Johannes Kremser, BSc

Influence of Phosphatidylcholine Chain Length and Composition on Intrinsic Lipid Curvatures

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Technical Physics

submitted to

Graz University of Technology

Supervisor

Georg Pabst, Assoc.-Prof. Dr.

Institute of Materials Physics

Graz, October 2020

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

29.10.2020

Date, Signature

Abstract

The intrinsic lipid curvature C_0 is a property of lipids which describes the bending radius of unstressed monolayers, resulting from the molecular shape of the lipids. It is an important parameter in the calculation of interfacial energies, which affect lipid/protein interactions. We explore a model-based approach to determine C_0 by applying a full q-range small-angle X-ray scattering data analysis of inverted hexagonal phases based on Bayesian statistics. The technique was tested on different mixtures of phosphatidylethanolamines with lamellar phase forming phosphocholines and compared to an all-atom molecular dynamics simulation. We included a non-linearity term for lipid mixtures based on the effective lipid headgroup size. We found that saturated lipids show a linear dependence of curvature and volume on chain length and that lipids with unsaturated chains deviate from this linear trend. Also, challenges and limitations of our approach are discussed. In particular, we observed that the consistency of the sample preparation is crucial for a successful determination of the curvature.

Acknowledgments

Contents

1. Introduction

The present work is dedicated to determine the intrinsic curvature of membrane lipids by means of a global analysis of data provided by X-ray experiments. It is based on previous work [7, 8] from Moritz Frewein and tries to improve the methodology introduced there. The main goal is to find an applicable method which is capable of evaluating the intrinsic curvatures for bilayer forming lipids.

1.1. Lipids and Intrinsic Curvature

Lipids, as membrane forming molecules, are a crucial part of all biological cells and therefore very important for every living organism. The intrinsic curvature, one of the molecular properties of lipids, is an important quantity and has impact on the partitioning of transmembrane proteins and various other lipid-protein and lipid-peptide interactions [16, 9, 3].

The concept of curvature is best understood when looking at a bent monolayer. From every point on the surface, a normal can be drawn and the principal curvatures, c_1 and c_2, can be defined, as shown in figure 1, as the maximum and minimum values of curvature respectively [26].
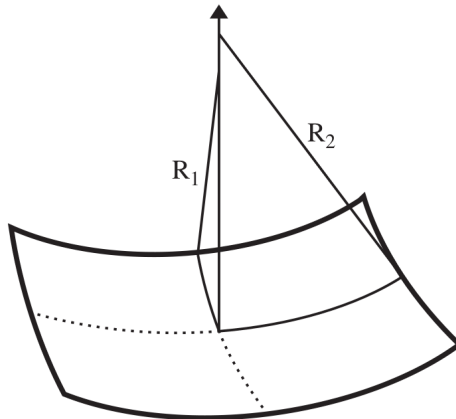
Figure 1: Principal curvatures at one point of a monolayer. At these points they are defined as $c_1 = 1/R1$ and $c_2 = 1/R_2$. Source: [26]

The two principal curvatures are perpendicular to each other and combined give the mean curvature C or the Gaussian curvature K at that point [26].

$$C = \frac{1}{2}(c_1 + c_2)$$
$$K = c_1 c_2 \tag{1}$$

If both, the mean and the Gaussian, curvatures are known at every point of the surface, the shape of the membrane is determined. It is important to point out that going from a flat

1

surface to a cylinder, the mean curvature will change but the Gaussian curvature remains 0. The Gaussian curvature cannot be altered by bending alone. Some distortion, such as stretching or compression, must occur, thus a flat surface cannot be wrapped into a sphere without distorting it [26].

It is considered that the free energy of the system, consisting of the lipid and water molecules, determines the favored shape of the membrane. The contributors to this free energy g_{tot} are the membrane curvature elasticity g_C, the packing of the hydrocarbon chains g_P and a free energy interaction term g_{inter} which accounts for any other interaction such as hydration and electrostatic forces [26].

$$g_{tot} = g_C + g_P + g_{inter} \tag{2}$$

The membrane curvature elasticity is well described by the 'Helfrich Ansatz' which simplifies the problem by reducing the membrane to an infinitely thin elastic surface. The energy cost for deformations is then only dependent on changes of the curvature of the surface, defined by its mean and Gaussian curvatures, C respectively K. This gives the curvature elastic energy per unit area [26, 12]:

$$g_C = 2\kappa(C - C_0)^2 + \kappa_G K \tag{3}$$

With the bending modulus κ, describing the energy needed to bend the surface, and the Gaussian modulus, accounting for the change of the Gaussian curvature. The mean curvature of the surface when totally relaxed C_0, later referred as the intrinsic curvature, is the quantity of the surface this work is determined to. For inverse structures such as the inverse hexagonal phase H_{II}, the intrinsic curvature is defined as:

$$C_0 = \frac{-1}{R_0} \tag{4}$$

With R_0 describing the radius of the curvature at the neutral plane, which occurs where bending and stretching fluctuations are decoupled [17]. The sign is a convention, being negative means that the monolayer is curved away from the hydrocarbon chains, see figure 2.
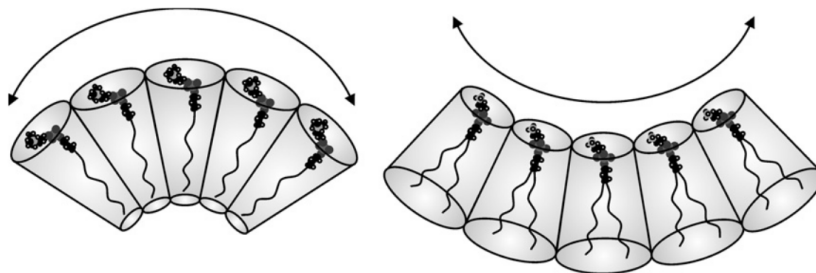
Figure 2: Convention adopted for the sign of the curvature definition. Left: positive; Right: negative (inverse); Source: [26]

1. Introduction

The formally mentioned inverse hexagonal phase H_{II} is one of several structures lipids can assemble when dissolved in water. Which phase a lipid will prefer during this self-assembly process is largely defined by the packing parameter S which depends on the molecular volume of the lipid chains V, the optimum surface area of the headgroup A_0 and the maximum length of the chains l [13, 26], see equation (5).

When in contact with water, the hydrophobic effect keeps the oily chains of the lipid molecules screened from the water as much as possible. This leads to a variety of supramolecular aggregates that are formed spontaneously via self-assembly [22]. As mentioned above, the main factor determining which aggregate is formed is the packing parameter. In figure 3 an overview of the possible aggregates dependent on the packing parameter is given.

$$S = \frac{\nu}{A_0 \, l} \tag{5}$$
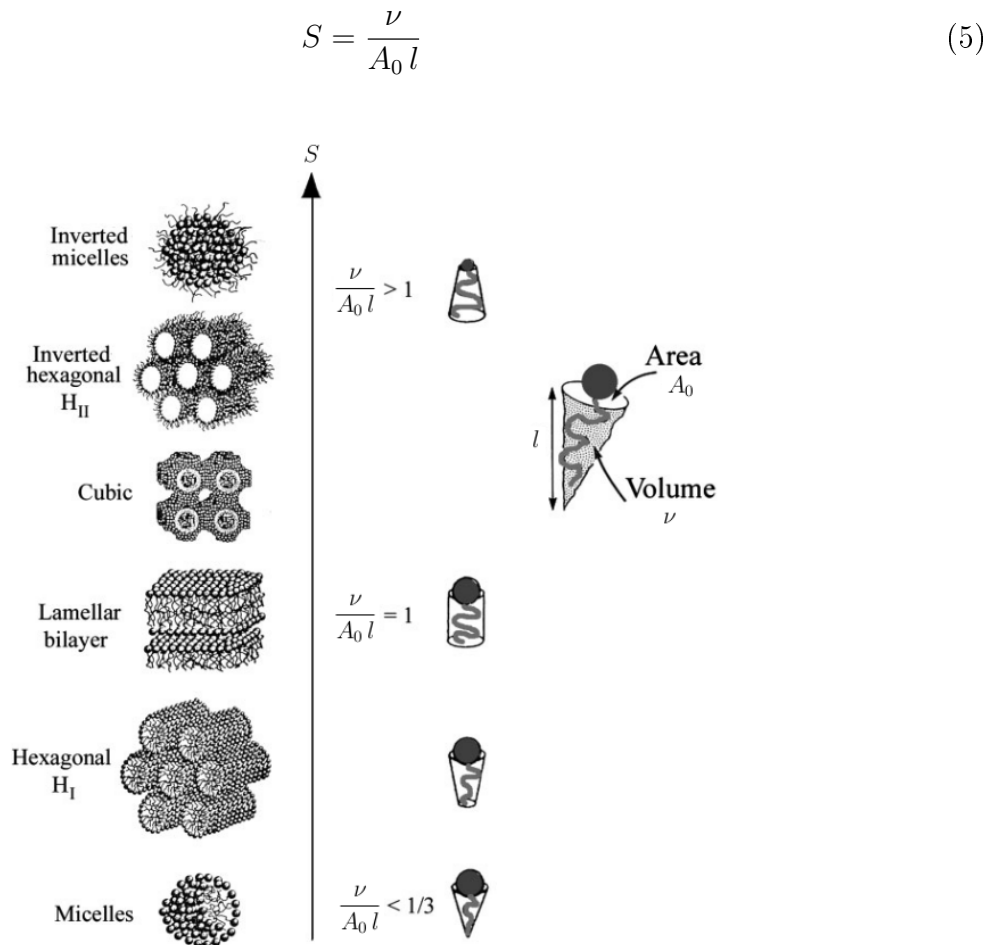
Figure 3: Possible Lipid phases dependent on the packing parameter. Source: [23]

As we have seen, lipids are very diverse and are often classified into non-bilayer and bilayer forming lipids [1]. Lipid bilayers are never free of stress which makes measuring a parameter, e.g. the intrinsic curvature, that depends on the shape of the molecule itself difficult. Luckily the monolayers in the H_{II} phase can provide a stress-free environment. The non-bilayer forming lipid DOPE (1,2-dioleoyl-sn-glycero-3 phosphoethanolamine) can assemble this aggregate when dissolved in water, see figure 4. In theory, the molecules in such a configuration

are free of outer stress but due to the hexagonal array, there will be voids between the lipid cylinders which would cause the lipids to stretch out, generating deformations of the tubes. To prevent this stress a filler molecule, tricosene in this work, is added [7].
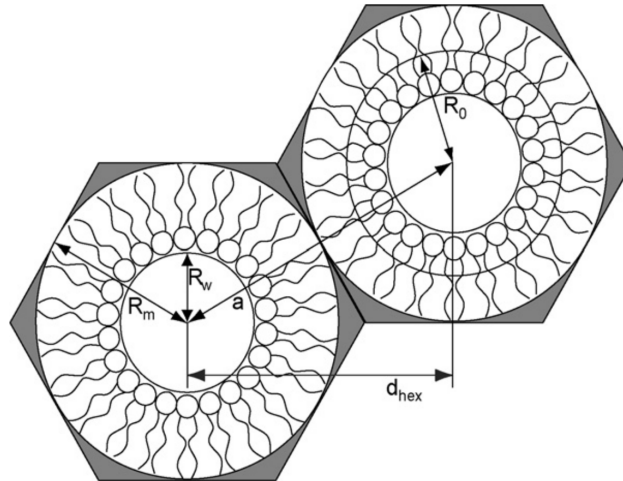
Figure 4: Schematic of the inverse hexagonal phase H_{II}. The H_{II} phase consists of tubes of lipid monolayer around a water core arranged in a hexagonal lattice. Tricosene is used to fill up the voids (shaded gray). R_0 denotes the radius to the neutral plane, defining the intrinsic curvature with $C_0 = \frac{-1}{R_0}$, a describes the lattice constant and R_w stands for the radius of the water core. [1]

1.1.1. Some Lipid types and the filler molecule

DOPE (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine) with its di-monounsaturated chains and rather small head, compared to the chain volume, occurs in the H_{II} phase on contact with water at room temperature. This attribute makes it a great candidate for curvature measurements. The chemical structure of DOPE is visualized in figure 5.
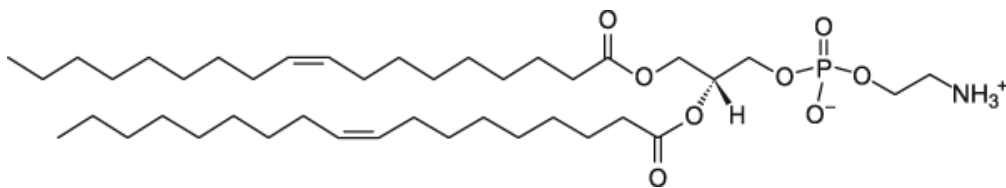
Figure 5: Chemical structure for DOPE. Source: [20]

The method for measuring the intrinsic curvature, which is described in this work, is mainly applied to phosphocholines. One lipid of this family, DPhPC (1,2-diphytanoyl-sn-glycero-3-phosphocholine), is shown in figure 6. The additional methyl groups on every other carbon molecule make this lipid a bit special. With this additional chain volume a more negative curvature is expected compared to other phosphocholines.
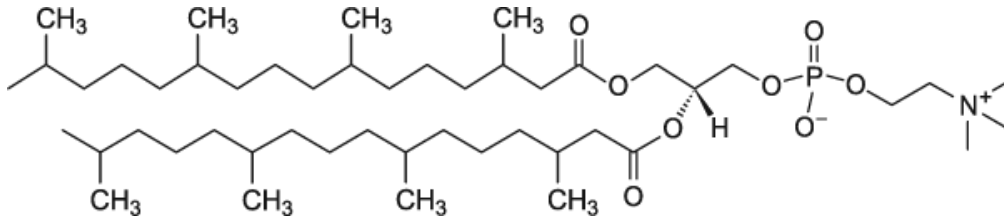
Figure 6: Chemical structure for DPhPC. Source: [20]

An example of lipids with completely saturated chains is DPPC (1,2-dipalmitoyl-sn-glycero-3-phosphocholine) shown in figure 7. Similar lipids, i.e. lipids with a PC head and saturated chains, are expected to have a curvature at around 0 or even slightly positive, depending on their chain length.
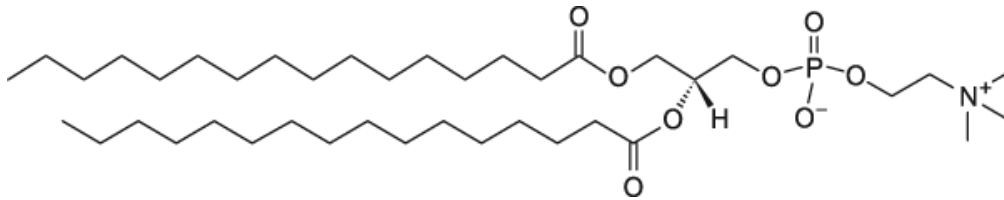
Figure 7: Chemical structure for DPPC. Source: [20]

Previous works showed that cis-9-tricosene inserts primarily into the interstices between the lipid cylinders in the H_{II} phase and so is effectively reducing packing frustration[1, 16]. The master thesis of M. Frewein [8] suggests $12\,w\%$ of added filler molecule which is the amount used in this work. The chemical structure of tricosene is printed in figure 8.
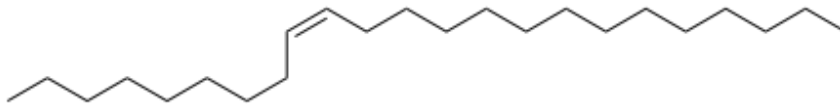
Figure 8: Chemical structure for cis-9-tricosene. Source: [5]

1.2. Lipid mixtures – Headgroup Interactions

To measure a bilayer forming lipid, usually a small amount of it is added as guest lipid and is expected to incorporate into the H_{II} phase, assembled by the host lipid, while distributing evenly in such. By doing so, the guest lipid will alter the cylinder radius according to its curvature. The aim of this work is to exploit this change in curvature to determine the intrinsic curvature of the bilayer forming lipid.

In previous works, a linear addition of the host and guest curvatures was assumed [16].

$$C_0 = (1 - x)C_{0,host} + xC_{0,guest} \,, \tag{6}$$

with x being the molar fraction of the added guest lipid in the mix.

$$x = \frac{N_{guest}}{N_{guest} + N_{host}}$$

The following considerations on non-linear headgroup interaction when mixing the lipids are based on an unpublished work by Primož Ziherl (University of Ljubljana, Slovenia) [30]. Let's examine a mixture of two different lipids. One of them, the one which makes up the main body of the mixture, has to form a H_{II} phase. For this consideration it is important to assume that the lipids mix well, the mixture still forms a H_{II} phase and that the radial dimension of the headgroup is the same for both lipids. When these assumptions are met, each of the two lipids can be characterized by their in-plane dimensions at the neutral plane. These dimensions are denoted as b_{host} and b_{guest} for the host and guest lipid respectively and are sketched in figure 9.
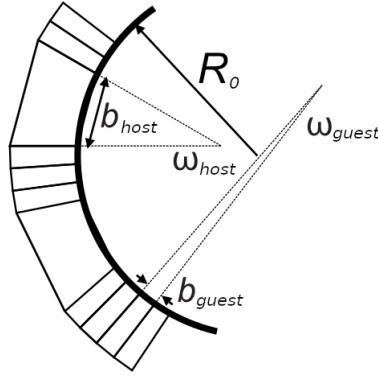
Figure 9: Schematic of the cross section of an inverted cylindrical monolayer micelle consisting of two different lipids. Showing different head widths b_A and b_B. The thick line is representing the neutral plane with radius R_0. Source: [30]

The total angle of the neutral plane contour is 2π and consists of the contribution of N_{host} molecules of the host and N_{guest} molecules of the guest lipid.

$$2\pi = N_{host}\omega_{host} + N_{guest}\omega_{guest} = (N_{host} + N_{guest})[(1-x)\omega_{host} + x\omega_{guest}] \qquad (7)$$

This argument applies as long as at least one lipid has a negative curvature and the mixture still forms an inverted cylindrical micelle.
The circumference of the contour at the neutral plane is

$$l = N_{host}b_{host} + N_{guest}b_{guest} = (N_{host} + N_{guest})[(1-x)b_{host} + xb_{guest}], \qquad (8)$$

with the in-plane dimensions of the lipid headgroups b_{host} and b_{guests}. Replacing $(N_{host} + N_{guest})$ with (7) gives

$$2\pi R_0 = 2\pi \frac{(1-x)b_{host} + xb_{guest}}{(1-x)\omega_{host} + x\omega_{guest}},$$

where R_0 denotes the radius of the neutral plane. With $C_0 = \frac{-1}{R_0}$ and the in-plane intrinsic curvature of the single components $C_{0,host/guest} = \frac{-\omega_{host/guest}}{b_{host/guest}}$ this can finally be written as the total intrinsic curvature of the mixture.

$$C_0(b_{guest}/b_{host}, x) = \frac{(1-x)C_{0,host} + x\frac{b_{guest}}{b_{host}}C_{0,guest}}{(1-x) + x\frac{b_{guest}}{b_{host}}} \tag{9}$$

The result is valid for $C_0(b_{guest}/b_{host}, x) < 0$ where the mixture still forms a H_{II} phase. For $b_{guest}/b_{host} = 1$ this reduces to the linear relationship in (6).

1.3. Problems of the previous work

The main problems of the previous work on defining a global model for evaluating the intrinsic curvature, which are also attacked in this work, are listed below:

- Convergence problems of the MCMC-algorithm:

  The MCMC algorithm described in [8] is rather simple and therefore has sometimes problems converging. This is addressed by switching to a full-blown external fit library [25].

- Mid- to high q-range is not coinciding well with the data for DOPE fits:

  Figure 10 shows some issues of the previous approach in the mid- to high q-range. This will be addressed by adjusting the part of the fit accounting for diffuse scattering.
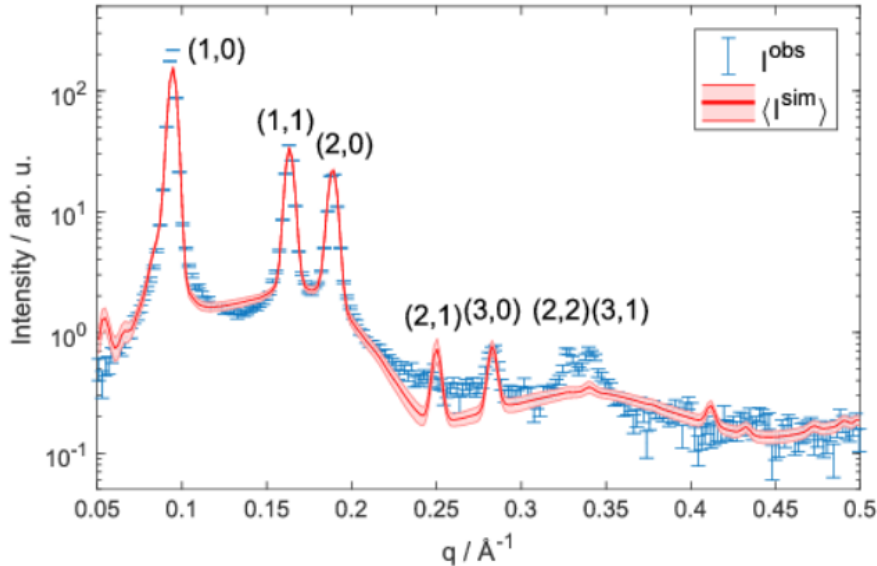
Figure 10: Result for pure DOPE at 35 °C of the original method. Source: [7]

- Odd high q-range behavior of the form factor:

  In figure 11, the shown intensity of the form factor displays an unexpected, additional minima between $0.3\,\text{Å}$ and $0.4\,\text{Å}$. As it is shown later, also simulations do not show this additional minima. This problem will be addressed by also applying the lipid unit cells polydispersity to the tricosene part, which was not done previously.
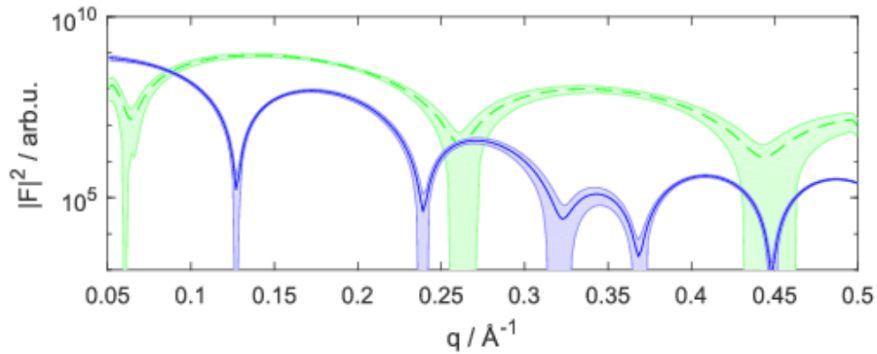
Figure 11: Hexagonal form factor for DOPE at $35\,^\circ\text{C}$. Source: [7]

- Insufficient agreement for results of the same lipid at different guest concentrations:

  Arguably the biggest problem was the bad agreement within results for different guest concentrations of the same lipid. An instability like that raises doubt in which result to trust. In the present work this problem will be addressed with a multitude of small adaptions like the simplification of the mixed head slabs and reparametrization.
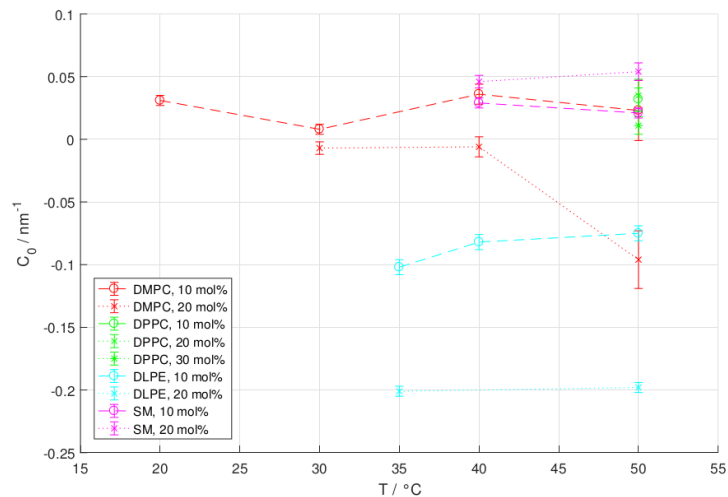
Figure 12: Results for different bilayer forming lipids acquired by the original method. Source:[8]

1.4. Small-angle X-ray scattering

H_{II} phases are very accessible systems for deriving the intrinsic lipid curvatures by small angle X-Ray scattering (SAXS) [7, 16]. As it will be described later in more detail, the

intensity of the scattering signal $I(\boldsymbol{q}) \propto |F(\boldsymbol{q})|^2 |S(\boldsymbol{q})|$ is proportional to the form factor $F(\boldsymbol{q})$ and structure factor $S(\boldsymbol{q})$. The form factor in turn, among other things, depends on the scattering length density (SLD) which is in the case of SAXS the electron density of the specimen. What is called the scattering vector $\boldsymbol{q}$ is proportional to the difference between the primary, the part of the incident beam which gets not scattered, and the scattered beam.

$$\boldsymbol{q} = \frac{2\pi}{\lambda}(\boldsymbol{S} - \boldsymbol{S_0}) \tag{10}$$

The modulus of the scattering vector becomes $q = \frac{4\pi}{\lambda}\sin\theta$. In figure 13 a simple schematic of a beam hitting 2 arbitrary scattering centers is shown.
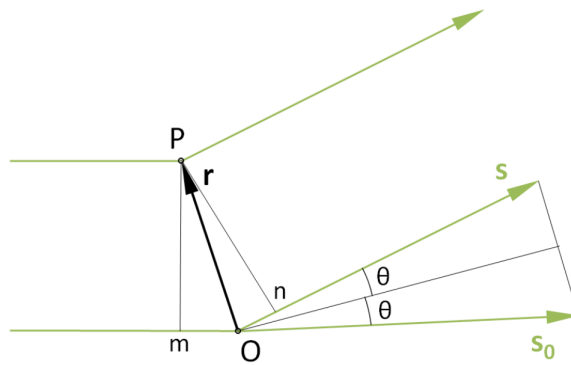
Figure 13: Elastic scattering at the scattering centers P and O. The primary beam and the scattered, secondary beam are denoted as S_0 respectively S. Source: [8]

1.5. Bragg Peaks and linear extrapolation

A non-global approach was developed by Benjamin Kollmitzer and is described in [16]. His method uses only Bragg peaks to calculate the overall intrinsic curvature of the system and ignores the diffuse scattering signal. The results of measurements of different guest lipid amounts are then linearly extrapolated to get the curvature of the lipid of interest.

Some results of the present work will be compared to this linear extrapolation and the differences will be discussed.

1.6. Molecular dynamics simulation for DOPE

The unpublished results of an all-atom molecular dynamics (MD) simulation of a system consisting of 128 DOPE molecules with 40 tricosene chains is provided by Douwe Jan Bonthuis (Graz University of Technology, Austria). This simulation helps to validate the results of DOPE produced by the present thesis.

2. Methods – From Lipid Dust to Curvature

This section deals with all methods used, from sample preparation to actual measurement up to the computational evaluation and interpretation of the data.

2.1. Experimental Techniques

2.1.1. Sample Preparation – RSE

The sample preparation follows the description in [8]. Lipids were purchased at Avanti Polar Lipids (Alabaster, AL) in the form of dry powder and used without any further purification. The total amount of lipid, a single species or a combination of two, used for one sample is 6 mg. The powder is dissolved in organic solvent, a chloroform-methanol mixture (9:1 for phosphoethanolamines and 2:1 for phosphocholines), at a concentration of 10 mg/ml. Different lipids are dissolved in separate vials. The filler molecule is dissolved in a 9:1 chloroform-methanol mixture at a concentration of 5 mg/ml, also in a separate recipient. Some water (0.3 ml, ultra-pure 18 mΩ) is preheated to approximately 65 °C. The dissolved lipid and tricosene samples are added to the water and immediately put into the RSE-apparatus, see figure 14. In there it gets vortexed at 600 rpm, exposed to a vacuum of 400 to 500 mbar and there is a constant flow of argon with 60 ml/min. This process lasts 5 min to make sure that all the solvent is evaporated.
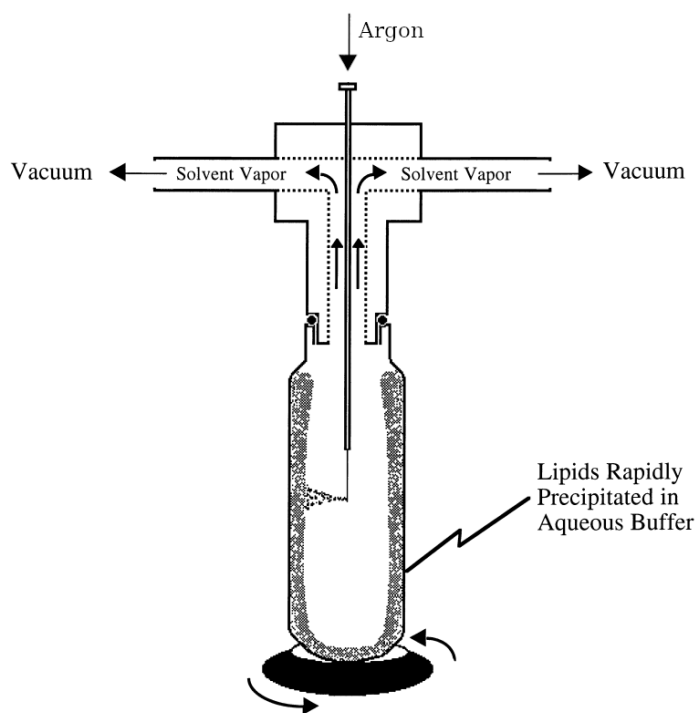
Figure 14: Cross-sectional schematics of the rapid solvent exchange (RSE) process. Source: [4]

2.1.2. SAXS Measurement

The SAXS measurements are performed on a SAXSpace camera from Anton Paar (Graz, Austria) equipped with an Eiger R1M detector system (Dectris, Baden-Daettwil, Switzerland) and a 30 W Genix 3D microfocus X-ray generator from Xenocs (Sassenage, France). The generated radiation originates from the Cu-Kα transition which has a wavelength of 1.54 Å creating a beam with a circular spot size of $\sim$300 μm on the detector. As sample holder paste cell from Anton Paar was used and samples are equilibrated for 10 min at each measured temperature. The total exposure time was 30 min, 6 frames with 5 min each, and the distance between sample and detector was set to 308 mm. Data correction and reduction was performed via the program SAXSanalysis (Anton Paar).
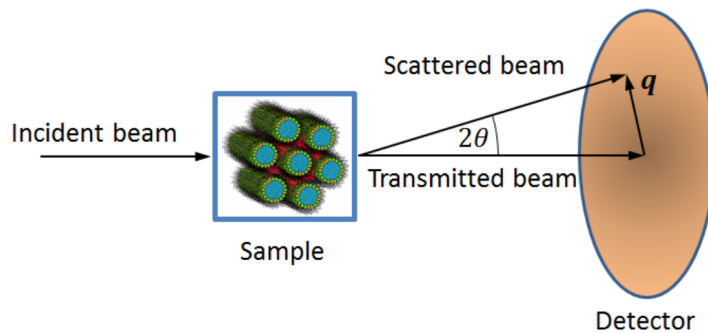
Figure 15: Schematic drawing of the SAXS-measurement. Source: [8]

2.2. Data Evaluation – Where is the Curvature

Now that the data is available, the question remains: Where is the curvature? In this section, the slab model used to describe the hexagonal phase is defined. Furthermore, it is explained how the mixing of lipids is modeled, parameterized and finally optimized via a Monte Carlo method to obtain the intrinsic curvature of lamellar phase forming lipids.

2.2.1. Modeling the Inverse Hexagonal Phase

What is needed here, is a parameterized, mathematical expression describing the hexagonal phase which can be optimized against measured data and in that way gain information about the system like the intrinsic curvature.
To define such a model, the H$_{II}$ phase is considered to be built out of lipid prisms arranged in a hexagonal lattice, see figure 16. A form factor $F(\boldsymbol{q})$ can be defined to describe a single prism, whereas a structure factor $S(\boldsymbol{q})$ is defined to account for the hexagonal lattice with q being the scattering vector [7].
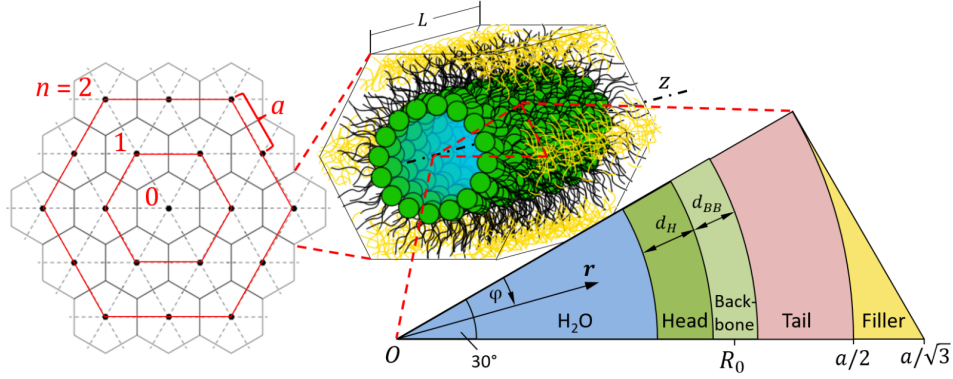
Figure 16: Scheme of the H_{II} phase model.
Source: [7]

The length of these lipid prisms is assumed to be very long compared to their diameter, therefore the Form Factor and Structure Factor are decoupled which gives the following relationship for the scattering intensity $I(\boldsymbol{q})$[6].

$$I(\boldsymbol{q}) \propto |F(\boldsymbol{q})|^2 \, S(\boldsymbol{q}) \tag{11}$$

In the following equations, the independent variables after the vertical bar are fitting parameters. In equation (12) for example this means that n and Δ will be fitted.

Structure Factor
The structure factor for this geometry is given by [6]:

$$S(q, \theta|n, \Delta) = 1 + \frac{1}{N_{hex}(n)} e^{-q^2 \Delta} \sum_{j \neq k}^{N_{hex}(n)} J_0(q|\boldsymbol{R}_j - \boldsymbol{R}_k|sin\theta) \tag{12}$$

Where θ is the angle between the scattering vector $\boldsymbol{q}$ and the long axis of the cylinders. N_{hex} is the total number of lattice points for a hexagonal lattice with n rings. To account for thermal displacements around the mean positions $\boldsymbol{R_j}$ of the lattice points, the variable Δ is used.

Form Factor
The form factor is used to account for the lipid prisms. These consist of a water core surrounded by a cylindrical shell of lipids with their heads pointing towards the water. When fitting cylinders in a hexagonal arrangement, there will be voids generated in the corners of the hexagons. The lipids would then have to stretch out to fill these voids which would cause stress and therefore affect the measured intrinsic curvature. To eliminate this stress, tricosene is added to the sample to fill up these voids.

2. Methods – From Lipid Dust to Curvature

The form factor describing a hexagonal prism, with length L, can be separated in its lipid core part F_{lipid} and the part filling up the spaces between the cylinders F_{inter}[7].

$$F(q, \theta | \boldsymbol{\rho}) = f(q, \theta) \int \rho(r, \phi) J_0(q\, r\, \sin \theta) r\, dr\, d\phi$$
$$= f(q, \phi) \left[F_{lipid}(q, \theta | \boldsymbol{\rho}) + F_{inter}(q, \theta | \boldsymbol{\rho}) \right] \tag{13}$$

$$f(q, \theta) = 4\pi\, \sin \left(\frac{L}{2} q \cos \theta \right)$$

$$F_{lipid}(q, \theta | \boldsymbol{\rho}) = \int_0^{a/2} r \Delta \rho(r) J_0(q\, r\, \sin \theta)\, dr =$$
$$\frac{1}{q \sin \theta} \left[\Delta \rho_M r_M J_1(q\, r_M \sin \theta) + \sum_{k=1}^{M-1} (\Delta \rho_k - \Delta \rho_{k+1} r_k) J_1(q\, r_k\, \sin \theta) \right] \tag{14}$$

With the SLD relative to water $\Delta \rho = \rho - \rho_w$ (ρ_w=0.33 Å), the radius of the cylinders r, the Bessel function of the first kind of i$^{\text{th}}$ order J_i. The sum here is summing over the section as seen in figure 16 thus M is the total number of slabs.

The form factor of the interstices is [7]:

$$F_{inter}(q, \theta | \boldsymbol{\rho}) = \frac{6}{\pi} \int_{a/2}^{a/\sqrt{3}} dr \int_0^{\pi/6 - \arccos(\frac{a}{2r})} r \Delta \rho_{inter} J_0(qr \sin \theta)\, d\phi =$$
$$\Delta \rho_{inter} \int_{a/2}^{a/\sqrt{3}} \left[\frac{\pi}{6} - \arccos \left(\frac{a}{2r} \right) \right] r J_0(qr \sin \theta)\, dr \tag{15}$$

As seen here the integral over r has to be integrated numerically, since integrating like that is rather expensive this cannot be done often. Fortunately F_{inter} does not depend on any fitting parameters and can therefore be calculated in advance.

As it will be described later, the SLD is built up from a simple slab model which gives sharp steps in the profile. In reality, this profile is of course much smoother due to molecular fluctuations. Analog to [7], it is accounted for these fluctuations by translating the shell boundaries by the distances x which are distributed like a Gaussian function $N(x | \mu = 0, \sigma_{fluc}^2)$ with mean μ and variance σ^2. This can be seen as a polydispersity of the slabs inside the lipid unit cell. Here these fluctuations are applied to the whole form factor.

$$F(q, \theta | \boldsymbol{\rho}, \sigma_{fluc}) = \int dx\, N(x | \mu = 0, \sigma_{fluc}^2)\, F(q, \theta | \boldsymbol{\rho}) \tag{16}$$

Diffuse scattering
Analog to previous publications [7, 16], the experimental data of this work shows significant diffuse scattering especially between the first two peaks (see figure 18). It is assumed that this

diffuse signal arises from the formation of a lamellar structure, a monolayer or a combination of both at the boundaries between the hydrocarbons and water. Depending on the lipid species, a diffuse form factor for a laterally uniform, infinitely long mono- or bilayer is calculated from (17). The SLD $\boldsymbol{\rho}_{diff}(z)$ depends on the lipid species and the presumed structure.

$$F_{diff}(q|\boldsymbol{\rho}_{diff}) = 4\pi^2 \int_{-\infty}^{\infty} \rho_{diff}(z)\, e^{iqz}\, dz \tag{17}$$

A simplified version of the slab model described in section 2.2.2 is used to define $\boldsymbol{\rho}_{diff}$. For the monolayer, it just consists of one slab for the hydrocarbon chains and another for the lipid head. The SLD is defined by the area per lipid in the monolayer A_{mono} and the lipid volume V of the H_{II}-phase-forming lipid, giving $\boldsymbol{\rho}_{mono} = \{A_{mono}, V\}$. Using equation (17) with this SLD gives the following relationship.

$$F_{mono}(q|\boldsymbol{\rho}_{mono}) = F_{diff}(q|\boldsymbol{\rho}_{diff} = \boldsymbol{\rho}_{mono}) =$$
$$\frac{4\pi^2 i}{q}\left[\Delta\rho_{HC,mono}\left(e^{-iqd_{HC,mono}} - 1\right) + \Delta\rho_{H,mono}\left(1 - e^{iqd_{H,mono}}\right)\right] \tag{18}$$

Analogously, the form factor for the bilayer phase can be defined by the area per lipid in the bilayer A_{bi} and the volume of the guest lipid V_g, giving $\boldsymbol{\rho}_{bi} = \{A_{bi}, V_g\}$.

$$F_{bi}(q|\boldsymbol{\rho}_{bi}) = F_{diff}(q|\boldsymbol{\rho}_{diff} = \boldsymbol{\rho}_{bi}) =$$
$$\frac{-4\pi^2 i}{q}\left[\Delta\rho_{HC,bi}\left(e^{iq(d_{H,bi}+2d_{HC,bi})} - e^{iqd_{H,bi}}\right)\right.$$
$$\left. +\Delta\rho_{H,bi}\left(e^{iqd_{H,bi}} - 1 + e^{i2q(d_{H,bi}+d_{HC,bi})} - e^{iq(d_{H,bi}+2d_{HC,bi})}\right)\right] \tag{19}$$

For pure H_{II} phase forming lipid species only a monolayer, surrounding the outermost boundary of the hexagonal structures, is assumed. For mixed lipid species, where the guest lipid does form bilayer structures, also a bilayer term is added. This bilayer is assumed to consist solely of the guest lipid and not to interact with the H_{II} structures. Of course, we cannot exclude that the diffuse signal originates from other aggregates formed during sample preparation.

Intensity
Putting all the parts together, while considering orientational averaging, the total scattering intensity can finally be obtained.

$$I(q) \propto \int_0^\pi |F(q,\theta|\boldsymbol{\rho},\sigma_{fluc})|^2\, S(q,\theta|n,\Delta)\sin\theta\, d\theta$$
$$+2c_{mono}\, F_{mono}(q|\boldsymbol{\rho}_{mono})\int_0^\pi F(q,\theta|\boldsymbol{\rho},\sigma_{fluc})s(q,\theta|n,\Delta)\sin\theta\, d\theta \tag{20}$$
$$+c_{mono}^2|F_{mono}(q|\boldsymbol{\rho}_{mono})|^2 + c_{bi}^2|F_{bi}(q|\boldsymbol{\rho}_{bi})|^2$$

Where c_{mono} and c_{bi} are used to scale the monolayer respectively bilayer form factors. For evaluations where solely H_{II} phase forming lipids are considered, c_{bi} is set to 0. The structure factor $s(q, \theta | n, \Delta)$ occurring in the mixed term is derived analogously to equation (12) and reads as follows:

$$s(q, \theta | n, \Delta) = \frac{1}{\sqrt{N_{hex}(n)}} e^{-q^2 \Delta / 2} \sum_{j \neq k}^{N_{hex}(n)} J_0(q |\boldsymbol{R}_j - \boldsymbol{R}_k| sin\theta) \tag{21}$$

2.2.2. Electron Density – Slab Model

For X-Ray scattering the SLD is the electron density of the specimen. The model for the SLD is based on the one from [7] but some adaptions were made. First and foremost it is now accounted for tricosene penetrating into the hydro-carbon chain region and the width of the lipid backbone is fixed to $d_{BB} = 4.6 \, \text{Å}$ [14].

The PE structure is parsed into 3 sections of a wedge shaped slab with opening angle α and height h. Besides the water core these sections are the headgroup, the glycerol backbone, and the hydrocarbon tails. It is also assumed that some added filler molecule, tricosene, interdigitates into the latter section which slightly alters the electron density there. This lipid unit cell, which holds exactly one lipid, is shown in figure 17.
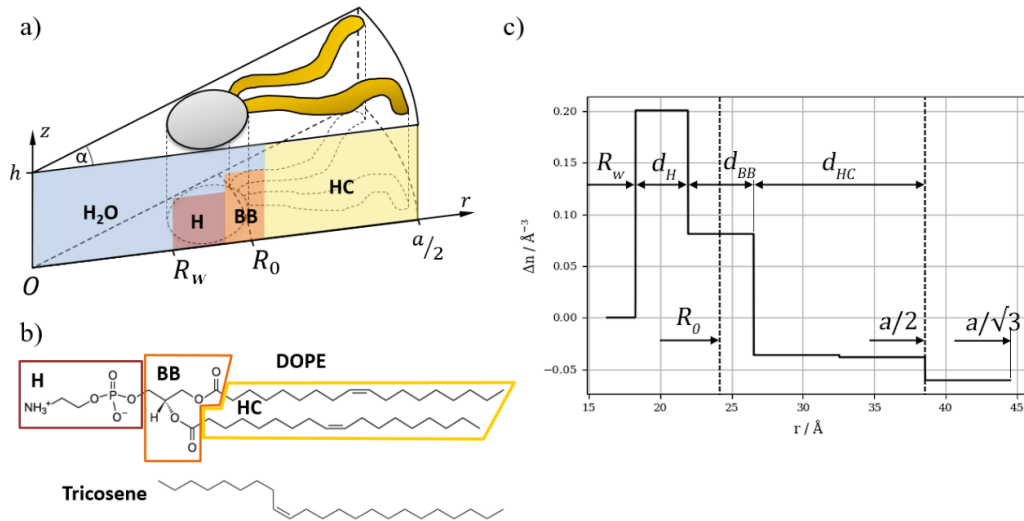
Figure 17: Modeling of the SLD for phosphatidylethanolamines: a) The lipid unit cell consisting of a cylinder sector with radius $a/2$, separated into the sections which build up the lipid. From left to right these sections are the water core, lipid head (H), lipid backbone (BB) and the hydrocarbon tail (HC). R_0 denotes the radial position of the neutral plane. b) Chemical structure of DOPE and tricosene. c) A corresponding electron density profile.
Source: [7]

The outer radius of this wedge-shaped structure is $a/2$. Fortunately the lattice constant a is

easily calculated from the Bragg peak positions via $q_{kl} = \frac{4\pi}{\sqrt{3}a}(k^2 + 2kl + l^2)$, with the position of the Bragg peaks q_{kl} with k and l being the Miller indices.

As mentioned earlier, the intrinsic curvature is measured at the neutral plane. The radial position of this plane is assumed to be at the center of the backbone slab. This assumption is backed by previous works from Frewein and Kollmitzer [7, 16].

The fit parameters which are defining the SLD for a single lipid species are the intrinsic curvature $C_0 = \frac{-1}{R_0}$, the radius of the water core R_w and the lipid volume V_{lipid}. Since the lattice constant is known from the peak positions of the data and the backbone is fixed, the remaining parameters, such as the width of the headgroup d_H and the hydrocarbon tails d_{HC}, are defined.

$$d_H = R_0 - \frac{d_{BB}}{2} - R_w \tag{22}$$

$$d_{HC} = \frac{a}{2} - R_0 - \frac{d_{BB}}{2} \; . \tag{23}$$

For X-rays the SLDs, i.e. the electron densities, of each sector are given by $n_k = N_k/V_k$ with k being one of the shells accounting for the headgroup, backbone or tail, the number of electrons of the kth shell N_k and its volume V_k. For DOPE the volumes for head and backbone are $V_H = 110\,\text{Å}^3$ and $V_{BB} = 135\,\text{Å}^3$ [19]. The remaining volume of the hydrocarbon tails is then $V_{HC} = V_{lipid} - V_{BB} - V_H$.

The density of tricosene was measured with the DMA 5000 M Density Meter from Anton Paar, and from this the electron density could be calculated with the molar mass and the number of electrons per molecule. This way the electron density is estimated to be $\rho_{tric} = 0.2759\,\text{Å}^{-3}$ at $20\,\text{deg}\,°\text{C}$ with a linear temperature dependency with a slope of $k = -0.000\,225\,\text{Å}^{-3}\,\text{K}^{-1}$.

Now the SLD is sufficiently defined with the fitting parameters C_0, R_w and V_{lipid}, giving $\boldsymbol{\rho} = \{C_0, R_w, V_{lipid}\}$. These are free parameters that will get optimized during fitting, the lipid backbone width is set to $4.6\,\text{Å}$, see section 2.2.2. All other parameters can be calculated from the lipid contribution to the volume of the kth shell [7].

$$V_k = \frac{\hat{A}(r_{k+1}^2 - r_k^2)}{2} - \tilde{n}_w^k V_w \tag{24}$$

With the number of water molecules in a shell $\tilde{n}_w^k$ and the molecular volume of water $V_w = 30\,\text{Å}^3$. The mantle area of a sector with unity radius, $\hat{A} = h\,\alpha$, can be calculated using equation (24) with $k = \text{HC}$ and $\tilde{n}_w^k = 0$ as follows:

$$\hat{A} = \frac{2(V_{HC} + V_{tric,HC})}{\frac{a^2}{4} - C_{0,HC}^{-2}} \tag{25}$$

With $R_0 + \frac{d_{BB}}{2} = \frac{1}{C_{0,HC}}$, the width of the backbone slab d_{BB}, the hydrocarbon chain volume V_{HC} and the lattice constant a. $V_{tric,HC}$ describes the volume per lipid which enters the region of the hydrocarbon tails. The intrinsic curvature $C_{0,HC}$ here is measured at the

hydrocarbon interface which is the plane between the lipid backbone and its chains. Since it is arbitrary where to measure the curvature, it is utilized that the chain volume is known. Equation (24) also defines the amount of water in the head and backbone regions, $\tilde{n}_w^H$ and $\tilde{n}_w^{BB}$. The difference here to [7] lies in the added tricosene volume in the hydrocarbon chain region. This assumption helps to achieve better convergence and is backed up by an all-atom molecular dynamics simulation [2], which will be discussed more detailed in the Results section. Furthermore, from geometrical considerations, it follows that the added $12\,\mathrm{w}\%$ of filler molecule generally does not fit into the volume between the rods, but is also incorporated into the hydrocarbon region.

With $\hat{A}$ it is possible to calculate the area per lipid at any position within the molecule, e.g. at the neutral plane:

$$A_0 = \hat{A} R_0 \tag{26}$$

2.2.3. Tricosene penetrating into Hydrocarbon Tails

As mentioned above, this model accounts for excess filler material, that means tricosene which exceeds the space between the cylinders, penetrating into the hydrocarbon tails region. The volume fraction between the lipid tubes follows from geometrical considerations, see the regions shaded in gray in figure 4.

$$\Gamma = \frac{2\sqrt{3} - \pi}{2\sqrt{3} - \frac{4}{a^2} R_w^2 \pi} \tag{27}$$

For this consideration, the water core gets subtracted to really calculate the volume fraction per lipid.
The $12\,\mathrm{wt}\%$ of tricosene added to the sample can be converted to the volume fraction v as follows [21].

$$v = \frac{m_{tric}/\rho_{tric}}{m_{lip}/\rho_{lip} + m_{tric}/\rho_{tric}} \tag{28}$$

With these two equations, it is possible to calculate the volume fraction of filler molecule exceeding the available space γ and therefore interdigitating into the hydrocarbon tails.

$$\gamma = \begin{cases} v - \Gamma\,, & \text{if } v > \Gamma \\ 0\,, & \text{otherwise} \end{cases} \tag{29}$$

Now that the volume fraction is known, it can be accounted for it in the electron density profile. The outer half of the HC slab will be a weighted linear combination of the electron density of the lipid tails and the one for tricosene with the guest concentration as weight.

2.2.4. Intrinsic Curvature of the Guest

As described in section 1.2, the curvatures of the host and guest lipid are added accounting for the possibly different head sizes. Due to the arbitrariness on where the intrinsic curvature is measured, the following considerations are made easier by defining all areas and curvatures

at the hydrocarbon interface and transforming the result to the neutral plane at the end. This makes some calculations easier, because then only the hydrocarbon slab has to be considered. Moving everything to the hydrocarbon interface transforms equation (9) into the following.

$$C_{0,HC} = \frac{(1-x)C_{host,HC} + x\,\xi\,C_{guest,HC}}{(1-x) + x\,\xi} \tag{30}$$

With the total intrinsic curvature $C_{0,HC}$, the curvature of the host lipid $C_{host,HC}$ and the curvature of the guest lipid $C_{guest,HC}$, all curvatures measured at the hydrocarbon interface, the mol fraction of the guest lipid in the hex phase x and the head-ratio of the host and guest lipid ξ.

The head-ratio is the ratio of the head widths (see figure 9), here at the hydrocarbon interface. The lipid head areas are the areas resulting from a cross section through the lipid, they are considered to be quadratic.

$$\xi = \frac{b_{guest}}{b_{host}} = \sqrt{\frac{A_{guest,HC}}{A_{host,HC}}} \tag{31}$$

With A_i describing the hosts respectively guests head area measured along the hydrocarbon interface.

It is assumed that the areas per lipid add up linearly:

$$A_{HC} = (1-x)A_{host,HC} + xA_{guest,HC} \tag{32}$$

Putting together equations (31), $A_{guest,HC} = \xi^2 A_{host,HC}$ from (32) and $A_{HC} = \hat{A}R_{HC}$, see (26), leads to the expression:

$$\hat{A} = -C_{0,HC}(1 - x + \xi^2 x)A_{host,HC} \tag{33}$$

Setting equal equations (25) and (33) finally gives a quadratic equation for the total intrinsic curvature measured at the hydrocarbon interface $C_{0,HC}$.

$$C_{0,HC}^2 + C_{0,HC}\frac{8V_{HC}}{a^2 A_{host,HC}(1 - x + \xi^2 x)} - \frac{4}{a^2} = 0 \tag{34}$$

With its discriminant being:

$$\delta = \left(\frac{8V_{HC}}{a^2 A_{host,HC}(1 - x + \xi^2 x)}\right)^2 - \frac{16}{a^2} \tag{35}$$

For physical values of the parameters only real solutions ($\delta > 0$) for the curvature are expected and since one of the requirements for this approach is that the total curvature still has to be negative, only the negative solution is interesting.

$$C_{0,HC} = \frac{-4V_{HC}}{a^2 A_{host,HC}(1 - x + \xi^2 x)} - \frac{\sqrt{\delta}}{2} \tag{36}$$

The total intrinsic curvature at the neutral plane is then $C_0 = \frac{-1}{-1/C_{0,HC} - d_{BB}/2}$. Then everything is known to calculate the intrinsic curvature of the guest lipid using equation (30), with the known curvature of the host lipid $C_{0,host}$. The curvatures in this last equation are again at the neutral plane.

$$C_{0,guest} = \frac{C_0(1 - x + x\xi) - C_{0,host}(1 - x)}{x\xi} \tag{37}$$

2.2.5. SLD for Lipid-Mixtures

For mixed lipid species the unit cell consists of a kind of hybrid lipid. That means that the SLD is very similar as in the case for a single lipid. Only the electron densities for each section n_k, see figure 17, are now a linear combination of the host and guest lipid densities $n_{k,host}$ respectively $n_{k,guest}$.

$$n_k = (1 - x)n_{k,host} + xn_{k,guest} \tag{38}$$

This represents a simplification compared to previous works [7, 8]. Again the electron densities are calculated from the number of electrons per volume, $n_{k,host/guest} = N_{k,host/guest}/V_{k,host/guest}$, with volumes for head and backbone for DPPC being $V_{H,DPPC} = 192\,\text{Å}$ and $V_{BB,DPPC} = 139\,\text{Å}$ [18].

2.3. Computational Application

The goal is to gather information about the system, most importantly here the intrinsic curvature, by optimizing the parameterized model described above to the available SAXS data. Bayesian statistics offers a mighty tool for such analysis and shall be discussed briefly before the used framework is described.

2.3.1. Markov Chain Monte Carlo

The term 'Markov Chain Monte Carlo' (MCMC) describes a variety of algorithms for sampling from probability distributions. This is achieved by utilizing successive random events, whereas each generated point depends on its predecessor. This dependency is what distinguishes MCMC from ordinary Monte Carlo methods and allows a wide variety of algorithms where the only requirement is that the distribution the samples are generated from converges to the target distribution [27].

The ensemble of all saved generated points, e.g. a series of random variables $x_0, ... x_N$, in such an algorithm is called the Markov chain. The goal is to construct such a chain which is, after enough time, of the desired stationary distribution $\tilde{\pi}(x)$.

Metropolis-Hasting algorithm
Probably the most popular MCMC algorithm is the Metropolis-Hasting algorithm [11]. It consists roughly of these steps:

1. Start with an arbitrary initial state $x = x_0$.

2. Propose a new random state x'. The proposal depends on the last state x.

3. Accept the new state if it fulfills following acceptance criterion[27]:

$$A(x', x) = min\left(1, \frac{p(x')p_p(x|x')}{p(x)p_p(x'|x)}\right) \tag{39}$$

4. If x' is not accepted, the next step of the chain will again be x.

The steps 2 and 3 are repeated until the chain converges or a maximum number of Markov steps is reached.
Here $p(x)$ denotes the likelihood function, i.e. the probability of measuring the data with a given parameter x. The variable $p_p(x'|x)$ stands for the proposal function which describes the probability of proposing a state x' given the state x.

This simple univariate example can easily be generalized to the multivariate model described in section 2.2. The likelihood function is then defined as:

$$p(d_l|\boldsymbol{x}, \tilde{\sigma}_l) = \prod_{l=1}^{N_q} \frac{1}{\tilde{\sigma}_l\sqrt{2\pi}} \exp\left[-\frac{1}{2\tilde{\sigma}_l^2}(I(q_l|\boldsymbol{x}) - d_l)^2\right] \tag{40}$$

with the observed SAXS data d_l, the weighted error $\tilde{\sigma}_l$, the number of data points N_q, the parameter set $\boldsymbol{x}$ and the model function $I(q_l|\boldsymbol{x})$.
The proposal function $p_p(\boldsymbol{x}_i|\boldsymbol{x}_j)$ describes the probability of proposing a multivariate parameter set depending on the state before and some prior knowledge of each parameter defining its lower and upper boundaries.

2.3.2. Weighting of the error

To optimize the acceptance during fitting the error provided by the SAXS measurement gets modified with a simple multiplicative scaling factor and a weight. The multiplicative factor is also optimized during the MCMC run.
The weight helps to emphasize the mid to high q-range more, otherwise the large peaks in the beginning have too much impact due to their high value to error ratio, and follows this function:

$$y = e^{-\gamma(q-q_0)} \tag{41}$$

Where q denotes the modulus of the scattering vector, q_0 its smallest value and γ an optimized parameter. This exponential decrease is applied to the errors and in doing so increases the signal to error ratio the more, the smaller it gets. The parts of the data with a high signal to error ratio have more impact on the fit.

2.3.3. With a little Help from my Python library – Pymc3

Although the general algorithm looks simple, the difficulty lies in the detail. It is not trivial to choose the proposal function which is crucial for getting a well converging chain. Since this is no work on theoretical computer science it is not necessary to implement an advanced MCMC sampler and rather use one of the many, very powerful options available.

For the present work a python library called 'PyMC3' is used [25]. It can be installed using the python package management and has a rather straight forward syntax.

2.3.4. Fitting Parameters

The possible fitting parameters for single as well as mixed lipid species are listed in this section. All parameters are drawn from flat distributions with defined lower and upper borders.

Single Lipid
The used fitting parameters for the single lipid case are shown in table 1, since the model is similar to the one from [7] the used fitting parameters are almost the same. The difference lies in the lamellar form factor. This work uses a monolayer form factor instead of one describing a bilayer.

Table 1: Fitting parameters for single lipid fits

Occurrence	Symbol	Description
Structure Factor	Δ	Mean square displacement of the lattice points
	n	Number of hexagonal shells (lattice size)
H_{II} form factor	C_0	Intrinsic curvature of the Lipid
	V	Lipid Volume
	R_w	Radius of the water core
	σ_{fluc}	Polydispersity of the unit cell
Mono-layer form factor	c_{mono}	Scaling constant for the monolayer phase
	A_{mono}	Area per lipid in the monolayer
Signal processing	Γ	Multiplicative scaling factor
	I_{add}	Additive constant accounting background

Lipid Mixture
In table 2 the fitting parameters used for lipid mixtures are shown. They differ from the ones above by some details. The head slab is now used for both lipids at once where their electron densities get linearly combined and the fitted head-ratio is used to calculate the intrinsic curvature of the whole system using equation (36). The total curvature is then used to calculate the curvature of the lipid of interest, see equation (37).
Lipid mixtures were mostly fitted using a joint analysis which will be described in the next section. Some variables were optimized as joint variables, i.e. the same parameter is opti-

mized for several data sets at once.

The exact guest lipid concentration x in the H$_{II}$ phase is unknown. It can vary due to preparation errors or to the forming of monolayer respectively lamellar phases. To account for this error during fitting, the concentrations are not added as regular fitting parameters but still they are varied, drawing from a Gaussian distribution around the desired concentration with a standard deviation of 10%. This keeps the fit more flexible and helps to avoid local minima.

Table 2: Fitting parameters for lipid mixtures
*Parameters fitted as joint variables. **If available, literature values are used.

Occurrence	Symbol	Description
Structure Factor	Δ	Mean square displacement of the lattice points
	n	Number of hexagonal shells (lattice size)
H$_{II}$ form factor	$*\xi$	Head-ratio of the lipids head widths
	R_w	Radius of the water core
	$*V_g$	Lipid Volume of the guest lipid
	σ_{fluc}	Polydispersity of the unit cell
Lamellar form factors	c_{mono}	Scaling constant for the monolayer phase
	c_{bi}	Scaling constant for the lamellar phase
	$**A_{bi}$	Area per lipid in the bilayer
Signal processing	Γ	Multiplicative scaling factor
	I_{add}	Additive constant accounting background

2.4. Fitting multiple data sets – Joint Analysis

If data for several concentrations is available it can be interesting to compare results for all data sets and see if they meet expectations. Since some parameters, like the volume or the head-ratio, are supposed to be consistent for all concentrations it is even more handy to be able to use multiple data sets to optimize these variables. Commonly optimized variables like that, where the same variable is used for different concentrations, will be called common variables whereas variables that are uniquely fitted for each different concentration will be called unique variables.

In this joint analysis the model is optimized to several data sets of different concentrations at once. The common parameters benefit from the additional information and while possibly the errors increase, they are better defined because more information was used for the fit. Other parameters like the scaling factors could be different for each concentration and therefore need to be fitted as unique parameters. This is increasing the number of parameters for the joint analysis.

One of the benefits of the joint analysis is that by doing just one MCMC run all the information over different concentrations is available. The quality of the result is immediately

visible when comparing the different results for the intrinsic curvature arising from unique variables.

3. Results

In this section the results of the global x-ray scattering analysis are shown. The following is the product of a Markov-Chain Monte Carlo fit algorithm applied to the SAXS data.

3.1. DOPE – Host for most Bilayer Lipids

DOPE will be used as the main host to measure other bilayer forming lipids. Those lipids are mixed into the hexagonal phase and their influence on the system is utilized to evaluate their intrinsic curvature. Although the model works best for matching chain lengths, DOPE is used for a variety of lipids with different lengths. For one lipid, DMPC that is, there is also a different host lipid with matching chain length used. In figure 18 the results for pure DOPE at 35 °C are shown.
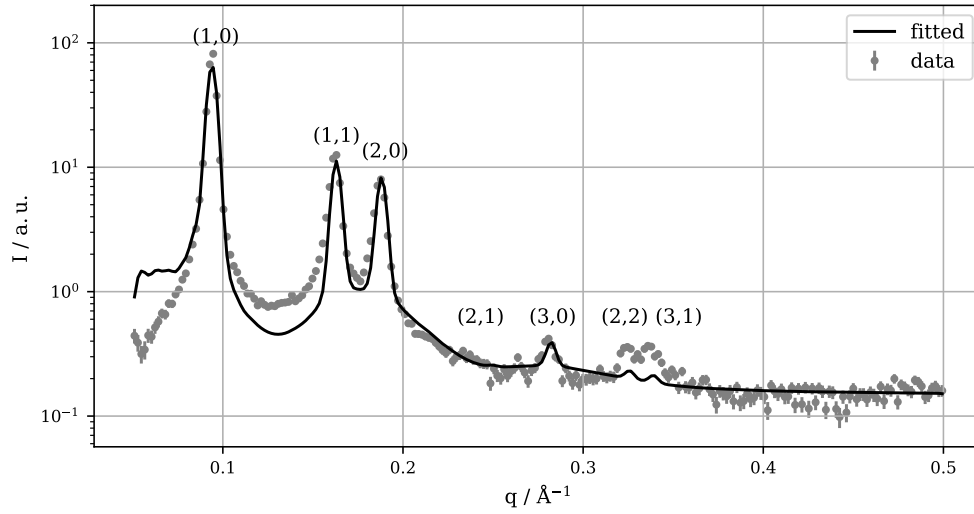
(a) SAXS data and fit. Numbers in brackets are the Miller indices.
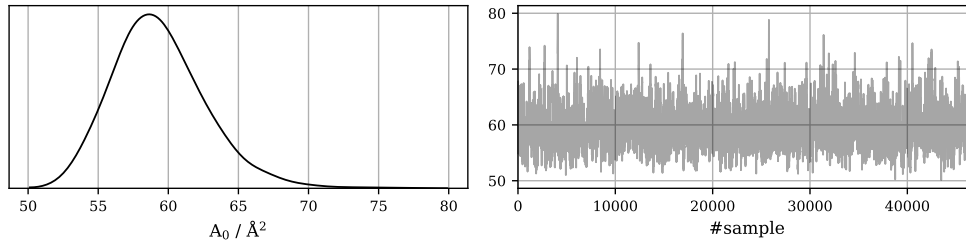
(b) Histogram of the head area, measured at the neutral plane, resulting from the fit. On the right the trace of all accepted Monte-Carlo steps is shown.
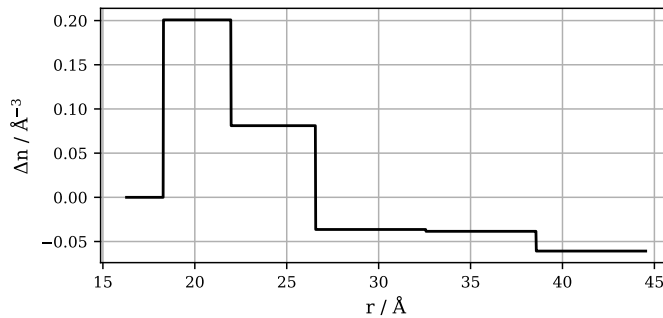
(c) Corresponding scattering length density profile. The electron density Δn is measured relative to the one of water.

Figure 18: Results of the global model for pure DOPE at 35 °C. This lipid is later used as host lipid for binary mixtures with bilayer forming lipids. The intrinsic curvature of this guest lipids can be evaluated from their influence on the host system.

One thing which is specific for DOPE is that the fourth peak, the one with miller index (2,1), is vanishing in the data. This was a discrepancy in the results of [7] where this peak

was present in the model. By substituting the bilayer with a monolayer form factor and emphasizing the higher q-range, see section 2.3.2, a better agreement around this peak is achieved. The model coincides slightly worse between the first two peaks but parameters like the head area are closer to expectations which supported the decision.

The SLD arising from fitting this data is shown in figure 18c. It is mainly defined by the intrinsic curvature, the radius of the water core, the width of the backbone, the lattice constant and the volumes, respectively the electron densities of each slab. The small kink at around $r = 32.5\,\text{Å}$ results from the excess filler molecule which is interdigitating into the lipid chains.

In figure 19 the parts which make up the models' intensity, i.e. the form factor of the cylinders, the form factor of the monolayer phase which accounts for the diffuse scattering signal and the structure factor, are plotted individually. The fact that the fourth peak of the structure factor aligns exactly with the second minimum of the monolayer form factor, thus eliminating the peak, is very encouraging and supports the change in the diffuse scattering calculation. In the next section, the cylinder form factor will be compared to one generated via a Molecular Dynamics simulation.
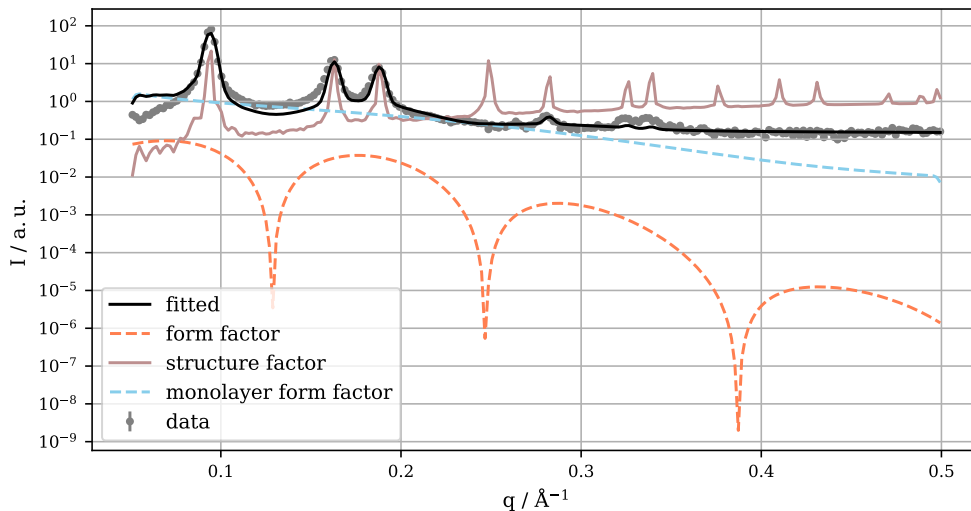
Figure 19: Different parts of the model plotted individually.

The complete results of the global analysis of pure DOPE data at different temperatures is stored in table 3, future tables with results of single lipid fits will only contain the most relevant parameters. Despite some concepts have changed and the model was re-parameterized, the results are generally in very good agreement with the ones presented by M. Frewein [7] which are shown in the last column. The fit quality was improved and the fit itself is more stable now which will be more obvious when lipid mixtures are analyzed.

The lattice constant a is calculated from the signals Bragg peak positions. Its main source for the error is the determination of these positions, so to account for deviations there, the process is repeated 10 times and the standard deviation over all calculations is used. Variables like the width of the lipid head slab d_h and the area per lipid at the neutral

plane A_0 are calculated from fit parameters, their errors are the standard deviations of their distributions.

Table 3: Results for pure DOPE at different temperatures. Variables above the dashed line are the lattice constant a which is calculated from Bragg peak positions, the width of the lipid head d_h (without backbone) and the area per lipid measured at the neutral plane A_0. The variables below the dashed line are fitting parameters, see table 1, with the mean as value and standard deviation as error.

Symbol	$T = 25\,°\text{C}$		$T = 35\,°\text{C}$		$T = 50\,°\text{C}$		$T = 25/35/50$ °C results from [7]
a / Å	79.8	± 0.5	77.1	± 0.5	73.6	± 0.5	80.4/76.9/73.5
d_h / Å	3.2	± 0.8	3.7	± 1.1	3.0	± 0.7	4.8/5.2/5.1
A_0 / Å^2	58	± 3	59	± 3	63	± 3	-/62/-
Δ	11	± 2	9	± 2	9	± 2	2.9/3.3/6.7
n	17	± 2	18	± 3	27.8	± 1.6	18/23/29
C_0 / Å^{-1}	-0.039	± 0.001	-0.041	± 0.001	-0.043	± 0.001	-0.039/-0.041/-0.043
V / Å^3	1152	± 11	1154	± 17	1166	± 16	1141/1142/1152
R_w / Å	19.9	± 0.6	18.3	± 0.9	17.8	± 0.7	
σ_{fluc}	0.159	± 0.008	0.153	± 0.018	0.166	± 0.014	0.163/0.163/0.176
c_{mono}	3.4	± 0.4	2.9	± 0.4	3.9	± 0.5	
A_{mono} / Å^2	77	± 6	63	± 5	88	± 6	61.1/64.9/65.4
Γ	3.8	± 0.4	3.4	± 0.6	8.8	± 1.2	8.18/9.65/7.13
I_{add} / a.u.	0.156	± 0.004	0.148	± 0.005	0.14	± 0.01	0.133/0.132/0.135

The result of the MD simulation for the area per lipid at the neutral plane is $A_{0,sim} = 54\,\text{Å}^2$, which within the error is in good agreement with the results found in the present work.

For measurements at temperatures other than the ones shown in this table, a linear interpolation is used, see figure 20. All presented variables show a good linear temperature dependency which is a nice indicator for the reliability of the results.
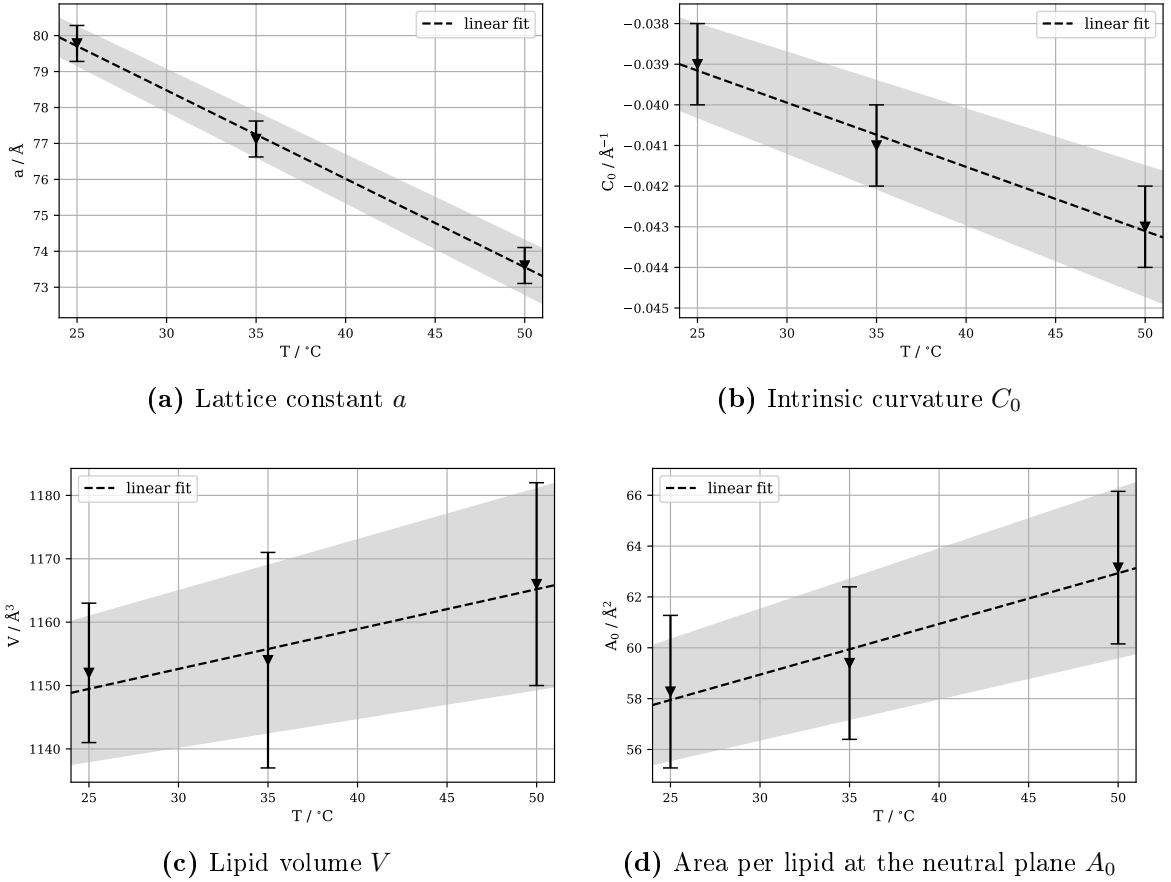
(a) Lattice constant a

(b) Intrinsic curvature C_0

(c) Lipid volume V

(d) Area per lipid at the neutral plane A_0

Figure 20: Linear interpolation for some DOPE results over temperature.

3.1.1. Comparison to Simulations

In this section the DOPE fit results at 35 °C are compared to the MD simulation provided by Douwe Jan Bonthuis [2]. In figure 21 the electron density coming from the fit is compared to the one arising from the simulation.

For an easier comparison, the step-function-like fit result is convoluted with a Gaussian function with $\sigma = 20$ Å. This rather large standard deviation suggests that the radial resolution is bad, nevertheless the smeared out electron density distribution is in a nearly perfect agreement with the one from the simulation. This indicates, that the applied polydispersity of the slabs mimics a real smooth electron density very well. On the one hand the standard deviation $\sigma = \sigma_{fluc} \cdot a$ used to apply the polydispersity converges to 12 Å which is lower than the 20 Å used for this comparison, on the other hand the radial position of the main peak is almost identical to the simulation and the standard deviation used for the convolution is not directly connected to the applied polydispersity.
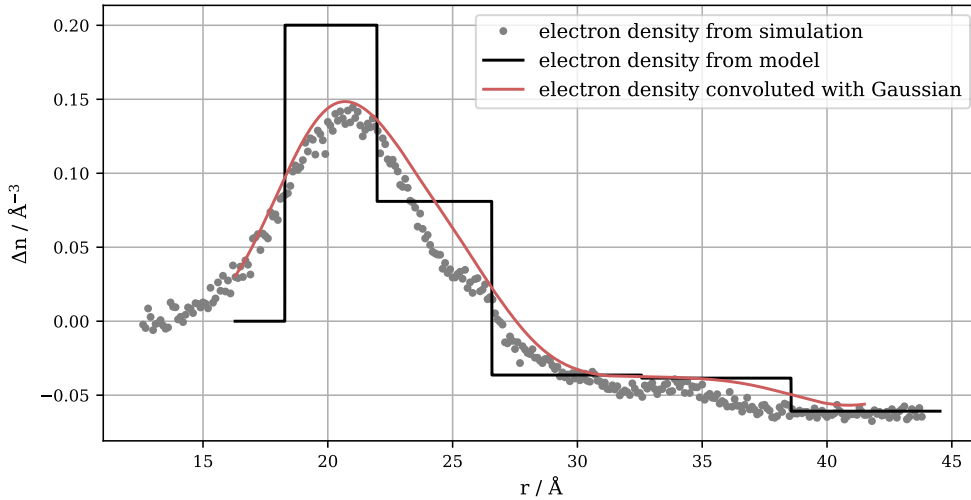
Figure 21: Electron density of DOPE at 35 °C compared to the simulation.

To increase the performance of the calculation, a different way of applying the polydispersity was tried. By simply multiplying a Gaussian function with the form factor, which should be analogous to the convolution of the electron density, the time needed for the calculation could be cut by magnitudes. While the performance increase is a benefit, this approach did deliver a worse agreement to the simulations. Also, the fit quality was decreased, so it seems not to be equivalent to (16) and therefore was rejected.

In figure 22 the form factor is compared to the simulation. The intensity of the shown curves is arbitrary, therefore they're scaled in a way that they align at their highest intensities. It looks like the polydispersity of the lipid unit cell shifts the minima slightly towards lower q values and doing so increasing the agreement with the simulation. Another thing that changes with polydispersity is the slope of the form factor which is very much-needed to resemble the data well in the high q-range. It was also possible to get rid of the additional minimum of the form factor, described in the beginning of this thesis, by applying the polydispersity also to the filler molecule slab. This was avoided in the work by M. Frewein due to its computational load.
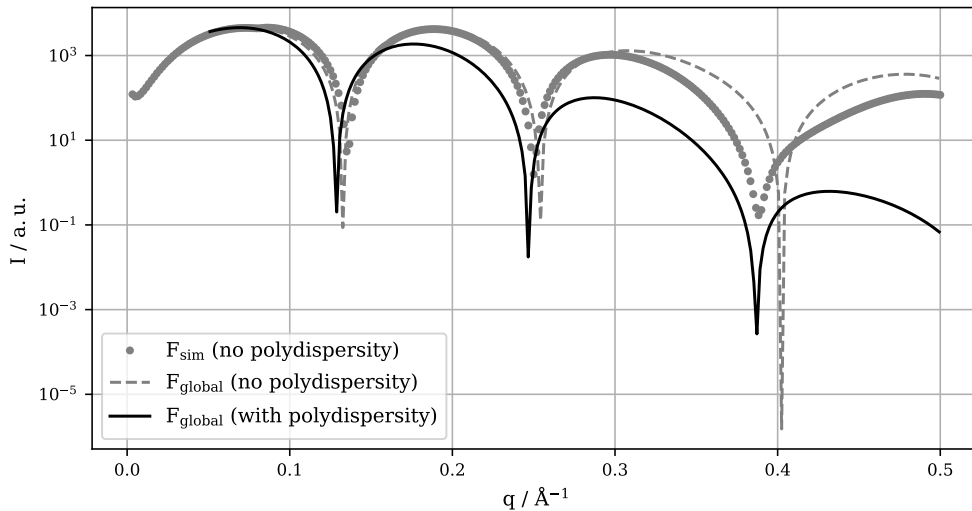
Figure 22: Form factor of DOPE at 35 °C F_{global} compared to the simulation F_{sim}.

To have another comparison to the simulated form factor, it is used to calculate the intensity by replacing the form factor from the global model there. In figure 23 this comparison is shown. The dashed black line shows the plain convolution of the form factor of the simulation F_{sim} with our structure factor, while the red line also includes the terms accounting for diffuse scattering. The parameters from table 3 at 35 °C were used. As it can be seen, also the simulated form factor cancels out the (2,1) peak and overall coincides well with the data. A certain deviation, as between the first two peaks, is to be expected because the model was not optimized for F_{sim}.
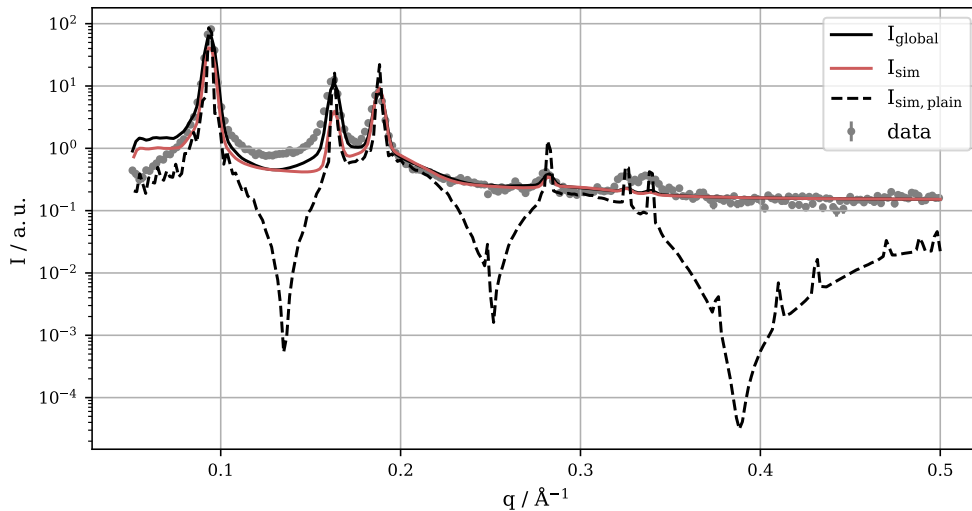

Figure 23: Replacing the form factor in the global model by the one resulting from the simulation. I_{global} arises from the regular global model, $I_{sim,plain} = S \cdot |F_{sim}|^2$ is the plain convolution of the structure factor from the global model with the form factor from the simulation and I_{sim} also includes the diffuse scattering.

3.1.2. Neutral Buoyancy – Successfully Failed Lipid Volume Verification

Neutral buoyancy occurs when the density of an object immersed in a fluid fully matches the fluids density and therefore balances the buoyant force and the force of gravity. In that state the object will float where it is, neither going up nor down.

This state of neutral buoyancy can be exploited to measure the specific lipid volume [29] and was previously performed for lipids in lamellar phases [10]. The method was considered to verify the lipid volume result of DOPE of the MCMC calculation. DOPE assembles H_{II} structures when in contact with water at room temperature for which no buoyancy data is available. In order for it to work, the specimen has to be heavier than regular water so it is possible to gradually increase the density of water, by deuterating it, and subsequently get to the neutral buoyancy state.

Equation (42) describes the composition of the considered densities for the solvent ρ_{solv} and the lipid ρ_{lipid}, x_i stands for the weight percentages of the ith part and ρ_i for its density. The density of DOPE is calculated using the volume from the fit results and its published molecular weight of 743.547 g/mol [20].

$$\frac{1}{\rho_{solv}} = \frac{1 - x_{D_2O}}{\rho_{H_2O}} + \frac{x_{D_2O}}{\rho_{D_2O}} \overset{!}{=} \frac{1 - x_{tric}}{\rho_{lipid}} + \frac{x_{tric}}{\rho_{tric}} = \frac{1}{\rho_{lip+tric}} \tag{42}$$

By balancing the left and right side, thereby fulfilling the requirement for neutral buoyancy, the needed amount of deuteration can be roughly estimated.

The experiment then involves going from a state where the specimen sinks to a floating one by deuterating the solvent. From the volume data provided by MCMC fits this seemed possible but in practice the lipid sample immersed in water was swimming in pure H_2O. This implies that the volume retrieved by fitting is too low. For our lipid specimen to swim a volume of at least $1203\,\text{Å}^3$ is required.

In this consideration the hexagonal structure was completely ignored. Furthermore, a liquid changes its properties when in a nanometer-scale confinement. According to [15] the density of water confined in a tube with a diameter of 3.6 nm, which is the diameter of the water core for DOPE at 35 °C, is smaller with a factor of 0.83 compared to bulk water. Taking into account the hexagonal structure and the confinement of water decreases the density used for ρ_{lipid} in (42). With this decreased density our DOPE sample is actually expected to swim in regular water, which is what we have observed.

Although it was not possible to conduct this experiment and use it to measure the lipid volume, the considerations made do actually confirm the volume result acquired for DOPE. While it was not possible to measure the volume of DOPE in the H_{II} phase, a lower boundary for the volume can be estimated to be around 1120 Å at 35°C, where our result is comfortably above as shown in table 3. If the volume were less than this boundary, the specimen would not swim.

3.2. 16:1PE – Host for DMPC

As mentioned earlier, DOPE will be mainly used as host, although the chain lengths may not match. DOPE has an 18:1 chain configuration, that means that it is considered compatible to saturated chains which are 16 carbon atoms long. The 1 behind the colon indicates that the chain has one double bond. For each degree of unsaturation it is assumed that the chain looses 2 C-atoms in length. The model does not account for different chain lengths, and the interactions at the lipid ends arising from it, in any way which could be a problem. Therefore, a shorter lipid, namely 16:1PE, is used as a host lipid for the chain length matching DMPC (14:0 chains) to compare the results with the ones where DOPE is used as the host.

In figure 24 the pure 16:1PE data is shown and in table 4 the results of the fit are stored. Between the first two peaks, an unexpected shoulder is visible. When compared to the data published by M. Frewein [7], it stands out that the lattice constant is lower while the curvature is higher. The first one being smaller indicates a problem with the filler molecule, the latter is an expected consequence of a smaller lattice constant. Either there was an error preparing the sample or the used tricosene was oxidized. Since a smaller lattice constant can be connected to stress in the H_{II} phase, the data for 16:1PE shown here is probably biased because of this stress.

Nevertheless, the data will still be accepted and used as host for DMPC because these two lipids have matching chain lengths and the 16:1PE-DMPC sample, which will be analyzed later, was prepared with the same filler molecule problem as its host.

Table 4: Results for 16:1PE at 35 °C. Variables above the dashed line are the lattice constant a, the width of the lipid head d_h and the area per lipid at the neutral plane A_0. Below the dashed line, some fitting parameters are shown, a complete list can be found in the appendix. Means as values and standard deviation as errors.

Symbol	$T = 35$ °C		$T = 35$ °C Results from [7]	
a / Å	74.0	$\pm\,0.5$	78.8	$\pm\,0.2$
d_h / Å	3.6	$\pm\,1.1$	4.5	$\pm\,1.7$
A_0 / Å^2	63	$\pm\,4$	60	$\pm\,7$
C_0 / Å^{-1}	-0.041	$\pm\,0.001$	-0.038	$\pm\,0.001$
V / Å^3	1064	$\pm\,12$	1042	$\pm\,8$
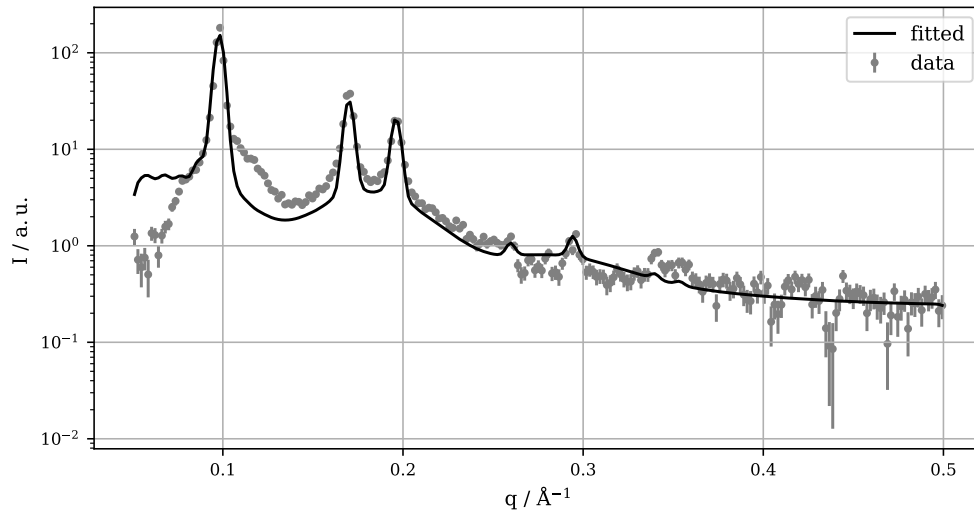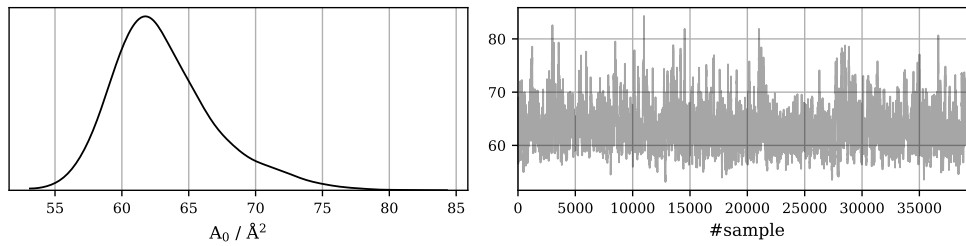
(a) SAXS data and fit.

(b) Histogram of the head area, measured at the neutral plane, resulting from the fit. On the right the trace of all accepted Monte-Carlo steps is shown.
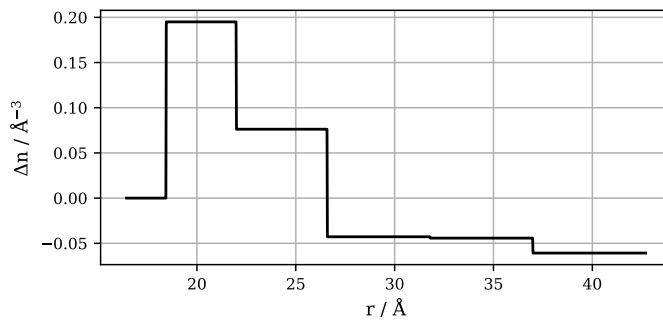
(c) Corresponding scattering length density profile. The electron density Δn is measured relative to the one of water.

Figure 24: Pure 16:1PE at 35 °C. Due to its matching chain configuration it will be used as host lipid for DMPC.

3.3. Phosphocholines

This work focuses on bilayer forming lipids with phosphocholine as their heads. In this section the results of these lipids shall be discussed.

3.3.1. DPhPC – 1,2-diphytanoyl-sn-glycero-3-phosphocholine

DPhPC is special in the sense that it incorporates very well into the H_{II} phase and thus allows high guest lipid concentrations. The original plan was therefore to use the DPhPC data to compare the global method with a previous approach, that neither includes diffuse scattering nor non-linear headgroup effects, to visualize their differences. The differences turn out to be surprisingly small and are discussed after the fit results are shown.

Figure 25 shows the results of the joint analysis of DOPE with 20 mol% and 40 mol% of DPhPC. This means that both data sets are used simultaneously to optimize the model, with volume and head-ratio as common variables, see section 2.4.
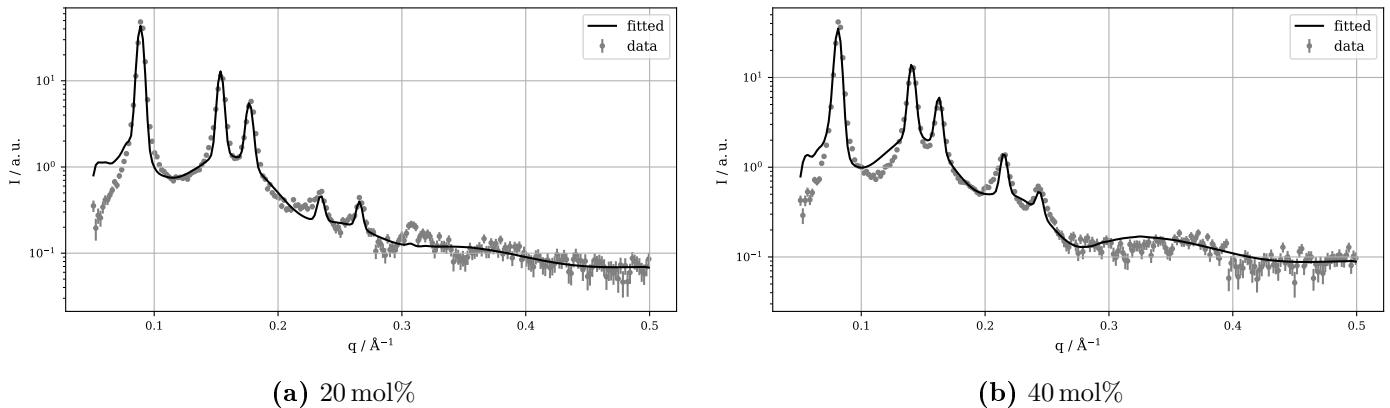

(a) 20 mol% (b) 40 mol%

Figure 25: SAXS data and fit results for DOPE mixed with DPhPC at 35 °C for different guest lipid concentrations

In figure 26 the intrinsic curvature distributions are shown. The result for 10 mol%, acquired in a regular fit of just this one data set, is plotted alongside the result of a joint analysis of the 20 mol% and 40 mol% data sets. All the given curvature results are coinciding very well. In table 5 these results are stored. It looks like the joint analysis shows lower errors on the otherwise very similar results. This can also be related to the different data sets used. The main point of this comparison is to verify that results for different amounts of added guest lipid do agree with each other because if that is the case, a single measurement at one concentration is enough to measure the intrinsic curvature of the guest lipid.

Table 5: Results for a DOPE–DPhPC sample from a joint analysis and a single data set fit. The shown variables are the lattice constant a, the head-ratio ξ, the width of the lipid head d_h, the guest intrinsic curvature $C_{0,DPhPC}$, the guest lipid volume V_{DPhPC} and the area per lipid at the neutral plane in the bilayer A_{bi}. The given percentages for some variables denote the amount of guest lipid. Values above the dashed line are calculated from fit parameter distributions. Means as values and standard deviation as errors.
* This variable is not fitted, a literature value from [18] is used

Symbol	$T = 35\,°\mathrm{C}$ Joint analysis		$T = 35\,°\mathrm{C}$ Single data set	
$a_{10\mathrm{mol\%}}$ / Å			78.9	$\pm\,0.5$
$a_{20\mathrm{mol\%}}$ / Å	81.9	$\pm\,0.5$		
$a_{40\mathrm{mol\%}}$ / Å	89.2	$\pm\,0.5$		
d_h / Å	4.7	$\pm\,0.5$	5.7	$\pm\,0.8$
$C_{0,DPhPC}$ / Å^{-1}	-0.0205	$\pm\,0.0014$	-0.020	$\pm\,0.002$
ξ	1.38	$\pm\,0.04$	1.69	$\pm\,0.19$
V_{DPhPC} / Å^3	1365	$\pm\,11$	1350	$\pm\,30$
*A_{bi} / Å^2	81.35		81.35	
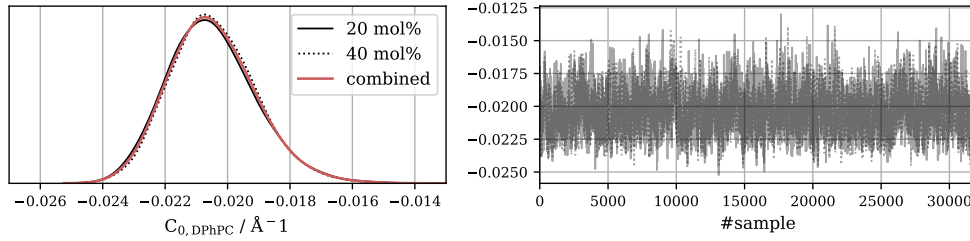
(a) Joint analysis of 20 and 40 mol%
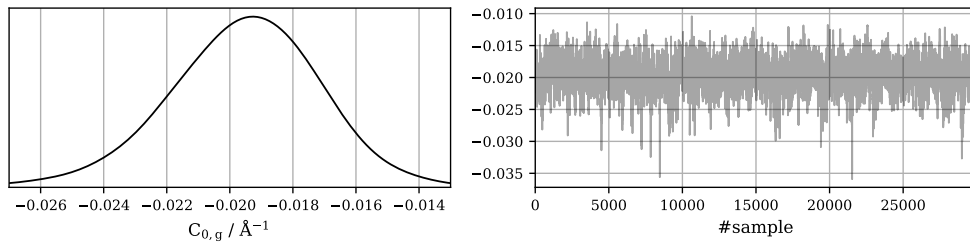
(b) 10 mol%

Figure 26: Histograms of the intrinsic curvature results for DPhPC resulting from fit at 35 °C. On the right, the trace of all accepted Monte-Carlo steps is shown.

Comparing to a non-global fitting Method

In figure 27 the linear extrapolation method introduced by B. Kollmitzer [16], see sec-

tion 1.5, is applied on DPhPC data at 35 °C and shown alongside the results of the global fitting method described in this thesis. The linear approach yields an intrinsic curvature of $\hat{C}_{0,dphpc} = (-0.020 \pm 0.003)\,\text{Å}^{-1}$. The hat denotes that it is the result from the extrapolation. Global fitting leads to an intrinsic curvature of DPhPC of $C_{0,dphpc} = (-0.0205 \pm 0.0014)\,\text{Å}^{-1}$. This is interestingly close to the result from the linear extrapolation. The deviation of the red curve in figure 27, which is equation (9) with the result of the global fit, from the curvatures calculated with Bragg peak positions arises from the non-linear headgroup interaction, for which the global model accounts, furthermore it is accounted for diffuse scattering. Because of the head-ratio being greater than one, $\xi = 1.38 \pm 0.04$ in this case, the read curve is indeed not straight but bending down.
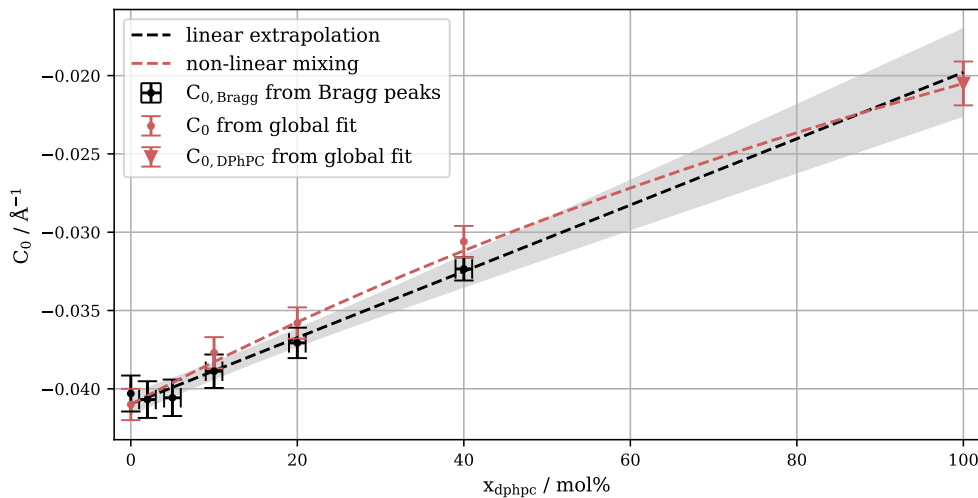

Figure 27: Linear extrapolation of the total intrinsic curvatures C_0 of mixtures of DOPE and DPhPC at different concentrations of the guest lipid x_{dphpc} at 35 °C. Every point (black colored) results from the Bragg peaks at its concentration. The resulting intrinsic curvature of DPHPC at 35 °C is $\hat{C}_{0,dphpc} = (-0.020 \pm 0.003)\,\text{Å}^{-1}$. The red dashed line results from (9) and represents the global fitting result which is $C_{0,dphpc} = (-0.0205 \pm 0.0014)\,\text{Å}^{-1}$.

The technique described in this work was expected to show lower values for the intrinsic curvature, due to the non-linear headgroup interactions, than yielded by linear extrapolation (black dashed line). Instead those two results turn out strikingly similar.
The benefits of the global fit are that for one the error is smaller and also only one measurement at one concentration is generally enough to evaluate the curvature. However, this is a special case. It is not possible for every lipid to mix that well into the H_{II} phase. The error of the linear extrapolation is therefore generally even higher than shown here.

3.3.2. DPPC – 1,2-dipalmitoyl-sn-glycero-3-phosphocholine

In figure 28 respectively table 6 the results for DPPC are shown. As expected, due to the fact that the 18:1 chain of DPPC matches in length to DOPE, the fit looks good and the

outcome is quite reasonable. Again, a joint analysis is used where all used data sets, $5\,\mathrm{mol\%}$, $10\,\mathrm{mol\%}$ and $15\,\mathrm{mol\%}$, are fitted at the same time.
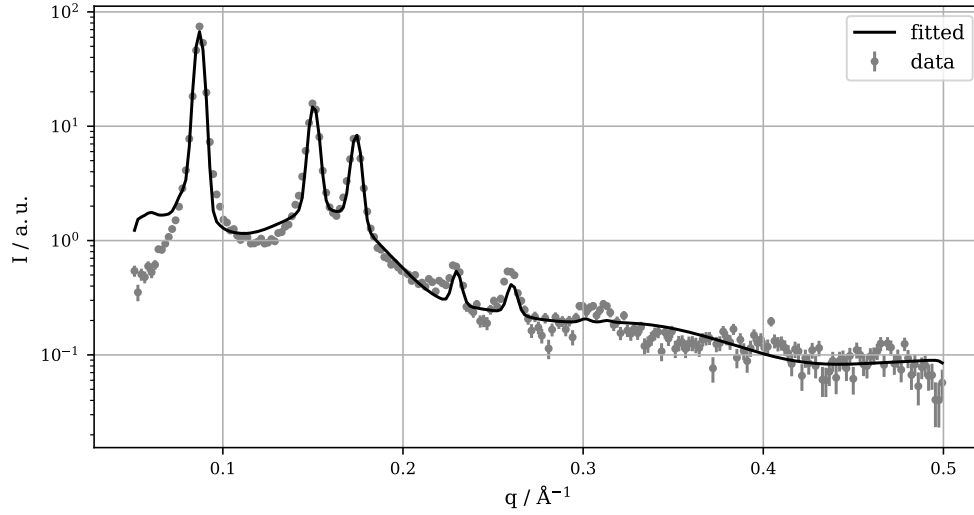
(a) SAXS data and fit of the $10\,\mathrm{mol\%}$ data.
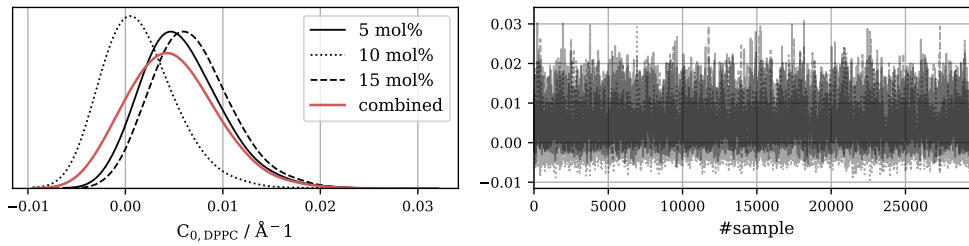
(b) Histogram of the curvature results for all concentrations. On the right the trace of all accepted Monte-Carlo steps is shown.
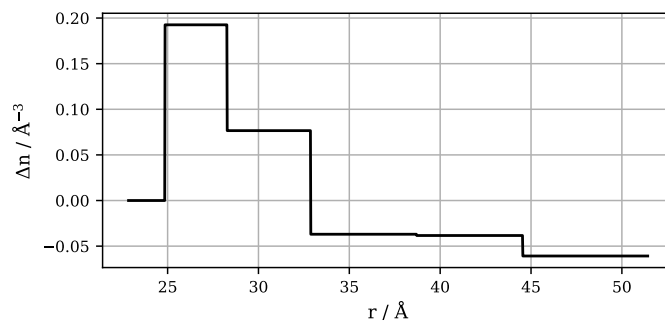
(c) Corresponding scattering length density profile. The electron density Δn is measured relative to the one of water.

Figure 28: DOPE with DPPC at $35\,^{\circ}\mathrm{C}$.

Table 6: Results for a DOPE–DPPC sample at different temperatures. The shown variables are the lattice constant a, the head-ratio ξ, the width of the lipid head d_h, the guest intrinsic curvature $C_{0,DPPC}$, the guest lipid volume V_{DPPC} and the area per lipid at the neutral plane in the bilayer A_{bi}. The given percentages for some variables denote the amount of guest lipid. Values above the dashed line are calculated from fit parameter distributions. Means as values and standard deviation as errors.
* This variable is not fitted, a literature value from [18] is used

Symbol	$T = 25\,°\mathrm{C}$		$T = 35\,°\mathrm{C}$		$T = 50\,°\mathrm{C}$	
$a_{5\mathrm{mol\%}}$ / Å	83.1	$\pm\,0.5$	80.6	$\pm\,0.5$	76.7	$\pm\,0.5$
$a_{10\mathrm{mol\%}}$ / Å	87.1	$\pm\,0.5$	83.5	$\pm\,0.5$	79.4	$\pm\,0.5$
$a_{15\mathrm{mol\%}}$ / Å	93.3	$\pm\,0.5$	89.1	$\pm\,0.5$	83.9	$\pm\,0.5$
d_h / Å	4.2	$\pm\,0.6$	3.8	$\pm\,0.6$	4.0	$\pm\,0.5$
$C_{0,DPPC}$ / Å^{-1}	0.004	$\pm\,0.005$	0.005	$\pm\,0.005$	0.003	$\pm\,0.005$
ξ	1.27	$\pm\,0.08$	1.12	$\pm\,0.07$	1.05	$\pm\,0.07$
V_{DPPC} / Å^3	1148	$\pm\,12$	1158	$\pm\,8$	1166	$\pm\,10$
*A_{bi} / Å^2	58.35		60.25		63.1	

Comparing to a non-global fitting Method

Similar to what was done with DPhPC, also here we want to compare the global fit result to the approach described by B. Kollmitzer, see section 1.5.
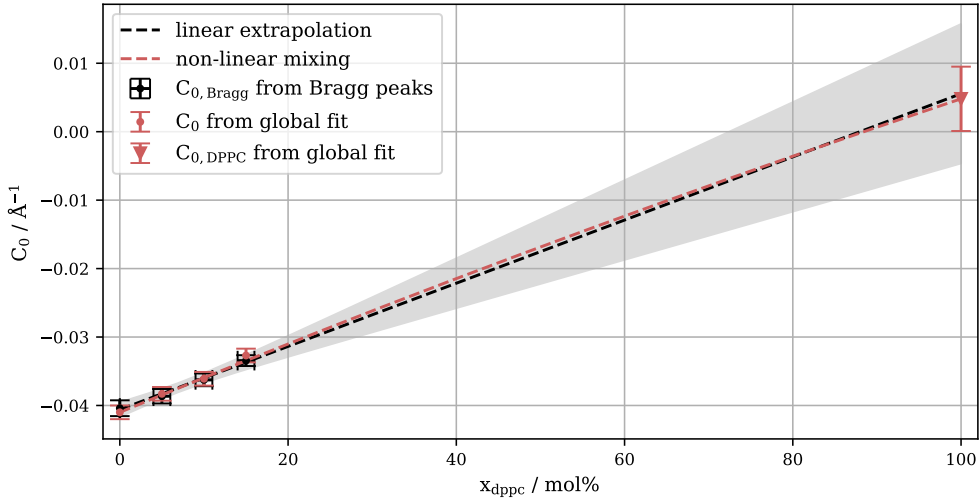
Figure 29: Comparison of joint analysis result $C_{0,DPPC}$ and linear extrapolation of the total curvatures for each concentration at $35\,°\mathrm{C}$.

Interestingly again, the curvature of the guest lipid, here DPPC, coming from the linear extrapolation, is almost the same as our result from the global model. As already mentioned, the global model has the advantage of less error and the data of one concentration is sufficient.

3.3.3. DOPC – 1,2-dioleoyl-sn-glycero-3-phosphocholine

Due to the fact that DOPC has the same chain configuration as the used host, DOPE, it is expected to work best with the model. The data is fitted using a joint analysis of 3 data sets with 5, 10 and 15 mol%. In figure 30 the fit result is compared to the 10 mol% SAXS data. Besides the very low q-range and the double-peak at around $0.3\,\text{Å}$, which are never ideal, the fit result describes the data very well.
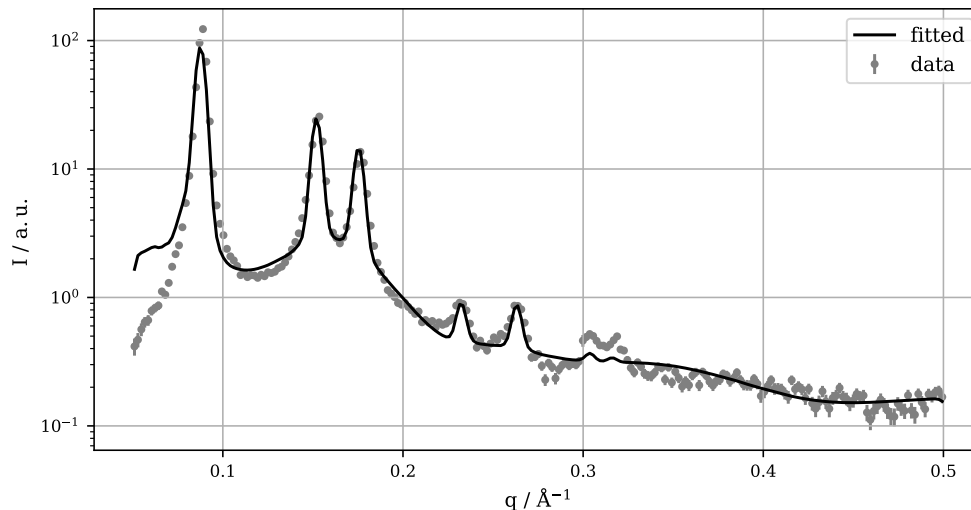
Figure 30: DOPE mixed with $10\,\text{mol}\%$ of DOPC at $35\,^\circ\text{C}$

Only some parameters optimized during a joint analysis are commonly used over different data sets, for most of them there is a distinct parameter for every set. These distinct parameters are the source for a possibly distinct curvature result for each of the different concentrations. This can be used to investigate the quality of the result. Ideally the curvature distributions resulting from different amounts of guest lipid are the same or at least similar. In figure 31 the distinct distributions of each used data set at different temperatures are compared. This is an example for a satisfying outcome, all results for all temperatures and all concentrations give a consensual result. Only the $5\,\text{mol}\%$ data at $25\,^\circ\text{C}$ is a bit too negative but still within the error. For $50\,^\circ\text{C}$ the intrinsic curvature result is slightly more negative which conforms to expectations.
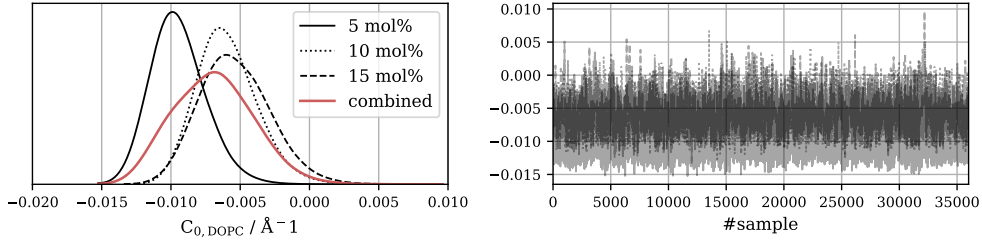
(a) Result of DOPE mixed with DOPC at 25 °C
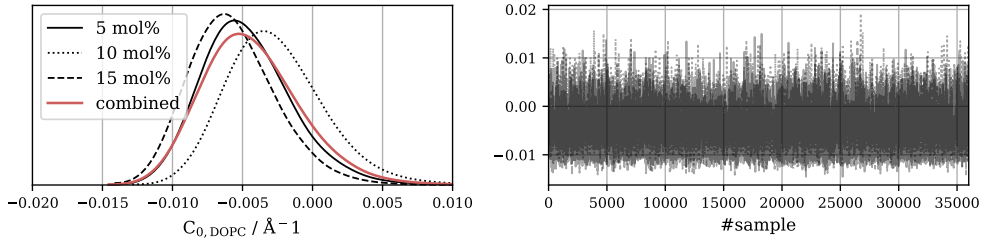
(b) Result of DOPE mixed with DOPC at 35 °C
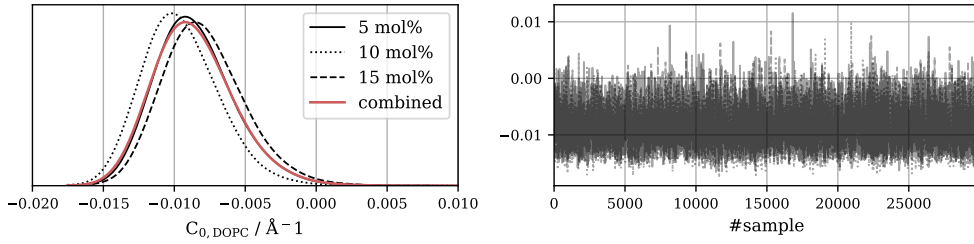
(c) Result of DOPE mixed with DOPC at 50 °C

Figure 31: Curvature distributions for DOPC

In table 7 the results for DOPC are shown. Again it is no complete list of all fitting parameters, since for all three concentration combined there are already 30 variables, but the most relevant ones are there. The head-ratio, the lipid volume for DOPC and the area per lipid in the lamellar phase are all commonly used for all concentrations. The latter one is also applied for the monolayer phase. For some variables there is a percentage information added as subscript, these are the formally mentioned distinct variables which are different for each guest concentration.

In order to get a single result for the curvature C_{DOPC} and the width of the lipid head d_h, they are calculated at every single MCMC-step available throughout all concentrations and then the mean values and standard deviations are evaluated. All shown results, despite the lattice constant, are the product of a joint analysis of data for 5, 10 and 15 mol% of guest lipid.

Table 7: Results for a DOPE–DOPC sample at different temperatures. The shown variables are the lattice constant a, the head-ratio ξ, the width of the lipid head d_h, the guest intrinsic curvature $C_{0,DOPC}$, the guest lipid volume V_{DOPC} and the area per lipid at the neutral plane in the bilayer A_{bi}. The given percentages for some variables denote the amount of guest lipid. Values above the dashed line are calculated from fit parameter distributions. Means as values and standard deviation as errors.

Symbol	$T = 25\,°\mathrm{C}$		$T = 35\,°\mathrm{C}$		$T = 50\,°\mathrm{C}$	
$a_{5\mathrm{mol\%}}$ / Å	82.0	$\pm\,0.5$	79.6	$\pm\,0.5$	75.8	$\pm\,0.5$
$a_{10\mathrm{mol\%}}$ / Å	85.5	$\pm\,0.5$	82.7	$\pm\,0.5$	78.1	$\pm\,0.5$
$a_{15\mathrm{mol\%}}$ / Å	88.1	$\pm\,0.5$	85.2	$\pm\,0.5$	80.8	$\pm\,0.5$
d_h / Å	4.9	$\pm\,0.5$	3.3	$\pm\,0.6$	4.5	$\pm\,0.6$
$C_{0,DOPC}$ / Å^{-1}	-0.007	$\pm\,0.003$	-0.004	$\pm\,0.004$	-0.009	$\pm\,0.003$
ξ	1.44	$\pm\,0.06$	1.13	$\pm\,0.08$	1.22	$\pm\,0.07$
V_{DOPC} / Å^3	1237	$\pm\,9$	1243	$\pm\,12$	1222	$\pm\,10$
A_{bi} / Å^2	68.8	$\pm\,1.1$	66.0	$\pm\,1.3$	66.8	$\pm\,1.3$

Looking at this result, it seems that DOPC has a similar temperature dependency as DOPE, although more subtle. The excellent fit results shown before and the good convergence of the fit, which gets explained in a bit, is in favor of the given values.

In figure 32 the marginal distribution of some fitting parameters are shown. They are a great tool to validate the quality of the fit. The 1-D distributions should all converge to a Gaussian function if the fit worked well, also then the errors can be estimated as the standard deviation of the distributions. The 2-D distributions give further insight into the fit result, foremost they have information about correlations between parameters which means overfitting could be exposed.
Said all that, the result for DOPC looks very promising. All parameters are converging well and no correlations are visible.
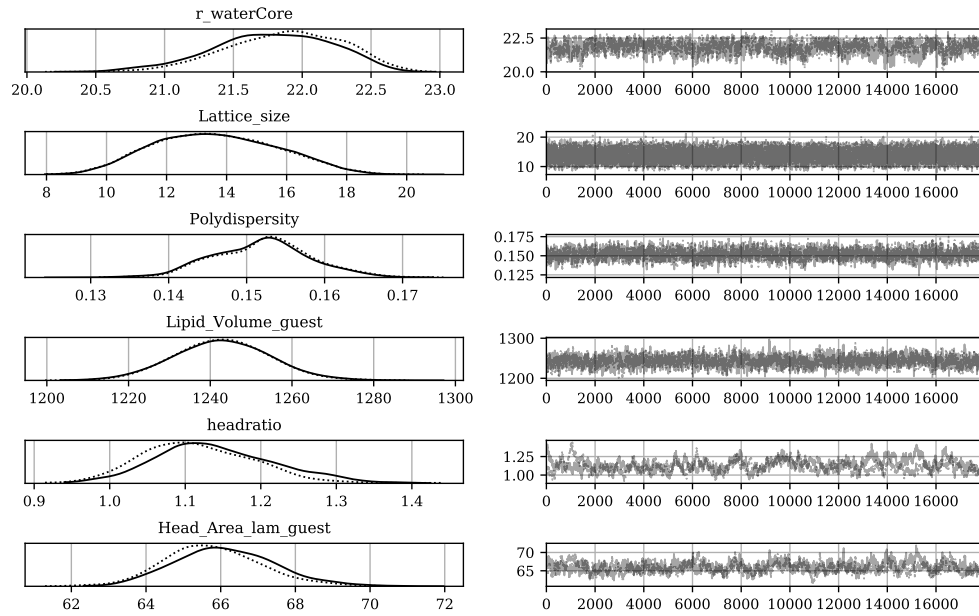
(a) Marginal distributions of some fit parameters
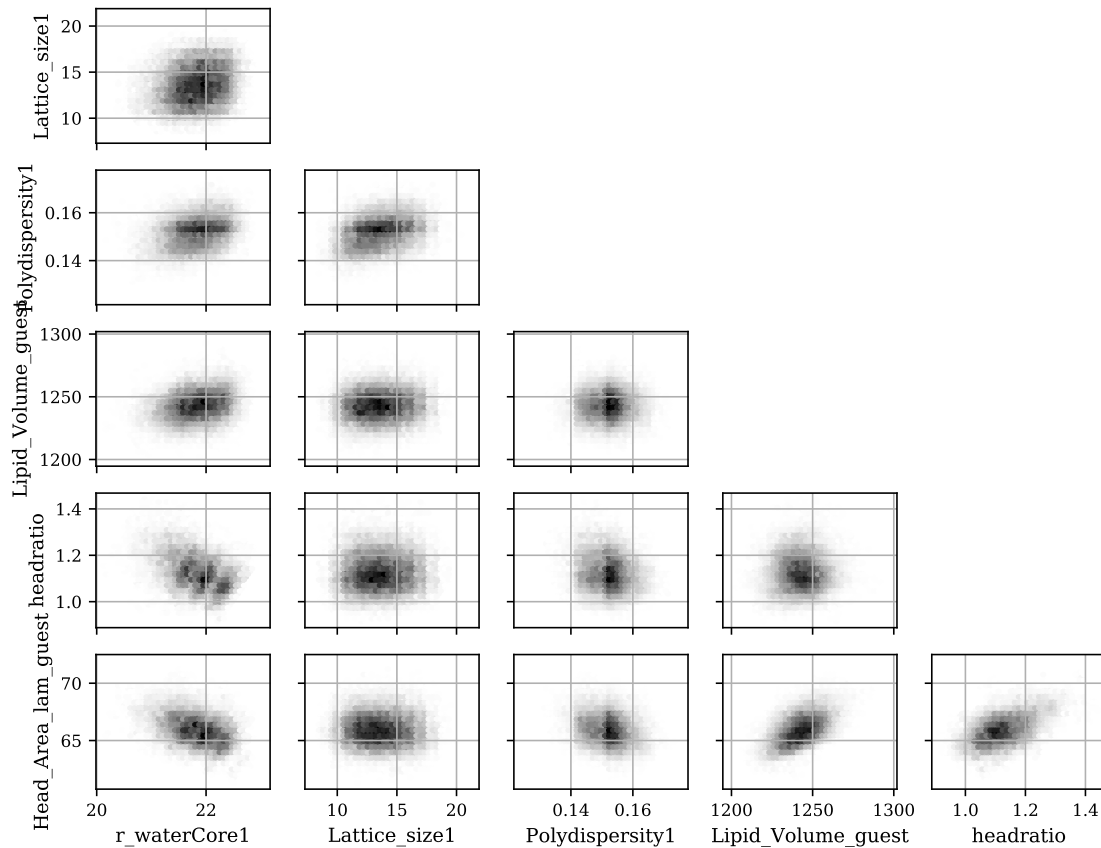
(b) Two dimensional marginal distributions of some fit parameters

Figure 32: Histograms over the MCMC-steps of an abstract of the fitting parameters for DOPE mixed with $10\,\text{mol}\%$ DOPC at $35\,^\circ\text{C}$

3.3.4. DMPC − 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMPC has a 14:0 chain configuration on both chains, meaning that its chains are built up of 14 C-atoms without any double bonds. The absence of double bonds makes the chains saturated. It is assumed that the kink of the chain introduced by a double bond between C-atoms, which is present in unsaturated chains, shortens the chain length a bit. So the chain configuration 16:1 has about the same length as 14:0 and so on. DOPE, which has two 18:1 chains, is therefore longer than DMPC, whereas 16:1PE, the name already implies 16:1 chains, is a good match. DMPC was measured with both lipids, DOPE and 16:1PE, as hosts to investigate the impact of different chain lengths on the model.

table 8 respectively figure 33 the results with the two different hosts are compared. The resulting curvatures are pretty much the same but keep in mind that the 16:1PE-DMPC data probably has an issue with its filler molecule, therefore the smaller lattice constant, also the agreement between the two concentrations is worse for this specimen. Some values in the table have percent information added as subscript because they are different for various guest lipid concentrations.

When looking at the curvature distributions it is visible that the ones with DOPE as host do agree better with each other, as seen in figure 33b, which makes this result more favorable over the one shown in figure 33a, regardless of the possible corruption of the latter sample. The fact that the result with DOPE as host seems more reliable is unexpected due to the chain length mismatch. But still, the differences between the results are small which is an indicator that the model is sufficient to use with phosphocholines even though the chain lengths do not match. Probably the impact is small here because PEs and PCs have a similar volume. Of course this could be different for other lipids such as ceramides.

Table 8: Results for DMPC with two different host lipids at 35 °C. The shown variables are the lattice constant a, the head-ratio ξ, the width of the lipid head d_h, the guest intrinsic curvature $C_{0,DMPC}$ and the guest lipid volume V_{DMPC}. The given percentages for some variables denote the amount of guest lipid. Values above the dashed line are calculated from fit parameter distributions. Means as values and standard deviation as errors.

Symbol	16:1PE − DMPC		DOPE − DMPC	
$a_{7.5\text{mol\%}}$ / Å	79.5	$\pm\,0.5$	83.3	$\pm\,0.5$
$a_{10\text{mol\%}}$ / Å	84.2	$\pm\,0.5$	85.0	$\pm\,0.5$
d_h / Å	4.3	$\pm\,0.9$	4.8	$\pm\,0.6$
$C_{0,DMPC}$ / Å^{-1}	0.006	$\pm\,0.005$	0.006	$\pm\,0.004$
ξ	1.74	$\pm\,0.14$	1.47	$\pm\,0.13$
V_{DMPC} / Å^3	1067	$\pm\,10$	1089	$\pm\,11$
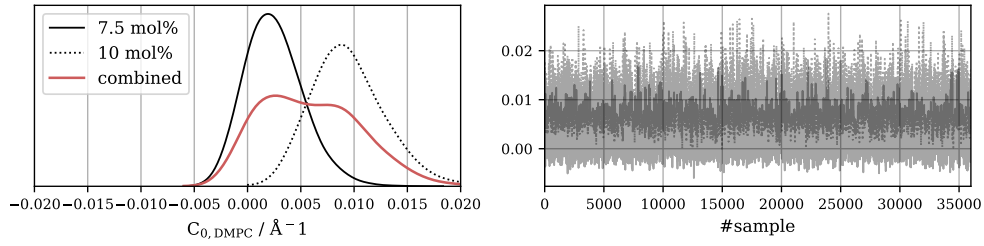
(a) DMPC with 16:1PE as the host
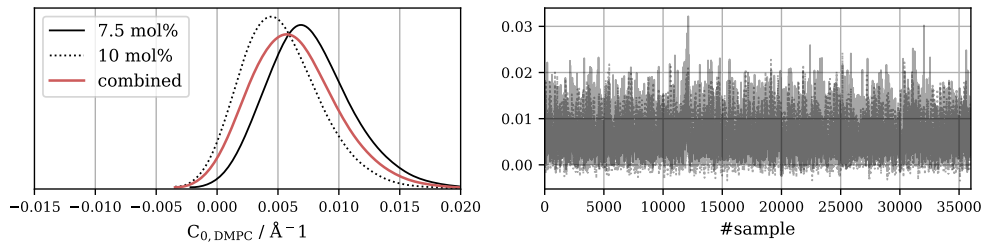
(b) DMPC with DOPE as the host

Figure 33: Resulting curvature distributions for DMPC with two different host lipids at 35 °C

In figure 34 the fit results are plotted alongside the SAXS data for both hosts to compare the overall agreement of both fits. There is hardly any difference in fit quality but the DOPE-DMPC SAXS-data looks a bit better, especially in the mid to high q-range. So in the end the DOPE-DMPC result is preferred over the one with 16:1PE but the difference is not significant. Perhaps the data quality was what tipped the scales here.
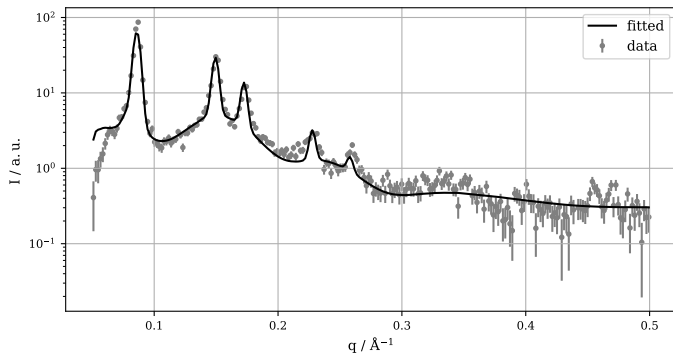
(a) $10 \, \mathrm{mol\%}$ DMPC with 16:1PE as host
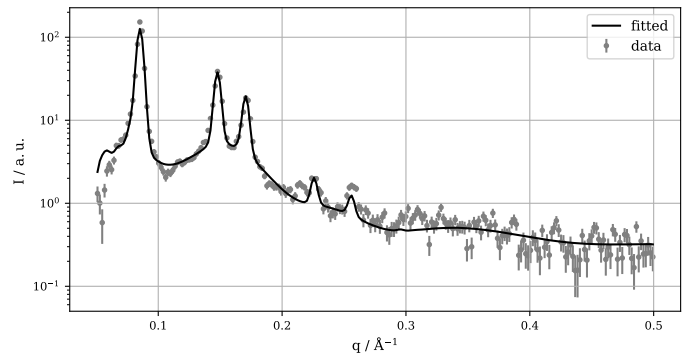
(b) $10 \, \mathrm{mol\%}$ DMPC with DOPE as host

Figure 34: Fit result for DMPC with two different host lipids at 35 °C

3.3.5. Conclusion on PC results

To conclude on the results the model produced for phosphocholines, the intrinsic curvatures and volumes for several different data sets are shown in figure 35. The x-axis in these plots

describes the number of C-atoms per lipid chain. POPC has one chain with 16 C-atoms and one with 18, which gives on average 17.5 atoms. The black dotted line represents a regression line of the saturated lipid results. DSPC was omitted for this linear regression because it was not measured at 35 degrees.

For both variables, intrinsic curvature and volume, the saturated lipids follow a linear trend depending on the number of C-atoms. DPhPC, with its additional CH_3 groups, meets the expectations by having a larger volume and a more negative curvature than saturated lipids at this chain length. Also the mono- and di-monounsaturated lipids, POPC and DOPC, show a trend in the expected directions by having a slightly lesser intrinsic curvature and a bit higher volume.

(a) Intrinsic curvatures
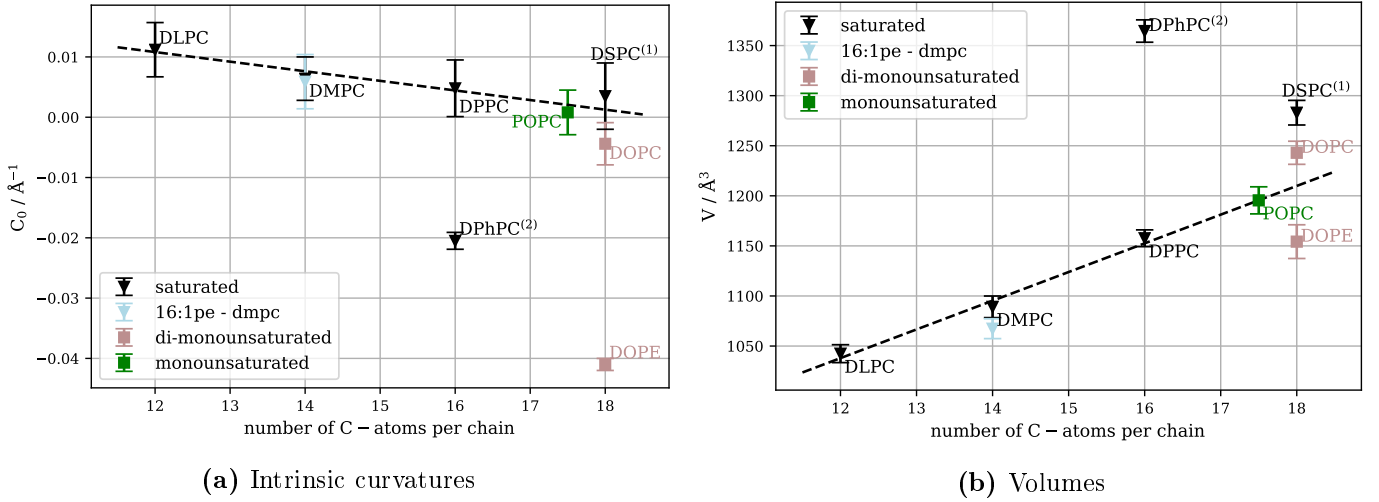
(b) Volumes

Figure 35: Comparison of results for different phospholines measured at 35 °C. (1) DSPC was measured at 30 °C. (2) DPhPC has additional CH_3 groups attached to its chains giving it more volume there which results in a more negative curvature.

The curvature follows a linear trend in respect of the number of C-atoms n_C

$$C_0(n_C) = -1.592 \cdot 10^{-3} \left[\mathring{A}^{-1} \right] n_C + 0.02991 \left[\mathring{A}^{-1} \right]$$

and the result for the linear regression for volume is:

$$V(n_C) = 28.65 \left[\mathring{A}^3 \right] n_C + 694.1 \left[\mathring{A}^3 \right]$$

For some lipids the temperature behavior of the intrinsic curvature is shown in figure 36. Within the errors all shown lipids seem to have a constant curvature with changing temperature.
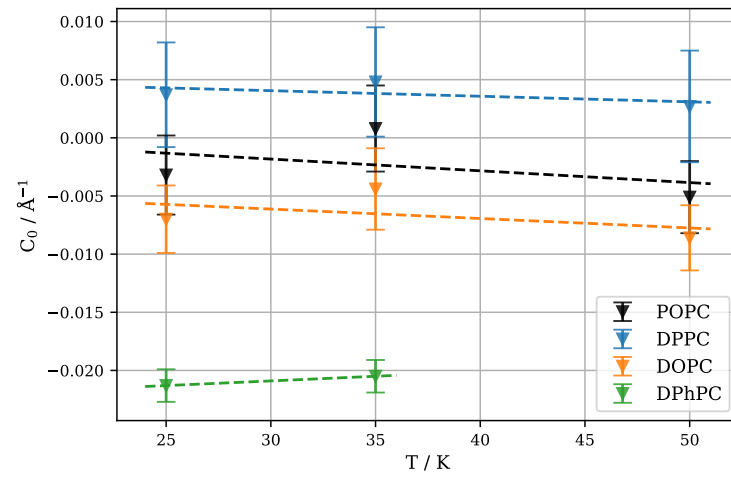
Figure 36: Summary of the temperature dependency of the intrinsic curvature results for some PCs

4. Discussion

Even if the sample preparation process (RSE) seems to work very well for producing H_{II} aggregates, there is some fluctuation in the aggregate dimensions.

A well studied and very decisive parameter is the amount of filler molecule added to the specimen [7, 21, 1, 28]. With increasing amount of this filler molecule the aggregates grow until they deviate from pure hexagonal structures [24]. Reproducibility of the lattice constant, which is basically the cylinder size of the rods, is not fully given. Oxidation of the filler molecule and deviations during the sample preparation lead to fluctuations of this variable. The outcome of the model for mixed lipid species crucially depends on the consistency of the lattice constant, therefore it is suggested to prepare a pure lipid sample every time it is used as host for a mixture. That way the host and mixed samples are prepared during the same process.

The presence of diffuse scattering definitely indicates another scattering source in addition to the hexagonal structure. Some problems arising from that are still to be solved. While the signal for lipid mixtures is very well described by the model, the low q-range for single lipid species shows significant deviations. Replacing the bilayer term here with one describing a monolayer did improve the fit quality and a monolayer is easier to argue for but still something is missing to fully describe the SAXS signal.

For mixed lipid species, the combination of mono-, bilayer and hexagonal form factors is doing a very good job in recreating the data. The amount of guest lipids incorporated in the H_{II} phase plays an important role in determining the intrinsic curvature. Since the exact distribution of the lipids, host and guest, between the H_{II}, mono-, and bilayer phases remains unknown the incorporated guest lipid concentration also stays in the dark. In contrast to former works using MCMC algorithms [8], some flexibility on the guest concentration is added by allowing to sample this parameter from a Gaussian distribution around the assumed value. This helps to compensate for concentration fluctuations and helps to remove discrepancies between results of data sets with different guest lipid amounts.

The model relies on a certain degree of similarity between the host and guest lipid. For one it does not account for differences in chain length which seemed to be fine for phosphocholines but definitely needs more investigation. Another thing is that the model works well for lipids with structurally similar heads, such as PE and PC, but complications can be expected with lipids where the head differs greatly from PE due to the simplifications made. The used head slab in lipid mixtures is just a linear combination of the host and guest instead of a more complex structure as in [8] where the head slab for mixtures is more finely divided. With lipids such as ceramides, a certain specialization of the model is probably still required.

Previously published results obtained by linear extrapolation [16] show a significant temperature dependence for partially unsaturated lipids such as POPC and DOPC. This couldn't be reproduced in this work. Further temperature investigations, which take into account the dependence on the consistency of the lattice constant mentioned at the beginning, are strongly recommended.

Conclusion and Outlook

The task of enhancing the technique to evaluate SAXS data of lipids in the H_{II} phase was performed and the procedure could be improved. The method was verified by an all-atom MD simulation for single lipid species and successfully applied to a variety of phosphocholines.

For mixed lipid species the model seems to be restricted to guest lipids with similar head structures. Chain length differences appeared to not have a large impact on the result but need more investigation. Next steps should include the application on more diverse lipids such as ceramides where probably some specialization of the model is required. To overcome difficulties in modeling more diverse guest lipids the procedure could be combined with MD simulations.

A. Complete result set DOPC

The complete result set of a joint analysis of DOPE mixed with DOPC at different temperatures is stored in table 9.

Table 9: Full results table for DOPE–DOPC data. The variables with mol% in their subscript are unique variables which means there is one for each concentration. Variables without that subscript are common variables, so they are evaluated using a combination of all data sets at this temperature.

Symbol	$T = 25\,°\text{C}$		$T = 35\,°\text{C}$		$T = 50\,°\text{C}$	
$a_{5\text{mol\%}}$ / Å	82.0	$\pm\,0.5$	79.6	$\pm\,0.5$	75.8	$\pm\,0.5$
$a_{10\text{mol\%}}$ / Å	85.5	$\pm\,0.5$	82.7	$\pm\,0.5$	78.1	$\pm\,0.5$
$a_{15\text{mol\%}}$ / Å	88.1	$\pm\,0.5$	85.2	$\pm\,0.5$	80.8	$\pm\,0.5$
d_h / Å	4.9	$\pm\,0.5$	3.3	$\pm\,0.6$	4.5	$\pm\,0.6$
$C_{0,DOPC}$ / Å^{-1}	-0.007	$\pm\,0.003$	-0.004	$\pm\,0.004$	-0.009	$\pm\,0.003$
ξ	1.44	$\pm\,0.06$	1.13	$\pm\,0.08$	1.22	$\pm\,0.07$
V_{DOPC} / Å^3	1237	$\pm\,9$	1243	$\pm\,12$	1222	$\pm\,10$
A_{bi} / Å^2	68.8	$\pm\,1.1$	66.0	$\pm\,1.3$	66.8	$\pm\,1.3$
$x_{5\text{mol\%}}$ / mol%	5.1	$\pm\,0.5$	5.0	$\pm\,0.5$	5.0	$\pm\,0.5$
$x_{10\text{mol\%}}$ / mol%	10.4	$\pm\,0.9$	9.8	$\pm\,1.0$	10.0	$\pm\,1.0$
$x_{15\text{mol\%}}$ / mol%	14.1	$\pm\,1.5$	15.1	$\pm\,1.5$	14.6	$\pm\,1.5$
$\sigma_{fluc,5\text{mol\%}}$	0.157	$\pm\,0.009$	0.153	$\pm\,0.008$	0.158	$\pm\,0.007$
$\sigma_{fluc,10\text{mol\%}}$	0.147	$\pm\,0.005$	0.152	$\pm\,0.006$	0.168	$\pm\,0.005$
$\sigma_{fluc,15\text{mol\%}}$	0.151	$\pm\,0.005$	0.155	$\pm\,0.006$	0.163	$\pm\,0.005$
$\Delta_{5\text{mol\%}}$	9	$\pm\,3$	7	$\pm\,2$	7.0	$\pm\,1.7$
$\Delta_{10\text{mol\%}}$	11.0	$\pm\,1.7$	9	$\pm\,2$	7.1	$\pm\,1.5$
$\Delta_{15\text{mol\%}}$	10	$\pm\,2$	13	$\pm\,2$	8.5	$\pm\,1.5$
$R_{w,5\text{mol\%}}$ / Å	19.7	$\pm\,0.5$	19.7	$\pm\,0.5$	18.1	$\pm\,0.5$
$R_{w,10\text{mol\%}}$ / Å	22.0	$\pm\,0.4$	21.8	$\pm\,0.4$	18.7	$\pm\,0.4$
$R_{w,15\text{mol\%}}$ / Å	23.4	$\pm\,0.4$	23.3	$\pm\,0.4$	20.2	$\pm\,0.4$
$c_{bi,5\text{mol\%}}$	2.5	$\pm\,0.3$	2.4	$\pm\,0.3$	2.4	$\pm\,0.3$
$c_{bi,10\text{mol\%}}$	2.5	$\pm\,0.2$	2.3	$\pm\,0.2$	2.3	$\pm\,0.2$
$c_{bi,15\text{mol\%}}$	2.7	$\pm\,0.2$	2.1	$\pm\,0.2$	2.1	$\pm\,0.2$
$c_{mono,5\text{mol\%}}$	3.4	$\pm\,0.4$	2.6	$\pm\,0.3$	2.4	$\pm\,0.5$
$c_{mono,10\text{mol\%}}$	2.7	$\pm\,0.4$	2.7	$\pm\,0.3$	2.8	$\pm\,0.3$
$c_{mono,15\text{mol\%}}$	2.8	$\pm\,0.3$	2.6	$\pm\,0.2$	3.2	$\pm\,0.3$
$n_{5\text{mol\%}}$	21	$\pm\,4$	15	$\pm\,2$	18.7	$\pm\,1.6$
$n_{10\text{mol\%}}$	24.8	$\pm\,0.7$	14	$\pm\,2$	19.1	$\pm\,1.4$
$n_{15\text{mol\%}}$	12.4	$\pm\,1.7$	21	$\pm\,4$	17.5	$\pm\,1.2$
$\Gamma_{5\text{mol\%}}$	3.5	$\pm\,0.3$	3.5	$\pm\,0.3$	4.1	$\pm\,0.3$
$\Gamma_{10\text{mol\%}}$	5.9	$\pm\,0.4$	5.7	$\pm\,0.5$	7.4	$\pm\,0.5$
$\Gamma_{15\text{mol\%}}$	7.0	$\pm\,0.5$	7.3	$\pm\,0.6$	8.1	$\pm\,0.5$
$I_{add,5\text{mol\%}}$ / a.u.	0.062	$\pm\,0.003$	0.068	$\pm\,0.004$	0.090	$\pm\,0.003$
$I_{add,10\text{mol\%}}$ / a.u.	0.130	$\pm\,0.004$	0.139	$\pm\,0.005$	0.178	$\pm\,0.004$
$I_{add,15\text{mol\%}}$ / a.u.	0.153	$\pm\,0.004$	0.155	$\pm\,0.005$	0.168	$\pm\,0.005$

References

[1] S. H. Alley et al. "X-ray diffraction measurement of the monolayer spontaneous curvature of dioleoylphosphatidylglycerol". In: *Chem Phys Lipids* 154.1 (2008), pp. 64–67. ISSN: 0009-3084. DOI: `10.1016/j.chemphyslip.2008.03.007`. URL: `18405663`.

[2] D. J. Bonthuis. *Inverse hexagonal phases - simulation setup*. unpublished (2019).

[3] M. F. Brown. "Soft Matter in Lipid-Protein Interactions". In: *Annu Rev Biophys* 46 (2017), pp. 379–410. ISSN: 1936122X. DOI: `10.1146/annurev-biophys-070816-033843`.

[4] J. T. Buboltz and G. W. Feigenson. "A novel strategy for the preparation of liposomes: rapid solvent exchange". In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1417.2 (1999), pp. 232–245.

[5] E. (Edgar181). *(Z)-9-tricosene biosynthesis*. `https://commons.wikimedia.org/wiki/File:(Z)-9-tricosene_biosynthesis.svg`. Accessed: 2020-10-04.

[6] N. Freiberger and O. Glatter. "Small-angle scattering from hexagonal liquid crystals". In: *J Phys Chem B* 110.30 (2006), pp. 14719–14727. ISSN: 1520-5207. DOI: `10.1021/jp0559332`. URL: `16869579`.

[7] M. P. K. Frewein, M. Rumetshofer, and G. Pabst. "Global small-angle scattering data analysis of inverted hexagonal phases". In: *J Appl Crystallogr* 52.Pt 2 (2019), pp. 403–414. ISSN: 0021-8898. DOI: `10.1107/S1600576719002760`.

[8] M. P. K. Frewein. "Global X-ray scattering data analysis of inverted hexagonal phases to measure intrinsic lipid curvatures". Master Thesis. Graz, Austria: Graz University of Technology, 2018.

[9] M. Frewein et al. "Lateral pressure-mediated protein partitioning into liquid-ordered/liquid-disordered domains". In: *Soft Matter* 12.13 (2016), pp. 3189–3195. DOI: `10.1039/C6SM00042H`.

[10] K. M. Hallinen, S. Tristram-Nagle, and J. F. Nagle. "Volumetric stability of lipid bilayers". In: *Phys. Chem. Chem. Phys.* 14.44 (2012), pp. 15452–15457. ISSN: 1463-9084. DOI: `10.1039/C2CP42595E`.

[11] W. K. Hastings. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". In: *Biometrika* 57.1 (1970), p. 97. ISSN: 00063444. DOI: `10.2307/2334940`.

[12] W. Helfrich. "Elastic properties of lipid bilayers: theory and possible experiments". In: *Z Naturforsch C* 28.11 (1973), pp. 693–703.

[13] J. N. Israelachvili. *Intermolecular and surface forces*. 3rd ed. Burlington, MA: Academic Press, 2011. URL: `http%3A//www.worldcat.org/oclc/706803091`.

[14] J. B. Klauda et al. "Simulation-based methods for interpreting x-ray data from lipid bilayers". In: *Biophys J* 90.8 (2006), pp. 2796–2807. ISSN: 0006-3495.

[15] A. W. Knight et al. "Water properties under nano-scale confinement". In: *Sci Rep* 9.1 (2019), p. 8246. DOI: `10.1038/s41598-019-44651-z`.

[16] B. Kollmitzer et al. "Monolayer spontaneous curvature of raft-forming membrane lipids". In: *Soft Matter* 9.45 (2013), pp. 10877–10884. URL: 24672578.

[17] M. M. Kozlov and M. Winterhalter. "Elastic-Moduli for Strongly Curved Monolayers - Position of the Neutral Surface". In: *J Phys II* 1.9 (1991), pp. 1077–1084. URL: ISI:A1991GC95400006.

[18] N. Kučerka et al. "Lipid bilayer structure determined by the simultaneous analysis of neutron and X-ray scattering data". In: *Biophys J* 95.5 (2008), pp. 2356–2367. ISSN: 0006-3495.

[19] N. Kučerka et al. "Molecular structures of fluid phosphatidylethanolamine bilayers obtained from simulation-to-experiment comparisons and experimental scattering density profiles". In: *J Phys Chem B* 119.5 (2015), pp. 1947–1956. ISSN: 1520-5207. DOI: 10.1021/jp511159q. URL: https://pubs.acs.org/doi/pdf/10.1021/jp511159q.

[20] A. P. Lipids. *Structure, Product description*. https://avantilipids.com. Accessed: 2020-10-04.

[21] K. Lohner et al. "Squalene promotes the formation of non-bilayer structures in phospholipid model membranes". In: *Biochim Biophys Acta* 1152.1 (1993), pp. 69–77. ISSN: 0006-3002.

[22] O. G. Mouritsen. *Life - as a matter of fat: The emerging science of lipidomics*. Berlin: Springer, 2005.

[23] O. G. Mouritsen. "Lipids, curvature, and nano-medicine". In: *European Journal of Lipid Science and Technology* 113.10 (2011), pp. 1174–1187.

[24] N. Negahbani. *Optimization of inverted hexagonal lipid phase preparation by rapid solvent exchange. Master's thesis, TU Graz*. 2016.

[25] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck. "Probabilistic programming in Python using PyMC3". In: *PeerJ Computer Science* 2 (2016), e55. DOI: 10.7287/peerj.preprints.1686v1.

[26] G. C. Shearman et al. "Inverse lyotropic phases of lipids and membrane curvature". In: *J Phys: Condes Matter* 18.28 (2006), S1105–S1124. URL: ISI:000238613000014.

[27] U. von Toussaint. "Bayesian inference in physics". In: *Rev. Mod. Phys.* 83.3 (2011), pp. 943–999. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.83.943.

[28] H. P. Vacklin et al. "The bending elasticity of 1-monoolein upon relief of packing stress". In: *Langmuir* 16.10 (2000), pp. 4741–4748. URL: ISI:000087006200050.

[29] M. C. Wiener et al. "Specific volumes of lipids in fully hydrated bilayer dispersions". In: *Biochim Biophys Acta* 938.2 (1988), pp. 135–142. ISSN: 00052736. DOI: 10.1016/0005-2736(88)90153-8.

[30] P. Ziherl. *Spontaneous curvature of lipid-mixture monolayers*. unpublished.