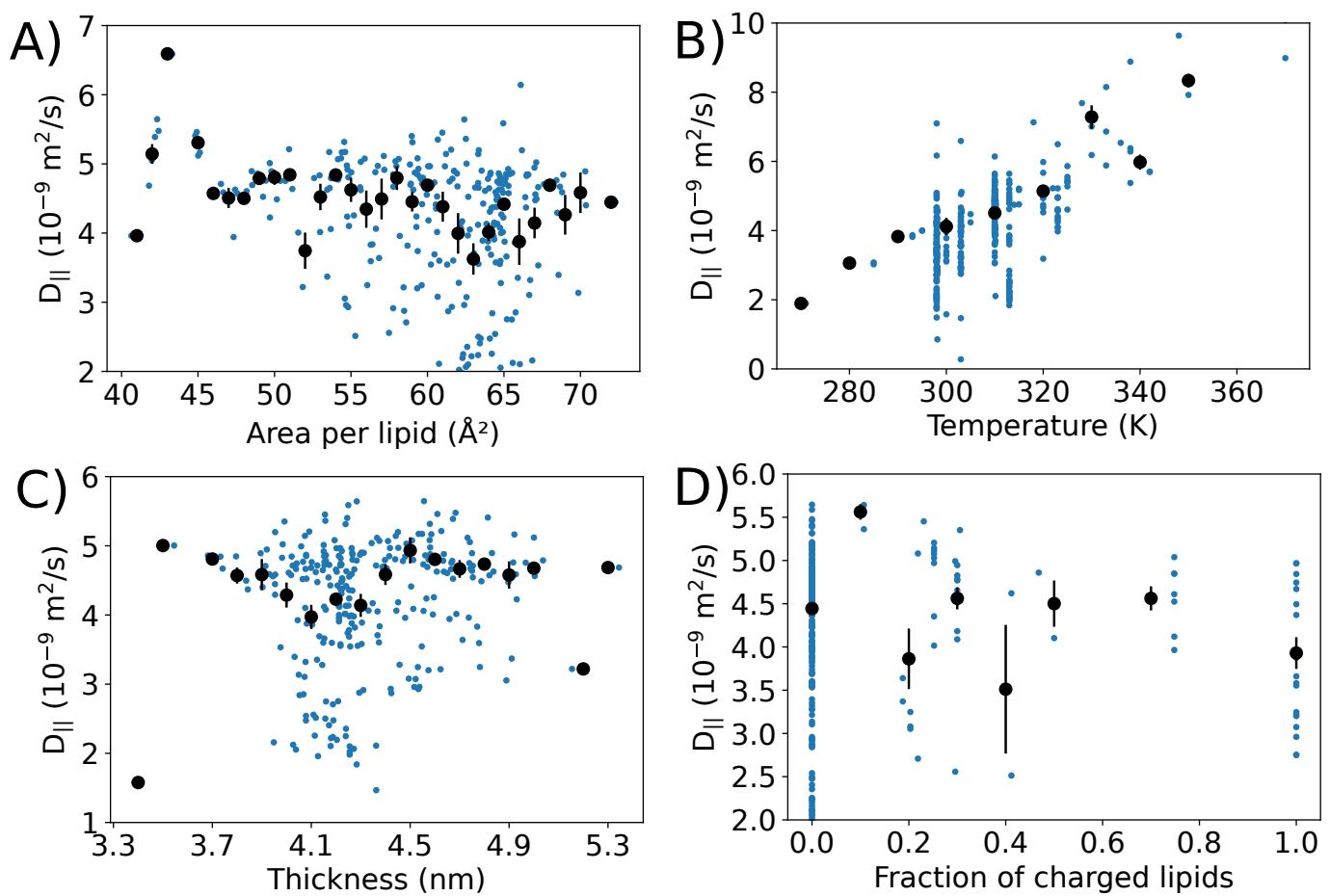


Figure S5. Lateral diffusion of water as a function of (A) area per lipid, (B) temperature, (C) membrane thickness, and (D) fraction of charged lipids in a membrane.

Databank content

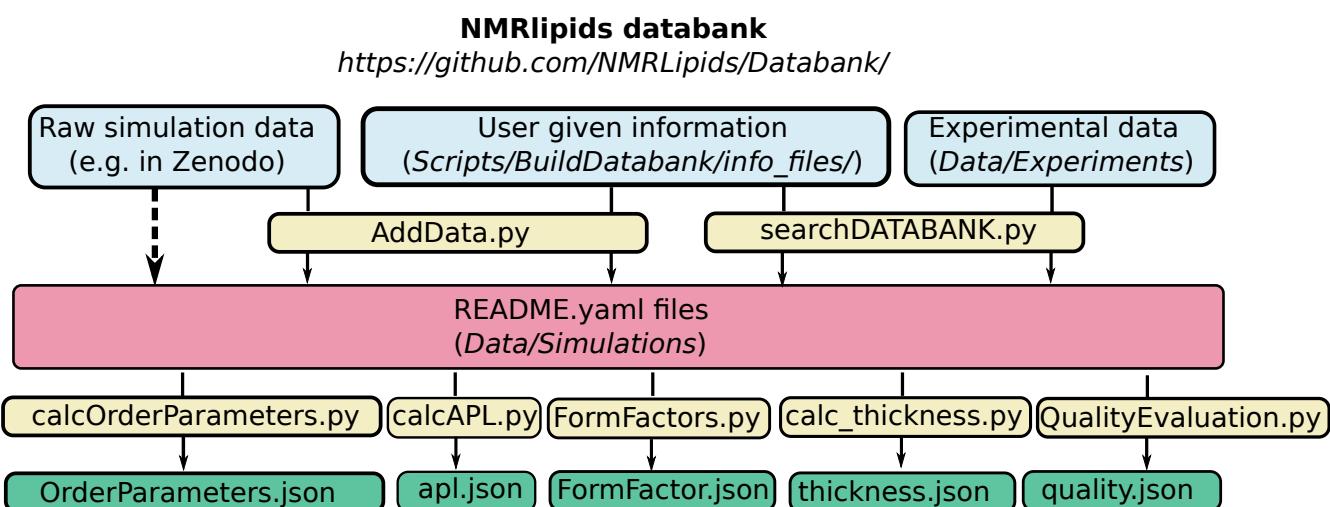


Figure S6. Structure of the NMRLipids databank. Manually added input data (blue boxes) includes basic information on the simulation, permanent links to the raw data, and experimental data if available. The databank entries (red box) and analysis results (green boxes), locating at <https://github.com/NMRLipids/Databank/tree/main/Data/Simulations> are automatically generated by the computer programs included in the NMRLipids databank (yellow boxes). Because raw data are not permanently stored but can be accessed based on the information in the databank, this connection is marked with the dashed line.

Code	Function	Input	Output
Databank layer			
Scripts/BuildDatabank/			
AddData.py	Create README.yaml from user given information in info.yaml	info.yaml	README.yaml
searchDATABANK.py	Pair simulations with available experimental data	README.yaml files	Updated README.yaml files
QualityEvaluation.py	Quality evaluate simulations that are paired with experimental data	README.yaml files	[lipid_name]_OrderParameters_quality.json, [lipid_name]_FragmentQuality.json, system_quality.json, FormFactorQuality.json
Scripts/AnalyzeDatabank/			
calcOrderParameters.py	Calculate C-H bond order parameters for all lipids in the simulation	README.yaml files	[lipid_name]_OrderParameters.json
calcAPL.py	Calculate area per lipid as a function of time	README.yaml files	apl.json
calc_FormFactors.py	Calculate x-ray scattering form factors	README.yaml files	FormFactor.json
calc_thickness.py	Calculate membrane thickness	README.yaml files	thickness.json

Table S1. Caption

key	description	type
DOI	DOI from where the raw data is found	user given (compulsory)
SOFTWARE	Software used to run the simulation (e.g. Gromacs, Amber, NAMD, etc.)	
TRJ	Name of the trajectory file found from DOI (trr or xtc for Gromacs, dcd for OpenMM)	
TPR (Gromacs)	Name of the tpr topology file found from DOI for Gromacs simulations	
PDB (OpenMM)	Name of the pdb file found from DOI for OpenMM simulations	
PREEQTIME	Pre-equilibrate time simulated before the uploaded trajectory in nanoseconds.	
TIMELEFTOUT	Equilibration period in the uploaded trajectory that should be discarded in analyses.	
COMPOSITION	Dictionary connecting universal molecule and atom names to the ones used in simulation	
DIR_WRK	Temporary working directory in your local computer.	
UNITEDATOM_DICT	Information for constucting hydrogens for united atom simulations using buildH program ² . Empty for all atom simulations.	
TYPEOFSYSTEM	Lipid bilayer or something else	
PUBLICATION	Give reference to a publication(s) related to the data.	User given (optional)
AUTHORS_CONTACT	Name and email of the main author(s) of the data.	
SYSTEM	System description on free text format	
SOFTWARE_VERSION	Version of the used software	
FF	Name of the used force field	
FF_SOURCE	Source of the force field parameters, e.g, CHARMM-GUI, webpage, citation to a publication, etc.	
FF_DATE	Date when force field parameters were accessed on the gives source (day/month/year).	
FFmolename	Molecule specific force field information, e.g., water model with FFSOL and sodium parameters with FFSOD.	
CPT	Name of the Gromacs checkpoint file.	
LOG	Name of the Gromacs log file.	
GRO	Name of the Gromacs gro file.	
TOP	Name of top file for Gromacs or psf file for OpenMM.	
CRD	Name of crd file for OpenMM.	
WARNINGS	Dictionary containing information about unusual features in the trajectory, such as ambiguous atom names, membrane normal not oriented in z-direction, old Gromacs version used, etc.	
TRAJECTORY_SIZE	Size of the trajectory file in bytes	automatically extracted data.
TRJLENGTH	Lenght of the trajectory (ps).	
TEMPERATURE	Temperature of the simulation.	
NUMBER_OF_ATOMS	Number of atoms in the simulation.	
DATEOFRUNNIG	Date when added into the databank	
EXPERIMENT	Potentially connected experimental data	
COMPOSITION	Numbers of lipid molecules in both leaflets and numbers of other molecules are added to the dictionary.	
ID	Unique ID number to ease the analyses.	

Table S2. Keys stored in the README.yaml files of simulations.

key	description
DOI	DOI of the publication related to the experimental data.
TEMPERATURE	Temperature of the experiment.
MOLAR_FRACTIONS	Dictionary of molar fractions of bilayer components
ION_CONCENTRATIONS	Dictionary of ion concentrations of the system
TOTAL_LIPID_CONCENTRATION	Total concentration of lipid components. If exact concentration is not known, but experiments are performed in excess water, 'full hydration' can be given.
COUNTER_IONS	Type of counter ions if present.

Table S3. Keys stored in the README.yaml files of experiments.

Abbreviation	Molecule name
Lipids and surfactants	
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
PYPC	1-(16:0)-2-(16:1 ^A 9)-sn-glycero-3-phosphocholine
PAzePCprot	1-palmitoyl-2-azelaoyl-sn-glycero-3-phosphocholine protonated
PAzePCdeprot	1-palmitoyl-2-azelaoyl-sn-glycero-3-phosphocholine deprotonated
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
DPPE	1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine
DPPG	1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt)
DEPC	1,2-dierucoyl-sn-glycero-3-phosphocholine
DRPC	1,2-(14:1 ^A 9)-sn-glycero-3-phosphocholine
DYPC	1,2-(16:1 ^A 9)-sn-glycero-3-phosphocholine
DLPC	1,2-dilauroyl-sn-glycero-3-phosphocholine
DLIPC	1,2-dilinoleoyl-sn-glycero-3-phosphocholine
DOG	1,2-dioleoyl-sn-glycerol
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
DDOPC	1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine
DOPS	1,2-dioleoyl-sn-glycero-3-phospho-L-serine
DSPC	1,2-distearoyl-sn-glycero-3-phosphocholine
DAPC	1,2-diarachidonoyl-sn-glycero-3-phosphocholine
SLiPC	1-(18:0)-2-(18:2 ^A 9, ^B 12)-sn-glycero-3-phosphocholine
DMTAP	1,2-dimyristoyl-3-trimethylammonium-propane
SOPC	1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPI	
SAPI	
SLPI	
SDG	1-stearoyl-2-docosahexaenoyl-sn-glycerol
SDPE	1-stearoyl-2-docosahexaenoyl-sn-glycero-3-phosphoethanolamine
CER	N-palmitoyl-D-erythro-sphingosine
CHOL	cholesterol
DCHOL	18,19-di-nor-cholesterol
DHMDMAB	dihexadecyldimethylammonium
Other molecules	
POT	potassium ion
SOD	sodium ion
CLA	chloride ion
CAL	calcium ion
CES	caesium ion
SOL	water

Table S4. Abbreviations for molecules used in the databank

Force field name and references
CHARMM36 ³
Slipids ^{4–8}
MacRog ⁹
Amber Lipid14/17 ^{10,11}
Charmm-Drude ¹²
ECClipids ^{13–15}
GROMOS-CKP ^{16–18}
Berger ¹⁹
ECC-CHARMM36 ²⁰
Orange ²¹
Poger ²²
GROMOS 43A1-S3 ²³
GAFFlipid ²⁴
OPLS3e ²⁵
Ulmschneider ²⁶
Chiu Gromos ²³

Table S5. List of current force fields used in simulations in the databank with references.

NMR experiments

Acyl chain order parameters of POPE and POPG were analyzed from the experiments that were previously used to determine headgroup order parameters¹⁵. The analysis of the crowded spectral region at 29-31 ppm was based on the previous assignment reported for POPC membranes²⁷.

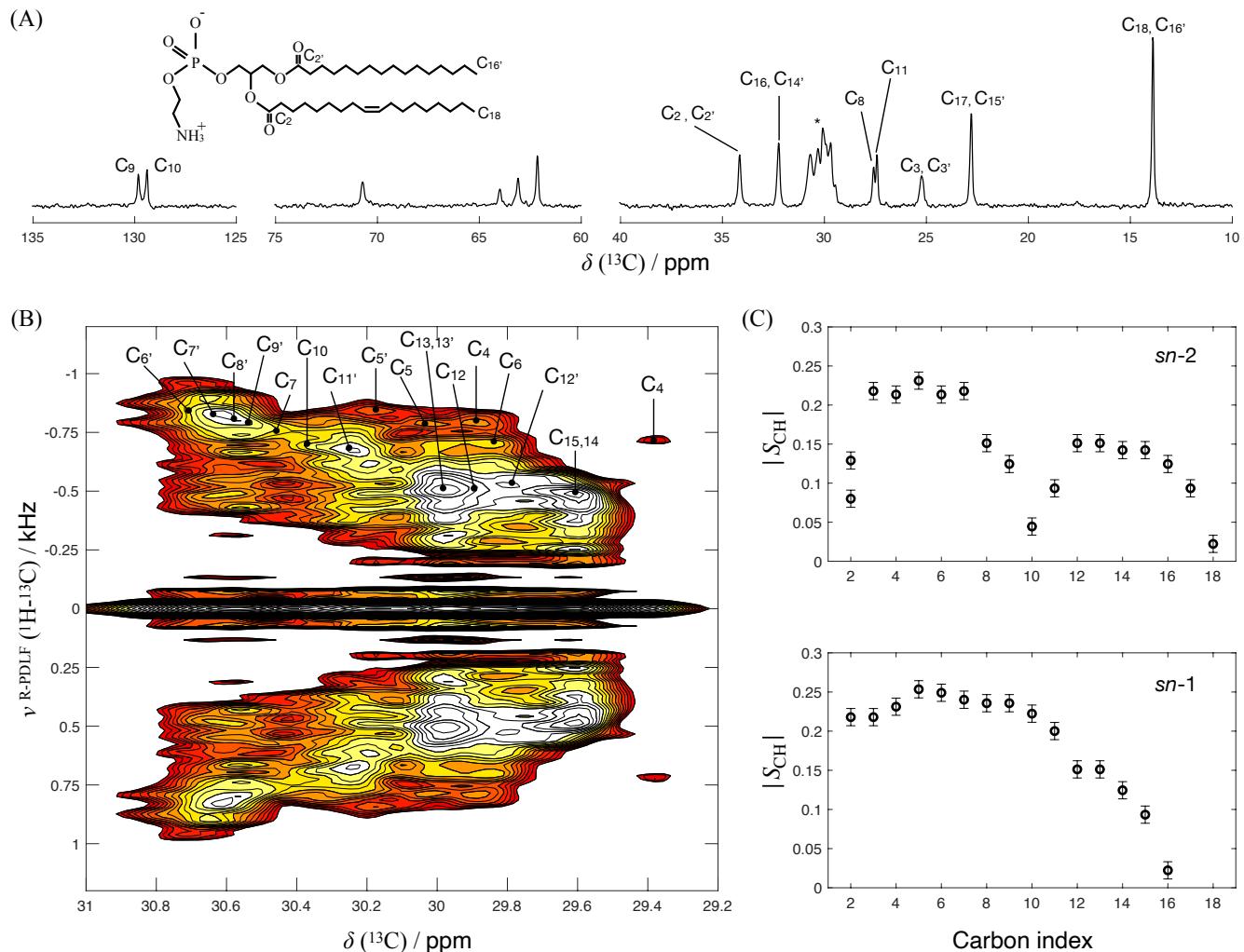


Figure S7. Determination of the POPE acyl chain order parameters from a R-PDLF spectrum measured at a magic angle spinning frequency of 5.15 kHz. (A) ^{13}C rINEPT spectrum with peak assignment. The labels used are shown in the chemical structure of POPE. The chemical shift of the methyl groups was defined as 13.8 ppm. (B) Contour plot of the R-PDLF spectrum for the crowded spectral region. The assignment was based on a previous assignment reported for POPC membranes²⁷. (C) C-H order parameter profile for the acyl chains of POPE. The splittings used for calculating the order parameters are shown in Fig. S8. The unassigned peaks belong to the headgroup and glycerol backbone carbons. A detailed assignment and order parameter analysis of these carbons was shown previously¹⁵.

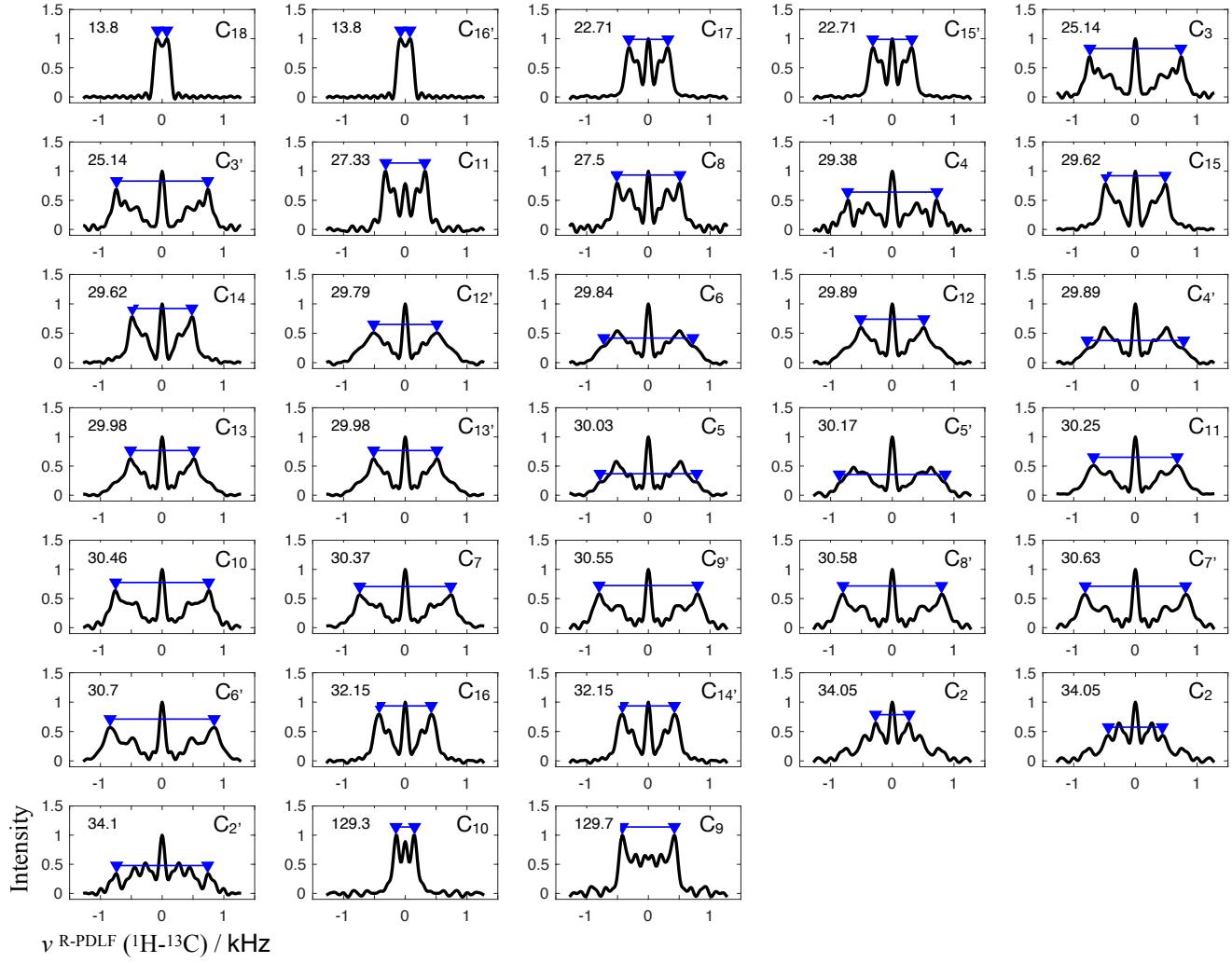


Figure S8. Dipolar spectra obtained from the 2D R-PDLF spectrum from POPE in Fig. S7. The number at the top left of each plot denotes the corresponding chemical shift. The carbon label for each splitting is displayed on the top right. The labels are the same as in Fig. S7.

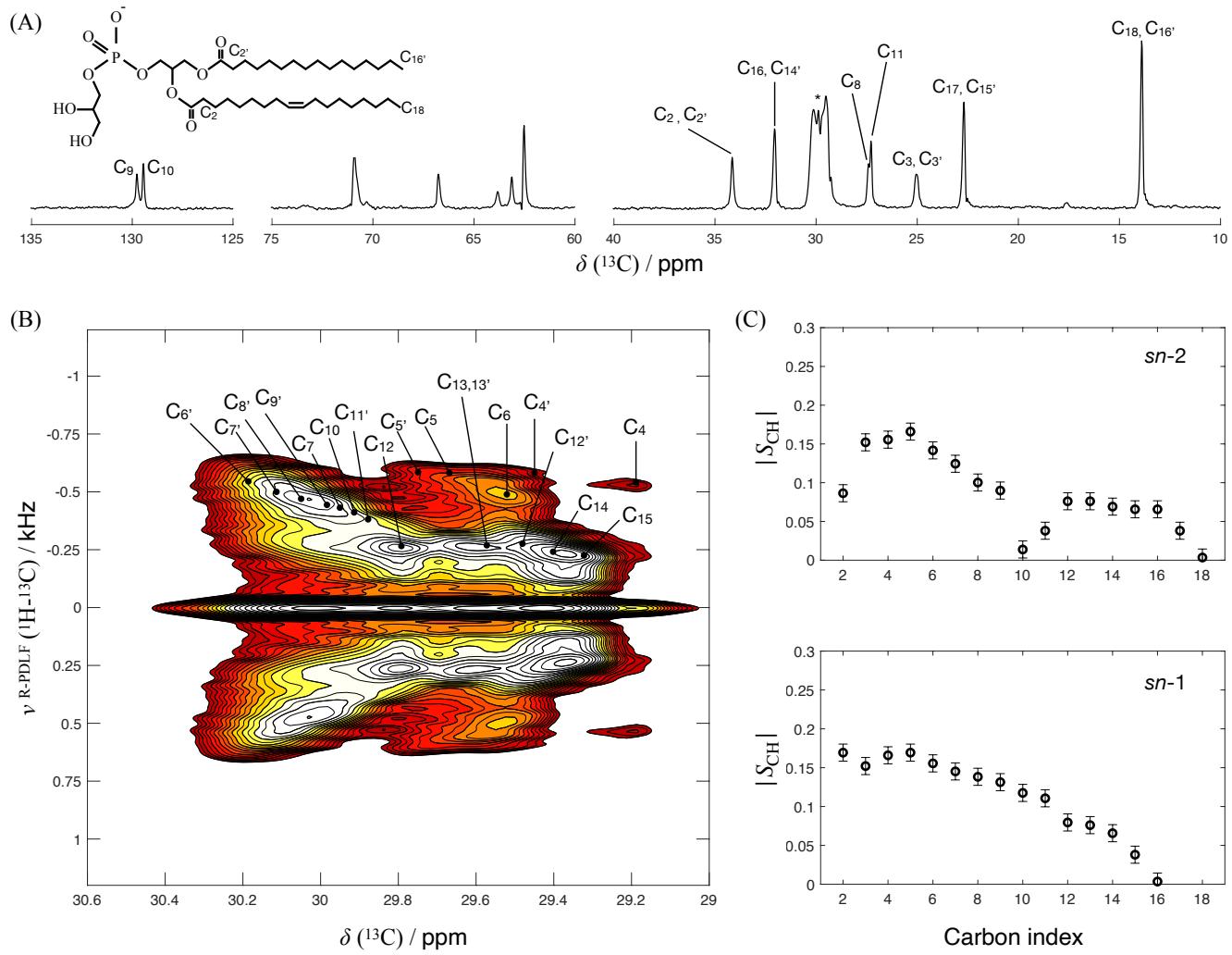


Figure S9. Determination of the POPG acyl chain order parameters from a R-PDLF spectrum measured at a magic angle spinning frequency of 5.15 kHz. (A) ^{13}C rINEPT spectrum with peak assignment. The labels used are shown in the chemical structure of POPG. The chemical shift of the methyl groups was defined as 13.8 ppm. (B) Contour plot of the R-PDLF spectrum for the crowded spectral region. The assignment was based on a previous assignment reported for POPC membranes²⁷. (C) C-H order parameter profile for the acyl chains of POPG. The splittings used for calculating the order parameters are shown in Fig. S10. The unassigned peaks belong to the headgroup and glycerol backbone carbons. A detailed assignment and order parameter analysis of these carbons was shown previously¹⁵.

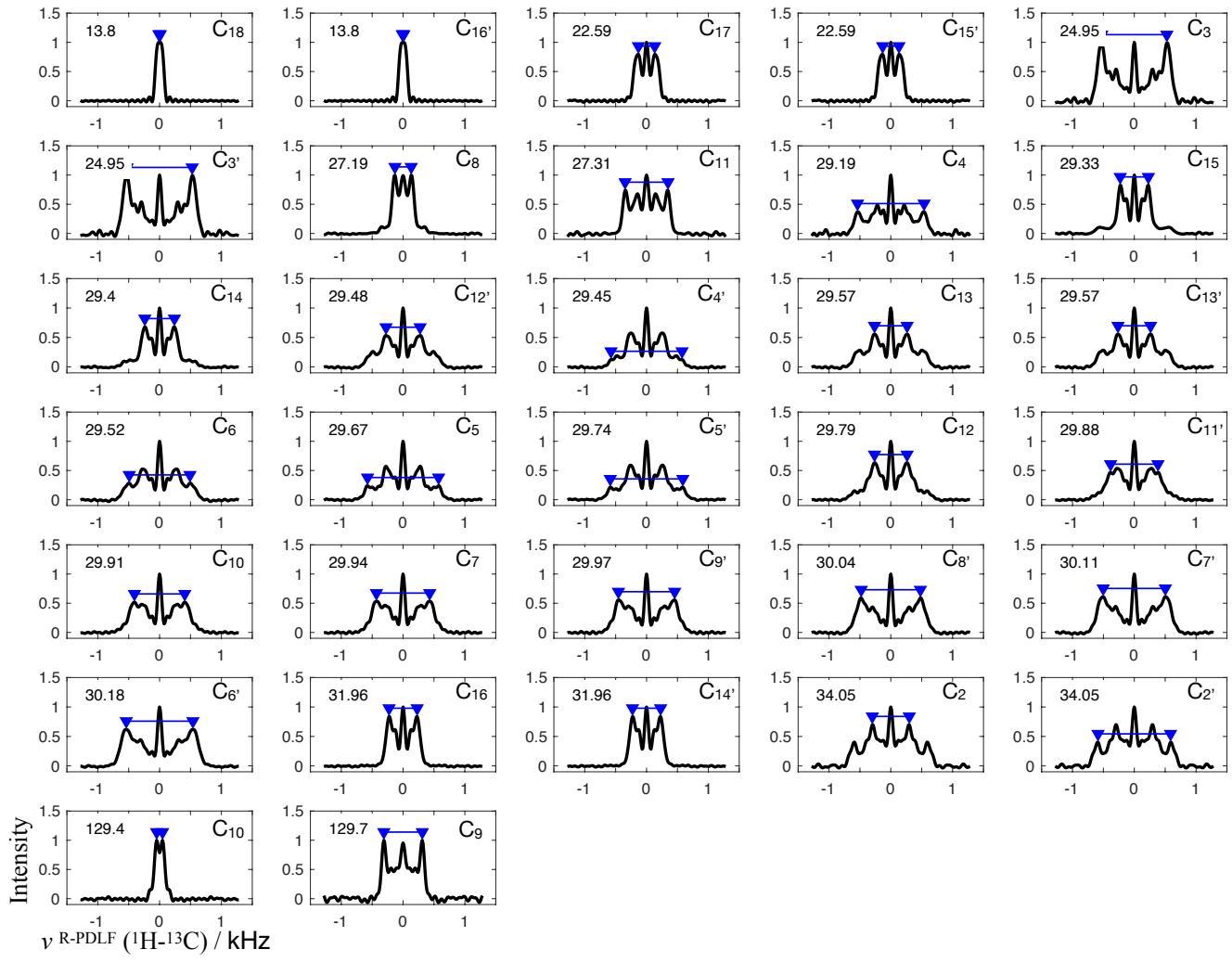


Figure S10. Dipolar spectra obtained from the 2D R-PDLF spectrum described in Fig. S9. The number at the top left of each plot denotes the corresponding chemical shift. The carbon label for each splitting is displayed on the top right. The labels are the same as in Fig. S9.

References

1. Javanainen, M. Simulations of POPC/cholesterol mixtures at 298 K, three system sizes, CHARMM36, DOI: [10.5281/zenodo.7035350](https://doi.org/10.5281/zenodo.7035350) (2021).
2. Santuz, H., Bacle, A., Poulain, P. & Fuchs, P. F. buildh: Build hydrogen atoms from united-atom molecular dynamics of lipids and calculate the order parameters. *J. Open Source Softw.* **6**, 3521, DOI: [10.21105/joss.03521](https://doi.org/10.21105/joss.03521) (2021).
3. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
4. Jämbeck, J. P. M. & Lyubartsev, A. P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B* **116**, 3164–3179 (2012).
5. Jämbeck, J. P. M. & Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **8**, 2938–2948 (2012).
6. Jämbeck, J. P. & Lyubartsev, A. P. Another piece of the membrane puzzle: extending lipids further. *J. Chem. Theo. Comput.* **9**, 774–784 (2012).
7. Ermilova, I. & Lyubartsev, A. P. Extension of the lipids force field to polyunsaturated lipids. *The J. Phys. Chem. B* **120**, 12826–12842 (2016).
8. Grote, F. & Lyubartsev, A. P. Optimization of lipids force field parameters describing headgroups of phospholipids. *J. Phys. Chem. B* **124**, 8784–8793 (2020).
9. Kulig, W., Pasenkiewicz-Gierula, M. & Róg, T. Cis and trans unsaturated phosphatidylcholine bilayers: A molecular dynamics simulation study. *Chem. Phys. Lipids* **195**, 12 – 20 (2016).
10. Dickson, C. J. *et al.* Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
11. Dickson, C. J., Walker, R. C. & Gould, I. R. Lipid21: Complex lipid membrane simulations with amber. *J. Chem. Theo. Comput.* **18**, 1726–1736 (2022).
12. Li, H. *et al.* Drude polarizable force field for molecular dynamics simulations of saturated and unsaturated zwitterionic lipids. *J. chemical theory computation* **13**, 4535–4552 (2017).
13. Melcr, J. *et al.* Accurate binding of sodium and calcium to a popc bilayer by effective inclusion of electronic polarization. *J. Phys. Chem. B* **122**, 4546–4557 (2018).
14. 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.* **16**, 738–748 (2020).
15. Bacle, A. *et al.* Inverse conformational selection in lipid–protein binding. *J. Am. Chem. Soc.* **143**, 13701–13709 (2021).
16. Chandrasekhar, I. *et al.* A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the gromos96 45a3 force field. *Eur. Biophys. J.* **32**, 67–77 (2003).
17. Kukol, A. Lipid models for united-atom molecular dynamics simulations of proteins. *J. Chem. Theory Comput.* **5**, 615–626 (2009).
18. Piggot, T. J., Piñeiro, Á. & Khalid, S. Molecular dynamics simulations of phosphatidylcholine membranes: A comparative force field study. *J. Chem. Theory Comput.* **8**, 4593–4609 (2012).
19. Berger, O., Edholm, O. & Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **72**, 2002 – 2013 (1997).
20. Nencini, R. *Development and testing of computer models of phospholipid membranes*. Master's thesis, Charles University of Prague, DOI: <http://hdl.handle.net/20.500.11956/106127> (2019). [Http://hdl.handle.net/20.500.11956/106127](http://hdl.handle.net/20.500.11956/106127).
21. Catte, A. *et al.* Molecular electrometer and binding of cations to phospholipid bilayers. *Phys. Chem. Chem. Phys.* **18**, 32560–32569 (2016).
22. Poger, D., Van Gunsteren, W. F. & Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **31**, 1117–1125 (2010).
23. Chiu, S.-W., Pandit, S. A., Scott, H. L. & Jakobsson, E. An improved united atom force field for simulation of mixed lipid bilayers. *J. Phys. Chem. B* **113**, 2748–2763 (2009).
24. Dickson, C. J., Rosso, L., Betz, R. M., Walker, R. C. & Gould, I. R. GAFFlipid: a general amber force field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter* **8**, 9617–9627 (2012).

25. Roos, K. *et al.* Opls3e: Extending force field coverage for drug-like small molecules. *J. Chem. Theory Comput.* **15**, 1863–1874 (2019).
26. Ulmschneider, J. P. & Ulmschneider, M. B. United atom lipid parameters for combination with the optimized potentials for liquid simulations all-atom force field. *J. Chem. Theory Comput.* **5**, 1803–1813 (2009).
27. Ferreira, T. M. *et al.* Cholesterol and POPC segmental order parameters in lipid membranes: solid state ^1H - ^{13}C NMR and MD simulation studies. *Phys. Chem. Chem. Phys.* **15**, 1976–1989 (2013).