Benchmarking the dynamics of phosphatidylcholine molecular dynamics force fields using open data

Hanne S. Antila,[†] Tiago Ferreira,[‡] Samuli Ollila,[¶] and Markus Miettinen[*,†]

†*Department of Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany*
‡*Tiago's affiliation here*
¶*Samuli's affiliation here*

E-mail: markus.miettinen@mpikg.mpg.de

Abstract

Brilliant abstract here

1 Introduction

Phospholipids are an important class of biomolecules not only as the building blocks of biomembranes but also as emerging candidates for micro- and nanotechnology, such as the use of liposomes as microcapsules in targeted drug delivery.[1] These molecules are composed of a hydrophilic phosphate head group, which is connected to two hydrophopic fatty acid tails via a glycerol backbone. The ability of lipids to self assemble into bilayer membrane (and other) configurations is a direct consequence of this dual nature. The conformations adopted by the molecules play critical in role in determining properties of these lipid assemblies, such as the dipole potential[2] and the curvature[?]

Although biological membranes are complex mixtures of multiple different lipids as well as other molecules, lamellar phospholipid bilayers with one or few lipid types have been successfully used as simplified models systems to decipher, *eg.*, possible molecular mechanisms behind anesthetics,[3?] the effect of cholestherol on membrane structure,[4?] and the functioning of membrane proteins[5] **1.Markus, do you want to add your favourite topics/references, espc. experimental.** The ability of these simple model membranes to reproduce the biological functioning of headgroup/glycerol structure in cells is backed up by experimental evidence.[6] In particular, classical molecular dynamics (MD) simulations of these model systems have been widely utilized[3,4,7? –10] to provide an atomistic view on the biomembranes, and hold vast potential in making further connections between the structure and the function.

Unfortunately, recent comparison of lipid order parameters from bilayer simulations to nuclear magnetic resonance (NMR) measurements have showed that none of the currently available lipid MD models (force fields) perfectly reproduces the conformational ensemble sampled by the headgroup.[8] Here, we take this comparison a step further and study the dynamics of those MD models. Our motivation is two-fold. Firstly, when investigating static properties of the bilayers, it is crucial to assess how well the simulations have converged. In order to extract reliable statistics, the conformations sampled has to represent the equilibrium distribution with enough transitions between states. Indeed, simulations of a single DOPC lipid using the CHARMM32b2 force field indicated that the conformations sampled do not replicate the equilibrium distribution even after 500 ns[11] and

bond dynamics of the Berger model was shown[9] to be too slow at the glycerol region of the PC headgroup compared to correlation times extracted from NMR experiments. Secondly, for complete picture of membrane functioning, understanding of the bilayer dynamics in addition to equilibrium measurements are needed. The ability of the MD model to reproduce the relative abundance of different dynamical processes is crucial for the correct interpretation of pathways leading to, eg., membrane deformation[?] and lipid-induced conformational[? ?] changes of membrane proteins.

Both the conformational dynamics of experimental lipid bilayers,[12–16] and the related lipid MD models,[12,14,15,17] have traditionally been assessed based on the spin-lattice relaxation rates R_1 (or the corresponding T_1 times), available through NMR measurements. At best, simulation and NMR experiments can be combined to provide an intepretation of the molecular motion.[15,17,18] However, relying on R_1 only has several drawbacks **2.do all flavors (31P-H,13C-H) have the same problem**. It relies on an underlying rotation-diffusion model, its sensitivity is typically limited to C–H bond reorientation with time scales ∼1-10ns, and measurements at several temperatures and magnetic field strengths are required to fully characterize the dynamics. To address these deficiencies, two of us recently introduced[9] the effective C–H correlation time τ_e—a model free quantity that encompasses lipid relaxation processes with time scales up to hundreds of nanoseconds. Most importantly increasing τ_e always signals some type of slowdown in the C–H bond dynamics, making the interpretation less ambiguous than for R_1 where slowdown in the dynamics can lead to either an increase or a decrease of R_1 value.[9]

Here, we utilize the effective correlation time to present the first comprehensive comparison of dynamics of 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC) MD models. We not only investigate the pure bilayers in room temperature but also look at the effect of cholesterol content, hydration and monovalent salt to see whether the model dynamics correctly responds to the change in conditions. An MD model fulfilling these requirements can provide a reliable tool to study, *eg*, membrane remodeling, or to interpret the experimental spin relaxation rates by connecting them to the underlying dynamical processes.

Within this study, we intentionally restrict ourselves to re-use existing, publicly available simulation trajectories and experimental results**3.public availability debatable?**. This is to demonstrate the power of open, well documented data in creating new knowledge at a lowered cost. The project was conducted as a open collaboration under the NMRlipids (`nmrlipids.blogspot.fi`) open science project and the main source of data was the collection of lipid bilayer simulations originating from the previous NMRlipids projects.[8,19]

2 Methods

3 Theoretical background

The NMR experiments investigating the lipid conformational dynamics take advantage of the fact that the relaxation of ^{13}C magnetization dominantly happens via the dipolar coupling of ^{13}C with the magnetic moments of the protons bound to it, with the symmetry axis of the interaction aligning with the C–H bond. The spectral density depicting the ^{13}C relaxation rates is expressed as

$$\mathrm{j}(\omega) = 2\int_0^\infty \cos(\omega\tau)\mathrm{g}(\tau)d\tau, \qquad (1)$$

4.are omega and tau obvious or should they be defined? which is the Fourier transformation of the C-H bond second order autocorrelation function

$$\mathrm{g}(\tau) = \langle P_2\left[\vec{\mu}(t)\cdot\vec{\mu}(t+\tau)\right]\rangle, \qquad (2)$$

where $\vec{\mu}(t)$ is the unit vector in the direction of the C–H bond at time t and P_2 is the second order Legendre polynomial. The angular brackets depict averaging over time. The autocorrelation function can be expressed as the product of two functions

$$\mathrm{g}(\tau) = \mathrm{g_f}(\tau)\mathrm{g_s}(\tau), \qquad (3)$$

where $g_f(\tau)$ characterizes fast decays owing to, for example, the rotations and $g_s(\tau)$ describes slow decays that originate from, *eg.*, the lipid diffusion. The two components, along with the oscillation due to magic angle scattering at the kHz region are depicted in Fig. 1. Correlation time of 4.2 ms has been estimated for multilamellar POPC samples at 300K for the slow modes, whereas in liquid crystalline lipid bilayers the faster $g_f(\tau)$ decays to a plateau value S^2_{CH} within few hundred nanoseconds.[9] The order parameters

$$S_{CH} = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle, \tag{4}$$

where θ is defined as the angle between the C–H bond and the bilayer normal, are measured in NMR experiments from this plateau. As the order parameter describes the conformational ensemble of the molecule, the fast-decaying component of the rotational correlation function intuitively depicts the time needed to sample these conformations. The characteristic time can be quantified via the effective correlation time

$$\tau_e = \int_0^\infty \frac{g_f(\tau) - S^2_{CH}}{1 - S_{CH}} d\tau \tag{5}$$

which is defined via the area of the normalized correlation function ($g_f'(\tau) = g_f(\tau) - S^2_{CH}/1 - S_{CH}$) and graphically depicted in Fig. 1b. It is easily seen that in the presence of more long-lived correlations, τ_e grows, signaling that more time is needed for full conformational sampling.

The bond correlation functions are easily accessible from MD simulations. In addition to extracting τ_e directly from the area, fitting a set of N exponentials, each with their own correlation times τ_i to the normalized correlation function

$$g'_{MD}(t) = \sum_{i=1}^{N} \alpha_i e^{-t/\tau_i} \tag{6}$$

provides an alternative way of quantifying τ_e

$$\tau_e = \sum_{i=1}^{N} \alpha_i \tau_i. \tag{7}$$

When the simulation trajectory is not long enough for the correlation function to reach the plateau, calculating the area gives a lower bound estimate for τ_e while the latter method includes also an estimate for the longer-time components via the fitting process. However, in practice the fit is often highly unreliable in terms of depicting the long tails of the correlation function, and this work we choose to calculate the effective correlation times using the area. **6.Figure here?**

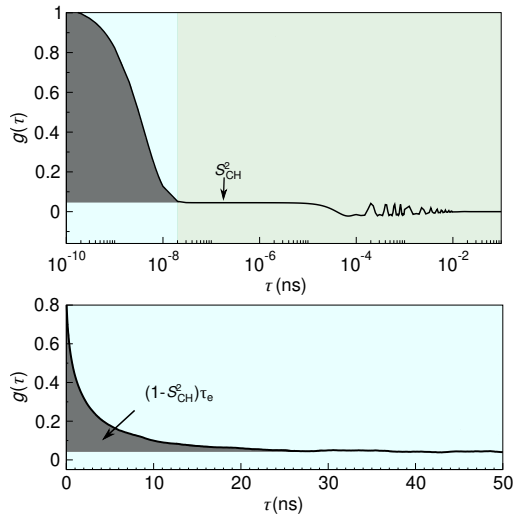

Figure 1: The autocorrelation function g_τ a) The fast mode (shaded blue) and the slow mode (shaded green) along with the oscillation owing to magic angle spinning. The fast mode decays to plateau quantifying the S_{CH} while the slow mode gives the final decay to zero b) Illustration of typical autocorrelation function obtained from a MD simulation. The gray area under the curve gives means of quantifying the τ_e.

5.re-draw image

The spin-lattice relaxation rate R_1 defines the time-scale on which ^{13}C longitudal magnetization equilibrates. It is defined as

$$R_1 = \frac{d_{CH} N_H}{20}\left[j(\omega_H - \omega_C) + 3j(\omega_C) + 6j(\omega_H + \omega_C)\right], \tag{8}$$

where N_H is the number of bound hydrogens, ω_H and ω_C are the Larmor frequencies for ^1H and ^{13}C, and d_{CH} is the dipolar coupling constant. For the methylene bond, $d_{CH}/2\pi$ approximately equals to -22kHz.

7.following needed only if we add R1 time comparison

The dependence of R_1 on the spectral density values at the Larmor frequencies means that the R_1 is reactive to the relative amounts of relaxation processes with time-scales near the inverses of these frequencies. Since the Larmor frequencies depend on the field strength used in the NMR measurements, this typically makes R_1 sensitive to 1-20 ns time-scales. Importantly, a change in R_1 indicates a difference in the relative amount if processes within the detection window, and therefore does not give information on the modulation of the total sampling rate.

3.1 Data aquisition

The simulation trajectories being analysed where collected from the Zenodo repository `zenodo.org` **8.is this how to cite Zenodo?**. A List of the simulations, as well as the references to the data files are presented in Table 1**9.make table** and additional computational details of each of the simulations are available at the referred Zenodo entry. All the experimental quantities were collected from the literature **10.except are they, or from Tiago?** sources referred at the respective figures.

The project was conducted as open collaboration using the `nmrlipids.blogspot.fi` blog as a communication platform. An open invitation to contributions was presented in the blog, and every contributor was offered a co-authorship. The input from each author is detailed in XXX **11.acknowledgements? SI?**. **12.Currently all simulation data was extracted from the zenodo repository, including Berger data (not from the 2015 paper)**

Table 1: List of simulated systems used in this work, along with the reference to the data

4 Results

Figure 2 presents a comparison of experimental POPC effective correlation times, in room temperature and under full hydration, to those obtained from five different molecular dynamics force fields. In general, MD models exhibit slower dynamics than what is observed experimentally in the headgroup region. The discrepancy is most drastic in the glycerol region of the molecule, while the dynamics of the tail C–H bonds are well reproduced. This is consistent with previous comparison performed with the Berger model,[9] as well as the insufficient conformational sampling of glycerol backbone torsions observed[11] in 500ns-long simulations of DOPC lipid utilizing the CHARMM32b2**??** force field. .The best overall performance is obtained from CHARMM36 and Slipids force fields. Note, however, that the values calculated from the simulations give the lower limit of τ_e, as we opted to quantify these from the area (see section 3), and the overestimation of the effective correlation time for some models might be more drastic than what is shown here.

In Figure 3 we present a comparison of experimental R_1 rates to those obtained from the simulations, under the same conditions as was done in Fig. 2 for τ_e. In contrast to the τ_e values, the force fields overall reproduce the experimental R_1 data well the glycerol and headgroup regions **14.what is up with the gamma carbons?** with slight tendency for faster dynamics in the headgroup, and slower motions at the glycerol region, compared to the experimental value. A good overall performance is obtained from the Slipids model. Based on the ability of force fields to replicate the R_1 rates, and the discrepancies observed for τ_e data, one can conclude that the force fields tend to reproduce the dynamics at 1ns scales for the glycerol and headgroup region but differences arise at longer processes contained in the τ_e values.

In the tail region the MD models dominantly produce slower R_1 rates, with the exception of lipid14 force field, which has the best agreement with the experimental data. We note that for the unresolved contributions from both tails, the experimental data should fall somewhere between the two simulated data points. While the good fit obtained for effective correlation times (Fig. 2) indicates that the force fields reproduce the general conformational dynamics well in the tail region, the differences detected

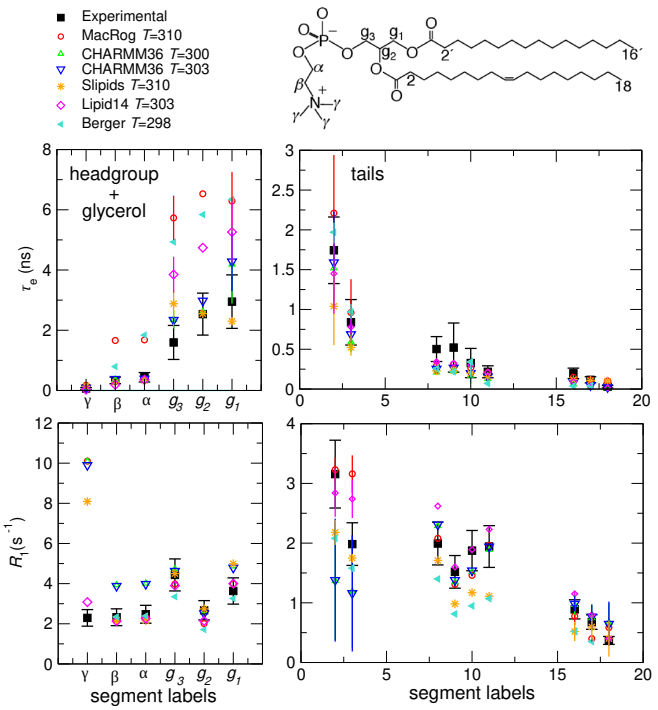

Figure 2: Top: The structure of POPC and the carbon nomelacture used in this work. Bottom: comparison of simulated and experimental correlation times. Further details of the simulation data is provided in Table 1. The experimental values are extracted from Ref. 9 and measured from fully hydrated POPC bilayers in L_α phase at 298 K. The experimental data for segments 2-3 and 16-18 represent the non-resolved contributions from both acyl chains, whereas labels 8-11 refer to the sn-2 (oleoyl) chain. The error bars for experimental values reflect error estimate of ±18% (see Ref. 9), whereas the error bars for simulated data points give the minimum and maximun value observed at each carbon while the symbol denotes the average.

13.but there is no 18 carbon in palmitoyl?

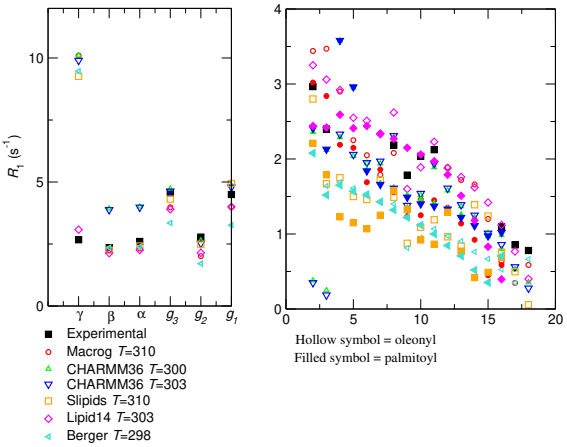

Figure 3: Experimental and simulated R_1 values for POPC. The experimental data is extracted from Ref.[9] and measurement conditions and the carbon labelling scheme is as in Fig. 2.

for R_1 rates signal that processes at the ∼1ns time scale are overrepresented**15.think if valid**.

A comparison to experimental dynamics under standard conditions is not sufficient to fully assess the quality of the dynamics obtained from an MD model: a more complete picture is acquired by evaluating whether the model responds correctly to the change in conditions. We therefore proceed to investigate how the dynamics change when cholesterol is added to the bilayer (Fig. 4). Experimentally, additional cholesterol is know to cause slower decorrelations of the hydrogens attached to POPC g_1, g_2, and g_3 carbons. All the force fields investigated in Fig. 4 qualitatively reproduce this. The increase upon addition of cholesterol is overestimated with the MacRog and CHARMM36 models while slight underestimation occurs in simulations using Slipids and Berger force fields, especially at g_1 and g_2 carbons. The change observed here, however, is particularly sensitive to the length of the trajectory as cholesterol-induced increase in effective correlation time is likely to lead to worse convergence of the correlation function within the limited simulation time, and more drastic underestimation of τ_e is expected than for simulations without cholesterol. This will, consequently, cause a tendency towards underestimation on the strength of the cholesterol-driven

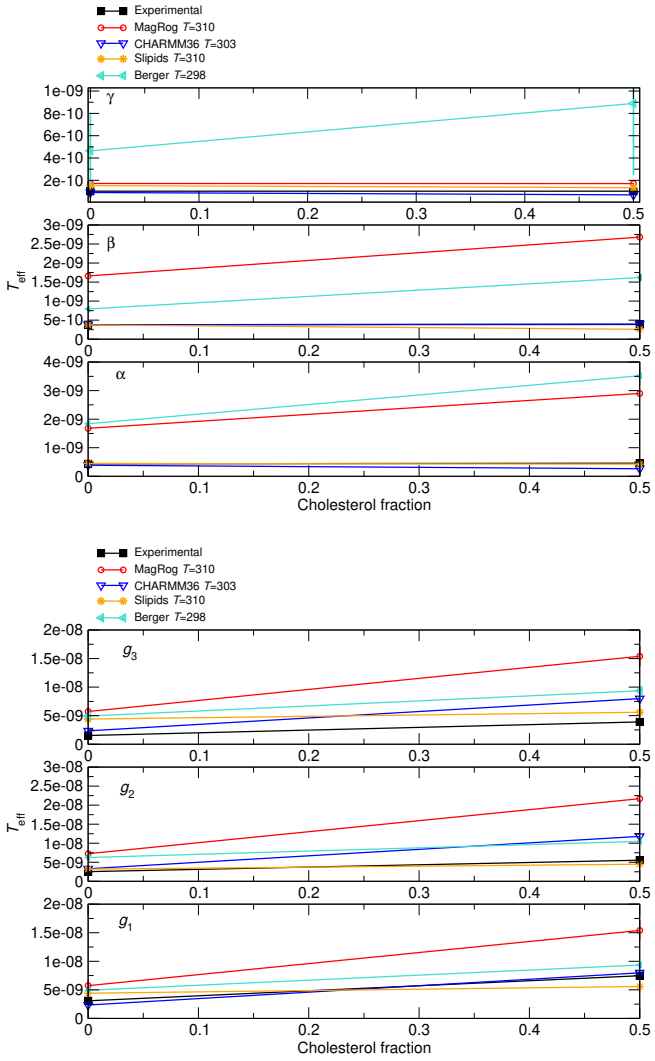

Figure 4: The effect of bilayer cholesterol content. The experimental values are from XXX and measured at XXXK

modulation of the effective correlation time.

For the α , β, and γ no change in τ_e is detected experimentally. Nevertheless, Berger and MacRog models exhibit a slow-down in dynamics also for these C–H bonds, while the τ_e in other models remains constant.

To investigate the effect of hydration on the C–H bond dynamics on the PC headgroup, we first present a comparison of experimental effective correlation times obtained from the POPC (measured in full hydration) and DMPC (1,2-Dimyristoyl-sn-glycero-3-phosphocholine, measured in low hydration) in Fig. 19. The values are the same within the experimental accuracy, which leads to two conclusions, 1) the motions of the headgroup bonds are unaffected by the differences in the tails between the two molecules and 2) the decrease of hydration level down to 13 waters per lipid does not considerably alter the correlation times for the PC headgroup. That said, the values in low hydration for the glycerol carbons, while still within the experimental accuracy, are consistently above those obtained in full hydration, which might imply detectable slowdown when hydration is further reduced.

As the available trajectories in low hydration are relatively short, we only investigate the effective correlation times γ and β carbons where the fast dynamics alleviates the risk of underestimation. The results presented in Fig. 19 show that all the force fields produce effective correlation times that are relatively unaffected by the hydration level above 15 waters per lipid (W/L), in line with the experimental observation. When the hydration is further reduced, the dynamics slows down. The slowdown is particularly clear with the MacRog and Berger**16.check new berger data** models. At these same levels of hydration, an increase in the lipid order parameters have been observed,[8] which together with the slow down may indicate the onset of a phase transition**17.is this a valid speculation?**. **18.Berger changes most but hard to interpret bc. only 2 data points and large error bars**.

Finally, we will investigate the response of the effective correlation times to increasing amounts of monovalent salt. Experimentally, the modulation of α and β carbon order pa-

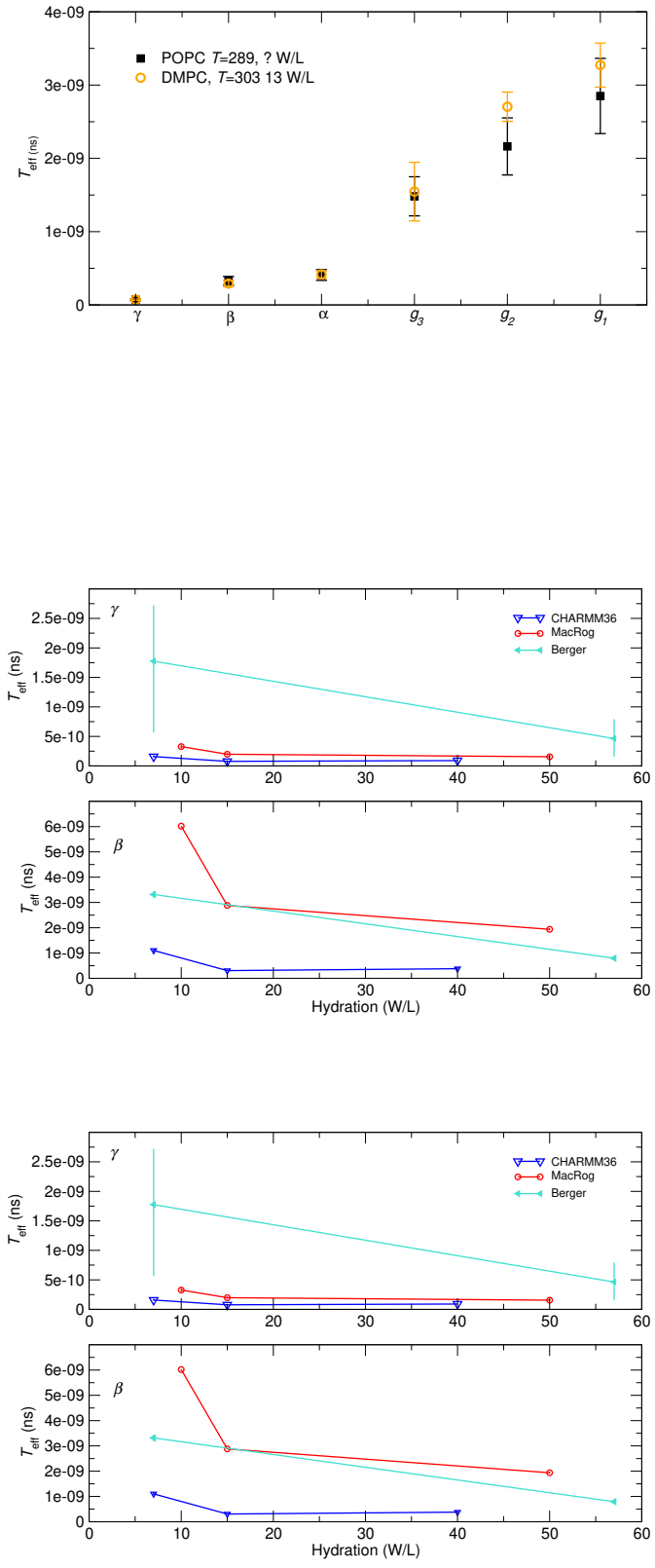

Figure 5: The effect of hydration on the effective correlation times. Top: comparison of experimental effective correlation times from DMPC in low hydration and POPC in full hydration. Neither the different chemistry of the lipid tails, nor the hydration level has no effect within the experimental accuracy. Bottom: The response of effective correlation times of γ

rameters increasing ion concentration have been used to quantify the ion binding to the lipid bilayers (the molecular electrometer concept [19?]). As expected from the uncharged nature, these order parameters are constant for POPC bilayers under salt addition in experiments, indicating negligible ion binding. Based on this, we anticipate the correlation times also to be unaffected by salt but to our knowledge no experimental measurements have been conducted to quantify this.

The molecular electrometer concept has been used to show that most molecular dynamics models overestimate the binding of monovalent ions to the PC headgroup:[19] the modulation the α and β carbon order parameters by increasing NaCl concentration was overestimated compared to the experiments, and accompanied by accumulation of ions at the bilayer surface, in the simulations. In Fig. 18 we compare two force fields, one that is known to exhibit overbinding[19] (MacRog) and one producing more realistic binding affinity (CHARMM36). Indeed, the effective correlation times extracted from the CHARMM36 model vary only a little when ion concentration is increased, whereas the MacRog model exhibits a clear slow-down. This indicates that similarly to the order parameters, the effective correlation times may be useful in investigating the ion binding affinity of lipid bilayers.

5 Conclusions

- The force fields have a tendency to towards slow dynamics
- Power of well documented, open data: one can draw conclusions using existing data, no need to always do new simulations.

Acknowledgement

This material is based upon work supported by XXX under Grant No. XXX. The project is/isn't part of the NMRlipids open collaboration (nmrlipids.blogspot.com)

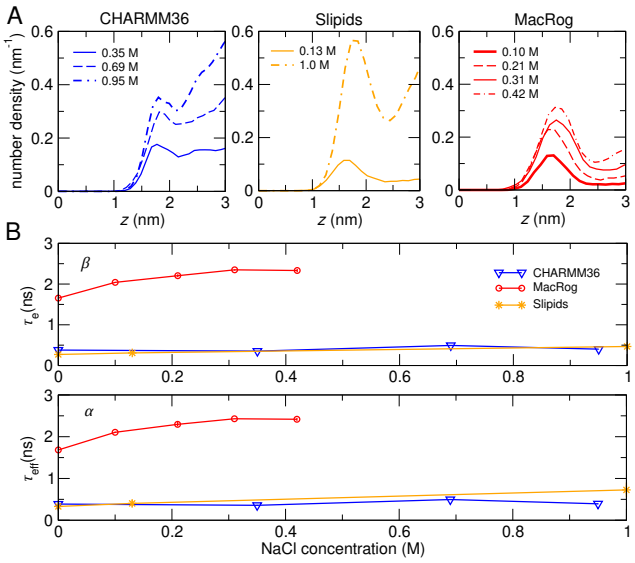

Figure 6: The effective correlation times of α and β C–H bonds under increasing concentration of NaCl from CHARMM36 and MacRog POPC simulations.

References

(1) Sercombe, L.; Veerati, T.; Moheimani, F.; Wu, S. Y.; Sood, A. K.; Hua, S. Advances and Challenges of Liposome Assisted Drug Delivery. *Frontiers in Pharmacology* **2015**, *6*, 286.

(2) Gawrisch, K.; Ruston, D.; Zimmerberg, J.; Parsegian, V.; Rand, R.; Fuller, N. Membrane dipole potentials, hydration forces, and the ordering of water at membrane surfaces. *Biophysical Journal* **1992**, *61*, 1213 – 1223.

(3) Chau, P.-L.; Hoang, P. N.; Picaud, S.; Jedlovszky, P. A possible mechanism for pressure reversal of general anaesthetics from molecular simulations. *Chemical Physics Letters* **2007**, *438*, 294 – 297.

(4) 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 1H13C NMR and MD simulation studies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1976–1989.

(5) Lindahl, E.; Sansom, M. S. Membrane proteins: molecular dynamics simulations. *Current Opinion in Structural Biology* **2008**, *18*, 425 – 431, Membranes / Engineering and design.

(6) 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 headgroup region by deuterium magnetic resonance. *Biochemistry* **1981**, *20*, 1826–1831.

(7) Lyubartsev, A. P.; Rabinovich, A. L. Recent development in computer simulations of lipid bilayers. *Soft Matter* **2011**, *7*, 25–39.

(8) Botan, A.; Favela-Rosales, F.; Fuchs, P. F. J.; Javanainen, M.; Kandu, 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. *The Journal of Physical Chemistry B* **2015**, *119*, 15075–15088, PMID: 26509669.

(9) Ferreira, T. M.; Ollila, O. H. S.; Pigliapochi, R.; Dabkowska, A. P.; Topgaard, D. Model-free estimation of the effective correlation time for CH bond reorientation in amphiphilic bilayers: 1H13C solid-state NMR and MD simulations. *The Journal of Chemical Physics* **2015**, *142*, 044905.

(10) Miettinen, M. S.; Lipowsky, R. Bilayer membranes with frequent flip-flops have tensionless leaflets. *Nano letters* **2019**, *?*, ?–?

(11) Vogel, A.; Feller, S. E. Headgroup Conformations of Phospholipids from Molecular Dynamics Simulation: Sampling Challenges and Comparison to Experiment. *The Journal of Membrane Biology* **2012**, *245*, 23–28.

(12) Feller, S. E.; Gawrisch, K.; MacKerell, A. D. Polyunsaturated Fatty Acids

in Lipid Bilayers: Intrinsic and Environmental Contributions to Their Unique Physical Properties. *Journal of the American Chemical Society* **2002**, *124*, 318–326, PMID: 11782184.

(13) Eldho, N. V.; Feller, S. E.; Tristram-Nagle, S.; Polozov, I. V.; Gawrisch, K. Polyunsaturated Docosahexaenoic vs Docosapentaenoic AcidDifferences in Lipid Matrix Properties from the Loss of One Double Bond. *Journal of the American Chemical Society* **2003**, *125*, 6409–6421, PMID: 12785780.

(14) Wohlert, J.; Edholm, O. Dynamics in atomistic simulations of phospholipid membranes: Nuclear magnetic resonance relaxation rates and lateral diffusion. *The Journal of Chemical Physics* **2006**, *125*, 204703.

(15) Klauda, J. B.; Roberts, M. F.; Redfield, A. G.; Brooks, B. R.; Pastor, R. W. Rotation of Lipids in Membranes: Molecular Dynamics Simulation, 31P Spin-Lattice Relaxation, and Rigid-Body Dynamics. *Biophysical Journal* **2008**, *94*, 3074 – 3083.

(16) Leftin, A.; Brown, M. F. An NMR database for simulations of membrane dynamics. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2011**, *1808*, 818 – 839, Including the Special Section: Protein translocation across or insertion into membranes.

(17) Klauda, J. B.; Eldho, N. V.; Gawrisch, K.; Brooks, B. R.; Pastor, R. W. Collective and Noncollective Models of NMR Relaxation in Lipid Vesicles and Multilayers. *The Journal of Physical Chemistry B* **2008**, *112*, 5924–5929, PMID: 18179193.

(18) Pastor, R. W.; Venable, R. M.; Karplus, M.; Szabo, A. A simulation based model of NMR T1 relaxation in lipid bilayer vesicles. *The Journal of Chemical Physics* **1988**, *89*, 1128–1140.

(19) Catte, A.; Girych, M.; Javanainen, M.; Loison, C.; Melcr, J.; Miettinen, M. S.; Monticelli, L.; Mtt, J.; Oganesyan, 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.

Graphical TOC Entry

TOC here if needed