Global small-angle scattering data analysis of inverted hexagonal phases

Moritz P. K. Frewein,[a,b] Michael Rumetshofer[c] and Georg Pabst[a,b]*

[a]University of Graz, Institute of Molecular Biosciences, Biophysics Division, NAWI Graz, 8010 Graz, Austria, [b]BioTechMed Graz, 8010 Graz, Austria, and [c]Graz University of Technology, Institute of Theoretical Physics and Computational Physics, NAWI Graz, 8010 Graz, Austria. *Correspondence e-mail: georg.pabst@uni-graz.at

A global analysis model has been developed for randomly oriented, fully hydrated, inverted hexagonal (H_{II}) phases formed by many amphiphiles in aqueous solution, including membrane lipids. The model is based on a structure factor for hexagonally packed rods and a compositional model for the scattering length density, enabling also the analysis of positionally weakly correlated H_{II} phases. Bayesian probability theory was used for optimization of the adjustable parameters, which allows parameter correlations to be retrieved in much more detail than standard analysis techniques and thereby enables a realistic error analysis. The model was applied to different phosphatidylethanolamines, including previously unreported H_{II} data for diC14:0 and diC16:1 phosphatidylethanolamine. The extracted structural features include intrinsic lipid curvature, hydrocarbon chain length and area per lipid at the position of the neutral plane.

1. Introduction

Elastic small-angle scattering (SAS) techniques are unrivaled for providing detailed structural insight into aggregates formed by amphiphiles in aqueous solutions (Glatter, 2018). In the field of membrane biophysics, significant efforts have been devoted to the development of SAS analysis methods for the biologically most relevant fluid lamellar phases, including domain-forming lipid mixtures and asymmetric lipid bilayers (Heberle & Pabst, 2017). In contrast, non-lamellar phases such as the inverted hexagonal (H_{II}) phase are less commonly found for membrane lipids under physiological conditions but are of significant biotechnological interest, for example, for gene transfection (Koltover *et al.*, 1998) or drug delivery systems (Yaghmur & Glatter, 2009). H_{II} phases are also highly amenable systems for deriving intrinsic lipid curvatures by small-angle X-ray scattering (SAXS) (Leikin *et al.*, 1996; Di Gregorio & Mariani, 2005; Kollmitzer *et al.*, 2013; Chen *et al.*, 2015), which is the main focus of the present contribution. The intrinsic lipid curvature C_0 is given by the negative inverse of the curvature radius, $-1/R_0$, of an unstressed monolayer at the position of the neutral plane, which corresponds to the location where molecular bending and stretching modes are decoupled (Kozlov & Winterhalter, 1991). Major interest in obtaining reliable C_0 values originates from its contribution to the stored elastic energy strain in planar bilayers (Marsh, 2006), transmembrane protein function (Dan & Safran, 1998; Frewein *et al.*, 2016) and overall membrane shape (Frolov *et al.*, 2011).

Structural details of H_{II} phases have been successfully derived using electron density map reconstruction based on
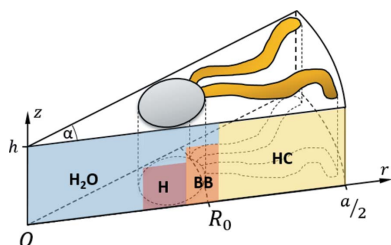
Bragg peak scattering only (Tate & Gruner, 1989, 1992; Rand *et al.*, 1990; Harper *et al.*, 2001). However, for highly swollen H_{II} phases or at elevated temperatures the number of observed Bragg peaks may become insufficient for a reliable analysis. This may be particularly the case for mixtures of cone-shaped (H_{II}-phase-forming) and cylindrically shaped (lamellar-phase-forming) or inverted cone-shaped (spherical-micelle-forming) lipids. Such mixtures are typically used for determining C_0 for non-H_{II}-phase-forming lipids [see *e.g.* Kollmitzer *et al.* (2013)]. In this case global analysis techniques which take into account both Bragg peaks and diffuse scattering become advantageous, as demonstrated previously also for lamellar phases (Pabst *et al.*, 2000).

Global analysis techniques have been reported previously for H_I phases, *i.e.* oil-in-water-type hexagonal aggregates (Freiberger & Glatter, 2006; Sundblom *et al.*, 2009). The specific need for developing a dedicated model for H_{II} phases comes from the observation that unoriented H_{II} phases contain previously unreported additional diffuse scattering originating most probably from packing defects. We have evaluated our global H_{II} model for phosphatidylethanolamines with differing hydrocarbon chain composition and as a function of temperature using Bayesian probability theory to increase the robustness of analysis. This method significantly increased the obtained information content compared with our previous analysis (Kollmitzer *et al.*, 2013) and allowed us to derive details about the structure, for example the lipid headgroup area, hydrocarbon chain length and molecular shape to name but a few.

2. Experimental methods

2.1. Sample preparation

Dioleoyl phosphatidylethanolamine (DOPE, diC18:1 PE), palmitoyl oleoyl phosphatidylethanolamine (POPE, C16:0-18:1 PE), dimyristoyl phosphatidylethanolamine (DMPE, diC14:0 PE) and dipalmitoleoyl phosphatidylethanolamine (diC16:1 PE) were purchased in the form of powder from Avanti Polar Lipids (Alabaster, AL, USA). *cis*-9-Tricosene was obtained from Sigma–Aldrich (Vienna, Austria). All lipids were used without any further purification. Note that dipalmitoleoyl phosphatidylethanolamine is deliberately abbreviated to diC16:1 PE in order not to be confused with dipalmitoyl phosphatidylethanolamine (diC16:0 PE).

Fully hydrated H_{II} phases were prepared using rapid solvent exchange (RSE) (Buboltz & Feigenson, 1999) as detailed previously (Leber *et al.*, 2018). In brief, stock solutions of lipids (10 mg ml^{-1}) and tricosene (5 mg ml^{-1}) were first prepared by dissolving both compounds in chloroform/methanol (9:1 *v/v*). Ultra-pure water (18 MΩ cm^{-2}) was filled into test tubes and equilibrated at 333–342 K using an incubator. Lipid and tricosene stock solutions were added to the test tubes containing preheated water (organic solvent/water ratio = 2.55) and then quickly mounted onto the RSE apparatus, described by Rieder *et al.* (2015). The organic solvent quickly evaporated under the following settings: temperature:

338 K; vortex speed: 600 r min^{-1}; argon flow: 60 ml min^{-1}; and final vacuum pressure: 400–500 mbar (1 mbar = 100 Pa). The full procedure was performed for 5 min, yielding a lipid pellet at the bottom of the test tube in excess water. All samples contained 12 wt% tricosene. Tricosene inserts preferentially into the interstical space between the rods in H_{II} phases, effectively reducing packing frustration as verified previously (Alley *et al.*, 2008; Kollmitzer *et al.*, 2013). Unstressed H_{II} phases are required for C_0 determination (Kozlov & Winterhalter, 1991).

2.2. Scattering experiments

Small-angle X-ray scattering (SAXS) experiments were performed via a SAXSpace compact camera (Anton Paar, Graz, Austria) equipped with an Eiger R 1 M detector system (Dectris, Baden-Daettwil, Switzerland) and a 30 W-Genix 3D microfocus X-ray generator (Xenocs, Sassenage, France), supplying Cu $K\alpha$ (λ = 1.54 Å) radiation with a circular spot size of the beam of ~300 μm on the detector. Samples were taken up in paste cells (Anton Paar) and equilibrated at each measured temperature for 10 min using a Peltier controlled sample stage (TC 150, Anton Paar). The total exposure time was 32 min (four frames of 8 min), with the sample-to-detector distance set to 308 mm. Data reduction, including sectorial data integration and corrections for sample transmission and background scattering, was performed using the program *SAXSanalyis* (Anton Paar).

3. Model

3.1. General aspects for H_{II} phases

We initially tested the applicability of a previously reported model-free approach (Freiberger & Glatter, 2006). However, although perfect fits to the experimental data were obtained, the corresponding pair distance distribution functions contained significantly negative values upon approaching the maximum particle size (D_{max}), which is not physically relevant (O. Glatter, personal communication). This encouraged us to proceed with data modeling.

To do so, we considered a bundle of hexagonal prisms consisting of a water core coated by lipids (Fig. 1). Its scattering intensity is characterized by the form factor of a single prism $F(\mathbf{q})$ and the structure factor of the whole bundle $S(\mathbf{q})$, where $\mathbf{q}$ is the scattering vector. Assuming that the prisms are long as compared to their diameters allows us to decouple form and structure factors (Freiberger & Glatter, 2006):

$$I(\mathbf{q}) \propto |F(\mathbf{q})|^2 S(\mathbf{q}). \qquad (1)$$

The structure factor of a two-dimensional lattice of infinitely long hexagonal prisms, averaged over all in-plane vectors, is given by (Oster & Riley, 1952; Marchal & Demé, 2003; Freiberger & Glatter, 2006)

$$S(q, \theta \mid n, \Delta) = 1 + \frac{1}{N_{\text{hex}}(n)} \exp(-q^2 \Delta)$$
$$\times \sum_{j \neq k}^{N_{\text{hex}}(n)} J_0(q|\mathbf{R}_j - \mathbf{R}_k| \sin \theta), \quad (2)$$

where θ is the angle between the scattering vector and the axis (z) normal to the hexagons, $N_{\text{hex}} = 1 + 3n(n + 1)$ is the total number of unit cells for n rings (Fig. 1), and J_0 is the zero-order Bessel function of the first kind. The $\exp(-q^2 \Delta)$ term is well known as the Debye–Waller factor, where Δ is the lateral mean-square displacement of the rotation axes of the unit cells around their mean positions $\mathbf{R}_j$. For the sake of readability we give the parameter dependencies after the vertical line in each equation, i.e. for equation (2) n and Δ. Analogously to Freiberger & Glatter (2006) we also considered a polydispersity of N_{hex}, which yields a smooth structure factor. However, since this affects only low q values and not significantly the final quality of the result, this was omitted in order to reduce overall computation times. An alternative structure factor, based on the positioning of peaks with flexible shapes on hexagonal lattice points, has been reported previously (Förster et al., 2005; Sundblom et al., 2009). However, its application involves a significantly higher number of adjustable parameters, which leads us to disregard this option.

The form factor for a hexagonal prism of length L is given by (Freiberger & Glatter, 2006)

$$F(q, \theta \mid \boldsymbol{\rho}) = f(q, \theta) \int \rho(r, \varphi) J_0(qr \sin \theta) r \, dr \, d\varphi$$
$$= f(q, \theta)[F_{\text{lipid}}(q, \theta \mid \boldsymbol{\rho}) + F_{\text{inter}}(q, \theta \mid \boldsymbol{\rho})] \quad (3)$$

with $f(q, \theta) = 4\pi \sin[(L/2)q \cos \theta]/(q \cos \theta)$, where $\rho(r, \varphi)$ is the in-plane scattering length density (SLD). Here, $\boldsymbol{\rho}$ denotes all parameters describing the SLD $\rho(r, \varphi)$. For evaluation, which because of symmetry is performed over 1/12 of the area of the hexagon, the form factor is split into a core–shell cylindrical part F_{lipid}, which accounts for the phospholipid only, and F_{inter}, which accounts for the interstitial space, often taken up by pure hydrophobic filler molecules (here, tricosene). We also found that the length of the cylinders does not affect F significantly for $L \geq 2500$ Å for cylinder radii between 35 and 45 Å, as occurring in the present paper. To shorten the computational times, we therefore fixed $L = 2500$ Å for all our further calculations.

The core–shell cylindrical part can be evaluated analytically (Székely et al., 2010):

$$F_{\text{lipid}}(q, \theta \mid \boldsymbol{\rho}) = \int_0^{a/2} r \Delta\rho(r) J_0(qr \sin \theta) \, dr$$
$$= \frac{1}{q \sin \theta} \left[\Delta\rho_M r_M J_1(qr_M \sin \theta) \right.$$
$$\left. + \sum_{k=1}^{M-1} (\Delta\rho_k - \Delta\rho_{k+1}) r_k J_1(qr_k \sin \theta) \right], \quad (4)$$

where M is the total number of shells, r_k are the shell radii, $\Delta\rho$ is the SLD relative to water ($\Delta\rho = \rho - \rho_W$; $\rho_W = 0.33$ Å^{-3} in the case of X-rays) and J_1 is the first-order Bessel function of the first kind.

Molecular fluctuations cause a smearing of the sharp boundaries between the individual slabs. Analogously to the approach adopted by Franks et al. (1982), these were taken into account by translating all shell boundaries $\{r_k\}$ by the distance x, whose value was assumed to be distributed by a Gaussian $\mathcal{N}(x \mid \mu, \sigma_{\text{fluc}}^2)$ of mean μ and variance σ_{fluc}^2:

$$F_{\text{lipid,fluc}}(q, \theta \mid \boldsymbol{\rho}, \sigma_{\text{fluc}})$$
$$= \int dx \, \mathcal{N}(x \mid \mu = 0, \sigma_{\text{fluc}}^2) F_{\text{lipid}}[q, \theta \mid \boldsymbol{\rho}'(x)]. \quad (5)$$

Here, $\mu = 0$ and $\boldsymbol{\rho}'(x)$ denotes the SLD including the radial shift x.

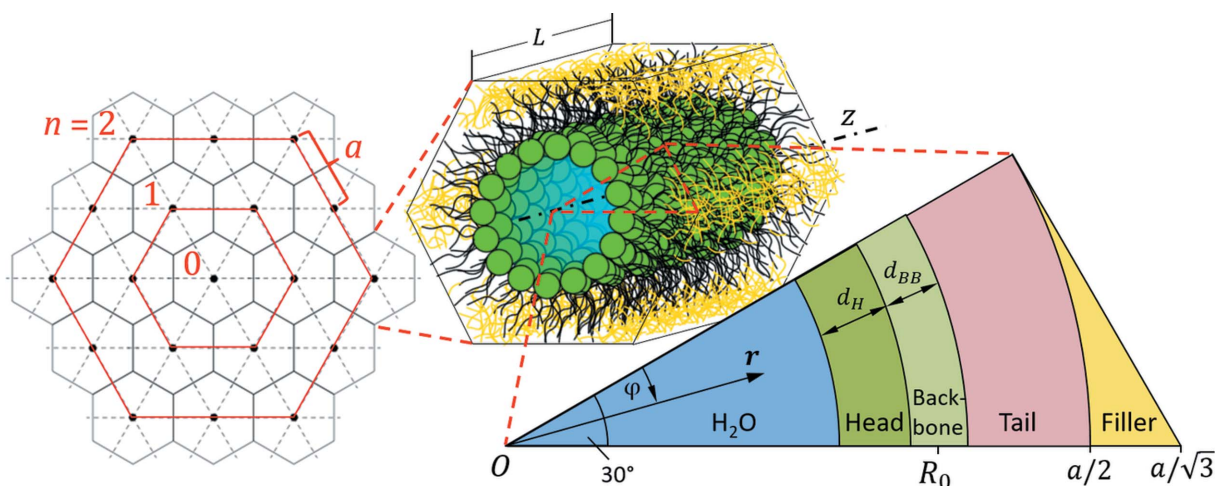The form factor of the interstices

Figure 1
Scheme of the H$_{\text{II}}$ phase model. The hexagonal lattice (left side) is defined by its lattice parameter a and the number of rings (lattice order) n. Its unit cell, shown in the center, is a regular hexagonal prism of length L and consists of a cylindrical water core, surrounded by lipids with their heads pointing toward the central water channel and a filler molecule occupying the interstices. We denote the axis of rotation by z. The unit cell is subdivided into areas of different SLD which depend on the molecular composition (see also Fig. 3). R_0 denotes the position of the neutral plane at the center of the lipid backbone.

$$F_{\text{inter}}(q, \theta \mid \boldsymbol{\rho}) = \frac{6}{\pi} \int\limits_{0}^{\pi/6} d\varphi \int\limits_{a/2}^{a/(2\cos\varphi)} r\Delta\rho_{\text{inter}} J_0(qr\sin\theta)\, dr \quad (6)$$

needs to be evaluated numerically, but remains constant for a given lattice constant a and SLD $\Delta\rho_{\text{inter}}$. However, since a can be determined accurately from Bragg peak positions, F_{inter} needs to be calculated only once for each scattering pattern.

H_{II} phases are well known to change their phase from '+' to '−' between the 10 and 11 reflections (Turner & Gruner, 1992), which brings about a minimum in $|F_{\text{lipid}}|^2$ between the two peaks.

All our present experimental data, as well as those previously reported (Kollmitzer *et al.*, 2013; Leber *et al.*, 2018), exhibited significant diffuse scattering between reflections 10 and 11 (Fig. 2). That is, experimental data from unoriented H_{II} phases show no form factor minimum in this q range. The additional scattering may also explain the failure of the model-free analysis approach discussed above and possibly arises from packing defects between hexagonal bundles, for example at grain boundaries. However, surface-aligned H_{II} phases do not exhibit such scattering contributions (Ding *et al.*, 2004), disfavoring such a scenario. Hence, this appears to be only a property of unoriented H_{II} phases, fully immersed in aqueous solution. We speculate that the outermost boundary of H_{II} structures may try to avoid contact of the hydrocarbon with water by forming a lamellar layer, *i.e.* in some ways similar to hexosomes (Yaghmur & Glatter, 2009). Indeed, we were able to account for the additional diffuse scattering by adding a form factor of a laterally uniform, infinitely extended, planar bilayer

$$F_{\text{BL}}(q \mid \boldsymbol{\rho}_{\text{lam}}) = 4\pi^2 \int \Delta\rho_{\text{lam}}(z)\exp(iqz)\, dz \quad (7)$$

to the total scattered intensity, where z is the coordinate normal to the lamellar phase and $\boldsymbol{\rho}_{\text{lam}}$ are the parameters describing the SLD of the lamellar phase. We cannot exclude that the additional diffuse scattering originates from unilamellar vesicles or other kinetically trapped aggregates formed during sample preparation.

Considering orientational averaging, we finally obtained for the total scattered intensity
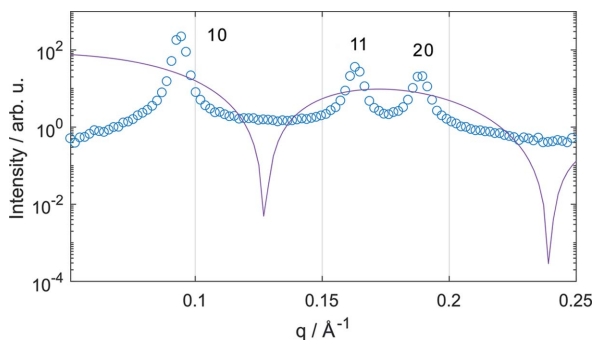

Figure 2
Overlay of the scattering pattern of DOPE (circles) and $|F_{\text{lipid}}(q)|^2$ (solid line). The phase change between the 10 and 11 reflections for the H_{II} phase leads to a minimum in the absolute square of the form factor, which is absent in the experimental data.

$$I^{\text{mod}}(q \mid n, \Delta, \sigma_{\text{fluc}}, \boldsymbol{\rho}, c_{\text{lam}}, \boldsymbol{\rho}_{\text{lam}})$$
$$\propto \int\limits_{0}^{\pi} |F(q, \theta \mid \boldsymbol{\rho}, \sigma_{\text{fluc}})|^2 S(q, \theta \mid n, \Delta)\sin\theta\, d\theta$$
$$+ 2c_{\text{lam}} F_{\text{BL}}(q \mid \boldsymbol{\rho}_{\text{lam}})$$
$$\times \int\limits_{0}^{\pi} F(q, \theta \mid \boldsymbol{\rho}, \sigma_{\text{fluc}}) s(q, \theta \mid n, \Delta)\sin\theta\, d\theta$$
$$+ c_{\text{lam}}^2 |F_{\text{BL}}(q \mid \boldsymbol{\rho}_{\text{lam}})|^2, \quad (8)$$

where c_{lam} denotes the fraction of the lamellar phase. The structure factor

$$s(q, \theta \mid n, \Delta) = \frac{1}{[N_{\text{hex}}(n)]^{1/2}}\exp(-q^2\Delta/2)\sum_{j}^{N_{\text{hex}}(n)} J_0(q|\mathbf{R}_j|\sin\theta) \quad (9)$$

was derived analogously to the H_{II} structure factor [equation (2)] and the form factor is $F(q, \theta \mid \boldsymbol{\rho}, \sigma_{\text{fluc}}) = f(q, \theta) \times [F_{\text{lipid,fluc}}(q, \theta \mid \boldsymbol{\rho}, \sigma_{\text{fluc}}) + F_{\text{inter}}(q, \theta \mid \boldsymbol{\rho})]$ [see equation (3)].

3.2. Composition-specific modeling

In this section we develop a model for the SLDs described by the parameters $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\text{lam}}$. For increased structural fidelity we considered the minimum number of parameters. We also constrained the SLDs by the specific molecular compositions. Assuming that tricosene partitions exclusively into the interstitial space, the PE structure was parsed into three cylindrical shells of a wedge-shaped lipid unit cell of opening angle α and height h (Fig. 3): (i) the headgroup (H), consisting of phosphate and ethanolamine groups, (ii) the glycerol backbone (BB), given by the carbonyl and glycerol groups, and (iii) the tails (HC) consisting of all methyl, methine and methylene groups. The outer radius of the wedge $a/2$ is evaluated in advance from the Bragg peak positions $q_{kl} = [4\pi/(3^{1/2}a)] \times (k^2 + 2kl + l^2)$, where k and l are the Miller indices. The position of the neutral plane R_0 was assumed to be in the center of the BB shell. This was motivated by bending/compression experiments, which obtained estimates for the location of the neutral plane within the lipid backbone regime (Kozlov & Winterhalter, 1991; Leikin *et al.*, 1996). In our model, the entire PE structure is described by the intrinsic curvature $C_0 = -1/R_0$, the width of the headgroup d_{H} and the width of the backbone d_{BB}. Further structural parameters of interest, such as the width of the hydrocarbon chain

$$d_{\text{HC}} = a/2 - d_{\text{BB}}/2 - R_0, \quad (10)$$

the lipid head-to-headgroup length

$$d_{\text{HH}} = 2(d_{\text{H}} + d_{\text{BB}} + d_{\text{HC}}) \quad (11)$$

and the radius of the water core

$$R_{\text{W}} = (a - d_{\text{HH}})/2 = R_0 - d_{\text{BB}}/2 - d_{\text{H}} \quad (12)$$

follow from these three parameters.

In the case of X-ray scattering the SLDs (electron densities) of each shell are given by $\rho_k = n_k^{\text{e}}/V_k$ with $k \in \{\text{H, BB, HC}\}$, where n_k^{e} is the number of electrons of a given quasi-molecular lipid fragment, $V_{\text{H}} = 110\ \text{Å}^3$, $V_{\text{BB}} = 135\ \text{Å}^3$ (Kučerka *et al.*,

2015) and $V_{HC} = V_{lipid} - V_{BB} - V_H$. Furthermore, we estimated $\Delta\rho_{inter}$ [equation (6)] of tricosene by molecular averaging over the fractional volumes of v_{CH}, v_{CH_2} and v_{CH_3} (Kučerka et al., 2015) (see supporting Table S1). In our model the electron density is sufficiently described by the parameters C_0, d_H, d_{BB} and V_{lipid}, and hence $\rho = \{C_0, d_H, d_{BB}, V_{lipid}\}$. All other parameters can be deduced from these by using the lipid contribution to the volume of the kth shell:

$$V_k = \frac{\hat{A}(r_{k+1}^2 - r_k^2)}{2} - \tilde{n}_W^k V_W, \qquad (13)$$

where $\tilde{n}_W^k$ is the number of water molecules within each shell and $V_W = 30$ Å^3 is the molecular volume of water. $\hat{A} = \alpha h$ is the mantle area of a sector of unitary radius and can be obtained using equation (13) with $k = HC$ and $\tilde{n}_W^{HC} = 0$:

$$\hat{A} = \frac{2V_{HC}}{a^2/4 - (R_0 + d_{BB}/2)^2}. \qquad (14)$$

Hence, equation (13) also defines $\tilde{n}_W^H$ and $\tilde{n}_W^{BB}$.

Using our parametrization, it is straightforward to derive the area per lipid at any position within the molecule. For example, the area per lipid in the neutral plane is calculated as

$$A_0 = \hat{A}R_0. \qquad (15)$$

Furthermore, following Israelachvili (2011), the molecular shape parameter is given by

$$\tilde{S} = \frac{V_{HC}}{\hat{A}(R_0 + d_{BB}/2)d_{HC}}, \qquad (16)$$

where $\tilde{S} = 1$ represents cylindrical (lamellar-phase-forming) molecules, and $\tilde{S} < 1$ or $\tilde{S} > 1$ typify molecules inducing negative or positive monolayer curvatures, respectively. In particular, $\tilde{S} > 1$ for amphiphiles forming aggregates with negative curvature, like the H$_{II}$ structure.

The form factor of the additional lamellar phase was calculated by integrating equation (7), using a simple SLD model consisting of head and tail slabs (Székely et al., 2010):

$$\begin{aligned}
F_{BL}(q) = \frac{4\pi^2}{iq} \Big(&\Delta\rho_{H,lam}\left[\exp(iqd_{H,lam}) - 1\right] \\
&+ \Delta\rho_{H,lam}\{\exp[i2q(d_{H,lam} + d_{HC,lam})] \\
&- \exp[iq(d_{H,lam} + 2d_{HC,lam})]\} \\
&+ \Delta\rho_{HC,lam}\{\exp[iq(d_{H,lam} + 2d_{HC,lam})] \\
&- \exp(iqd_{H,lam})\} \Big),
\end{aligned} \qquad (17)$$

where $\Delta\rho_{H,lam}$ and $\Delta\rho_{HC,lam}$ are the headgroup and hydrocarbon SLDs relative to water, respectively. These were derived as detailed above by counting the number of electrons in each slab and dividing by the corresponding volumes $V_{H,lam}$ or $V_{HC,lam}$. Assuming that $V_{HC,lam} = V_{HC}$ and $V_{H,lam} = V_H + V_{BB}$, the hydrocarbon slab thickness results from

$$d_{HC,lam} = V_{HC}/A_L \qquad (18)$$

and the headgroup thickness from

$$d_{H,lam} = (V_{H,lam} + n_{W,lam}V_W)/A_L. \qquad (19)$$

Hence, the area per lipid A_L and the number of headgroup-associated water molecules $n_{W,lam}$ are the only parameters for the lamellar phase, $\rho_{lam} = \{A_L, n_{W,lam}\}$.

4. Parameter estimation using Bayesian probability theory

The final model for scattered intensities of unoriented fully hydrated H$_{II}$ is given by

$$I^{sim}(q \mid \mathbf{x}) = \Gamma I^{mod}(q \mid n, \Delta, \sigma_{fluc}, \rho, c_{lam}, \rho_{lam}) + I_{inc}, \qquad (20)$$

where I^{mod} is given by equation (8), Γ is an instrumental scaling constant and I_{inc} accounts for incoherent scattering. In total, we have 12 model parameters, denoted by $\mathbf{x}$, which are listed in Table 1.

There are various ways of estimating the parameters $\mathbf{x}$. For a given data set $\mathbf{I}$ with the standard deviations σ in the presence of a well defined global minimum the method of least squares yields fitting parameters by minimizing a cost function $\chi^2(\mathbf{x} \mid \mathbf{I}, \sigma)$. However, such an approach led for our present data to a significant variation of results between consecutive optimization runs, indicating a cost function landscape with a weakly defined global minimum. Thus, besides unreliable $\mathbf{x}$ values, also error estimates and potential correlations between the parameters remained undetermined.
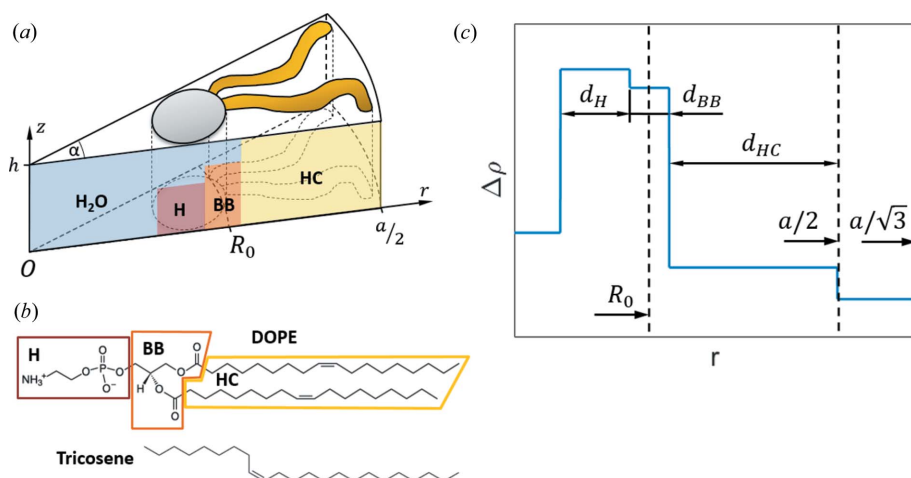

Figure 3
Composition-specific SLD modeling of phosphatidylethanolamines. (a) The unit cell of a single lipid has the shape of a cylinder sector of radius $a/2$. (b) Parsing of DOPE into head (H), backbone (BB) and hydrocarbon chain (HC) and chemical structure of tricosene. (c) Scheme of a corresponding electron density profile (see also Fig. 1).

research papers

Table 1
Overview of the model parameters for fully hydrated, unoriented H_{II} phases.

Occurrence	x	Meaning
Structure factor	Δ	Mean-square displacement of the lattice points
	n	Number of hexagonal shells (domain size)
H_{II} form factor	σ_{fluc}	Fluctuation constant of the lipid unit cell
	C_0	Intrinsic curvature
	d_H	Width of the lipid headgroup
	d_{BB}	Width of the lipid backbone
	V_{lipid}	Lipid volume
Lamellar form factor	c_{lam}	Lamellar form factor scaling constant
	A_L	Area per lipid of the lamellar phase
	$n_{W,lam}$	Number of water molecules in the headgroup slab of the lamellar phase
Signal scaling	Γ	Instrumental scaling constant
	I_{inc}	Incoherent background

To achieve higher confidence in our results we decided to use Bayesian probability theory; for a detailed introduction, see Jaynes & Bretthorst (2003), Gregory (2005), Sivia & Skillings (2012) and von der Linden *et al.* (2014). In brief, we were interested in deriving the probability $p(\mathbf{x} \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$, meaning the probability of the parameters $\mathbf{x}$ given the set of experimental data $\mathbf{I}$ with standard deviations $\boldsymbol{\sigma}$ and additional information $\mathcal{I}$ which might be present, such as the finite width of a lipid molecule. In the framework of Bayesian probability theory, Bayes' theorem shows how to calculate this quantity, also called the posterior:

$$\overbrace{p(\mathbf{x} \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})}^{\text{posterior}} \propto \overbrace{p(\mathbf{I} \mid \mathbf{x}, \boldsymbol{\sigma}, \mathcal{I})}^{\text{likelihood}} \overbrace{p(\mathbf{x} \mid \not{\boldsymbol{\sigma}}, \mathcal{I})}^{\text{prior}}. \quad (21)$$

Bayes' theorem constitutes the rule for learning from experimental data. The prior probability $p(\mathbf{x} \mid \mathcal{I})$ represents the prior knowledge about the unknown quantities $\mathbf{x}$. We crossed out $\boldsymbol{\sigma}$ in the prior, since the prior does not depend on the standard deviations of the data. The likelihood $p(\mathbf{I} \mid \mathbf{x}, \boldsymbol{\sigma}, \mathcal{I})$, representing the probability for the data $\mathbf{I}$ given $\mathbf{x}$ and $\boldsymbol{\sigma}$, includes all information about the measurement itself. The prior probabilities $p(\mathbf{x} \mid \mathcal{I})$ were assumed to be uniformly distributed between lower $\mathbf{x}_{min}$ and upper $\mathbf{x}_{max}$ constraints for all parameters. Therefore,

$$p(\mathbf{x} \mid \mathcal{I}) = \prod_i \frac{\Theta(x_i - x_{i,min}) - \Theta(x_i - x_{i,max})}{x_{i,max} - x_{i,min}}, \quad (22)$$

where $\Theta(x)$ is the Heaviside step function. For each parameter, x_i, $x_{i,min}$ and $x_{i,max}$ denote physically meaningful boundaries. In particular, we constrained the parameters d_H and d_{BB} by the conditions

$$\tilde{n}_W^H \geq 0 \quad \text{and} \quad \tilde{n}_W^{BB} \geq 0. \quad (23)$$

This means that the volumes of the head and backbone shell [equation (13)] have to be large enough to accommodate the respective molecular group.

We consider the likelihood $p(\mathbf{I} \mid \mathbf{x}, \boldsymbol{\sigma}, \mathcal{I})$. Since we did not trust the experimentally derived error estimates $\boldsymbol{\sigma}$ for the

scattered intensities, we assumed that their real values $\tilde{\boldsymbol{\sigma}}$ are connected to $\boldsymbol{\sigma}$ by a scaling factor η. Using the marginalization rule of Bayesian probability theory we obtain

$$p(\mathbf{I} \mid \mathbf{x}, \boldsymbol{\sigma}, \mathcal{I}) = \int d\eta \, d\tilde{\boldsymbol{\sigma}} \, p(\mathbf{I} \mid \mathbf{x}, \boldsymbol{\phi}, \tilde{\boldsymbol{\sigma}}, \not{\eta}, \mathcal{I}) \, p(\tilde{\boldsymbol{\sigma}}, \eta \mid \not{\mathbf{x}}, \boldsymbol{\sigma}, \mathcal{I})$$
$$= \int d\eta \, d\tilde{\boldsymbol{\sigma}} \, p(\mathbf{I} \mid \mathbf{x}, \tilde{\boldsymbol{\sigma}}, \mathcal{I}) \frac{\overbrace{p(\tilde{\boldsymbol{\sigma}} \mid \eta, \boldsymbol{\sigma}, \mathcal{I})}}{\prod_i \delta(\tilde{\sigma}_i - \eta\sigma_i)} \overbrace{p(\eta \mid \not{\boldsymbol{\phi}}, \mathcal{I})}^{\propto 1/\eta}$$
$$\propto \int d\eta \, \underbrace{p(\mathbf{I} \mid \mathbf{x}, \tilde{\boldsymbol{\sigma}} = \eta\boldsymbol{\sigma}, \mathcal{I})}_{\mathcal{N}(\mathbf{I} \mid \mathbf{x}, \eta\boldsymbol{\sigma})} (1/\eta), \quad (24)$$

where $d\tilde{\boldsymbol{\sigma}}$ is shorthand for $\prod_i d\tilde{\sigma}_i$. Here, we have specifically made use of the Jeffreys prior $p(\eta) \propto 1/\eta$ (Jeffreys, 1946), where η is a scaling invariant, meaning that we have *a priori* no idea about the order of magnitude of η. This scaling connects the likelihood [equation (24)] to the multivariate Gaussian:

$$\mathcal{N}(\mathbf{I} \mid \mathbf{x}, \eta\boldsymbol{\sigma}) = \prod_{i=1}^{N_q} \frac{1}{\eta\sigma_i(2\pi)^{1/2}} \exp\left\{-\frac{1}{2\eta^2\sigma_i^2}[I^{sim}(q_i \mid \mathbf{x}) - I_i^{obs}]^2\right\}, \quad (25)$$

where η has to be integrated out, respecting the Jeffreys prior, and I_i^{obs} denotes the observed intensity at q_i. Here, N_q is the number of data points for a given scattering pattern.

For illustration, consider an arbitrary function $\mathcal{O}(\mathbf{x})$ with the parameters $\mathbf{x}$. The expectation value of $\mathcal{O}(\mathbf{x})$ is then calculated by evaluating the integral

$$\langle\mathcal{O}(\mathbf{x})\rangle = \int d\mathbf{x} \, d\eta \, \mathcal{O}(\mathbf{x}) \, p(\mathbf{x}, \eta \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I}), \quad (26)$$

where

$$p(\mathbf{x}, \eta \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I}) = \frac{1}{Z} \mathcal{N}(\mathbf{I} \mid \mathbf{x}, \eta\boldsymbol{\sigma}) \frac{1}{\eta} p(\mathbf{x} \mid \mathcal{I}) \quad (27)$$

with the normalization constant Z. For example, using $\mathcal{O}(\mathbf{x}) = x_i$ produces the expectation value $\langle x_i \rangle$ for parameter x_i.

A suitable technique for performing these integrals and sampling from the probability distribution $p(\mathbf{x}, \eta \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ is Markov chain Monte Carlo (MCMC), which is based on constructing a Markov chain with the desired distribution of $\mathbf{x}$ in equilibrium. We used the Metropolis–Hastings algorithm for generating the Markov chain $\{\mathbf{x}^k, \eta^k\}$. Starting with a parameter set $\mathbf{x}^{k=1}$ and $\eta^{k=1}$, every new parameter set $k+1$ can be proposed by varying parameters in the old parameter set k. The new parameter set $k+1$ is accepted with the probability

$$P_{acc} = \min\left[1, \frac{p(\mathbf{x}^{k+1}, \eta^{k+1} \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})}{p(\mathbf{x}^k, \eta^k \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})}\right]. \quad (28)$$

The first 10−20% of a Markov chain has to be discarded to ensure that the rest of the Markov chain is independent of the initial state $\mathbf{x}^{k=1}$ and $\eta^{k=1}$, *i.e.* the Markov chain is equilibrated to the desired distribution.

In addition, the states in the Markov chain have to be uncorrelated, which can be ensured by taking only every N_{run}th state of the Markov chain. N_{run} can be controlled by evaluating the autocorrelation function or using techniques like binning and jackknifing. Finally, the observable can be estimated by

$$\mathcal{O} := \langle \mathcal{O}(\mathbf{x}) \rangle \simeq \frac{1}{N_{\text{Markov}}} \sum_{k=1}^{N_{\text{Markov}}} \mathcal{O}(\mathbf{x}^k), \qquad (29)$$

i.e. taking the mean value of N_{Markov} uncorrelated Markov chain elements. The confidence intervals can be estimated from

$$\Delta\mathcal{O} := \sigma_{\mathcal{O}}/N_{\text{Markov}}^{1/2}. \qquad (30)$$

The variance

$$\sigma_{\mathcal{O}}^2 = \langle \mathcal{O}(\theta)^2 \rangle - \langle \mathcal{O}(\theta) \rangle^2 \qquad (31)$$

can in turn be estimated from the Markov chain. Alternatively, the uncertainty can be determined from independent MCMC runs.

Since the Markov chain $\{\mathbf{x}^k, \eta^k\}$ is a representative sample drawn from $p(\mathbf{x}, \eta \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ it can be used to plot the probability distribution, for example the marginal probability distribution $p(x_i, x_j \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ for the parameters i and j by plotting the two-dimensional histogram of the samples $\{x_i^k\}$ and $\{x_j^k\}$. This allows us to unravel correlations between the parameters i and j, *i.e.* it allows the analysis of mutual parameter dependencies that could lead to ambiguous results using the least-squares method. Additionally, the cost function

$$\chi^2 = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{\left[I(q_i \mid \mathbf{x})^{\text{sim}} - I_j^{\text{obs}}\right]^2}{\tilde{\sigma}_i^2} \qquad (32)$$

is saved for every run.

5. Results and discussion

5.1. Tests of the analysis

We first explored our model and the Bayesian analysis on the well studied H_{II} structure of DOPE (Turner & Gruner, 1992, 1989; Harper *et al.*, 2001; Kollmitzer *et al.*, 2013). We emphasize that the choice of our model restricts the algorithm to a certain functional space for describing the scattering pattern. This constraint can lead to some systematic deviations from the experimental data, and Bayesian model comparison can be used for choosing the appropriate model.

Clearly, our model is able to account well for most features of the scattering pattern up to $q \simeq 0.5 \text{ Å}^{-1}$ (Fig. 4). In particular, the bilayer form factor compensates well for the form factor minimum between the 10 and 11 peaks, but adds also some diffuse scattering at higher q values. The small peak observed in the calculated intensity at very low q is an artifact resulting from the structure factor. This could be removed by averaging over a distribution of domains (Freiberger & Glatter, 2006), but does not affect the overall structural results and has therefore been omitted to reduce computational cost (see also above). Furthermore, the proximity of a form factor minimum to the 21 peak of the H_{II} phase nearly causes an extinction in the scattering data. Note that tricosene-free DOPE samples exhibit a clear 21 reflection [Fig. S4(c)]. However, because of strain-induced distortions of the hexagonal prisms such samples cannot be analyzed with the present model. The maximum aposterior (MAP) solution still shows a

slightly more pronounced 21 peak, since a perfect fit in this q range would lead to significant deviations between the model and experimental data close to the 20 peak, which due to its smaller errors has a higher significance in contributing to the overall goodness of the MAP solution. Additionally, our MAP solution underestimates the contributions of the 22 and 31 peaks owing to the proximity of the cylinder form factor to two minima. To account for this we tested SLD models of greater complexity, by considering either a separate slab for the methyl terminus of the hydrocarbon chain or a linear decrease of the electron density in the hydrocarbon regime. However, this did not lead to a significant improvement of the agreement between the model and the experimental data in this q range. In order to avoid overfitting, we therefore remained with the SLD model as described in Section 3. Table S2 lists the corresponding expectation values $\langle \mathbf{x} \rangle$ and variances $\sigma_{\mathbf{x}}$. In order to check for reproducibilty, we prepared a fresh DOPE sample. The results listed in Table S2 show that all structural lipid parameters are identical within experimental uncertainty [see also Fig. S4(b)]. Differences in lattice parameters, such as a and Δ, relate to slight variations of tricosene content.

One of the benefits of the Bayesian analysis compared with the least-squares method is the possibility to reveal
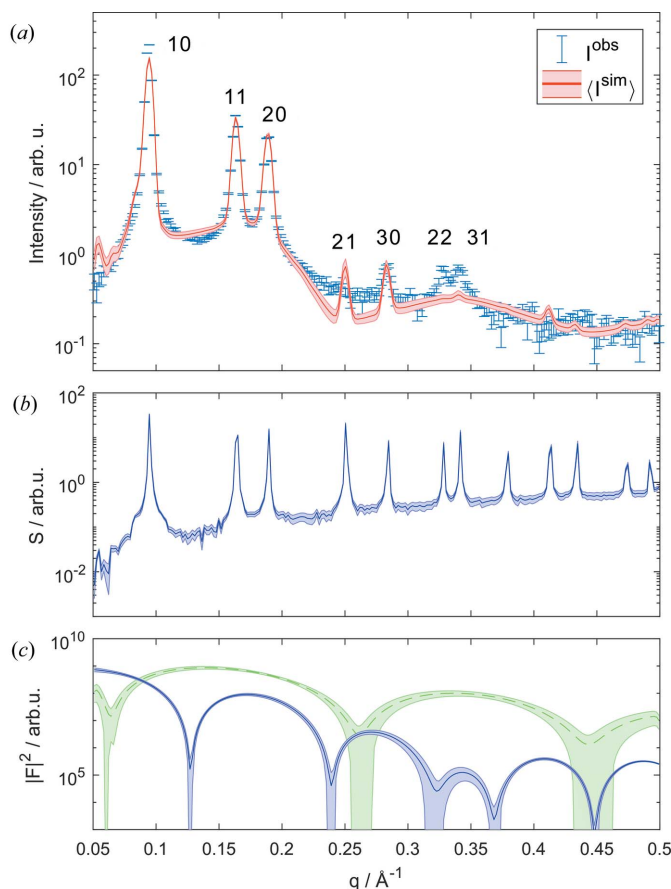

Figure 4
Expectation value and error bands of the intensity of fully hydrated DOPE at 308 K (*a*), and the involved structure (*b*) and form factors (*c*) (blue: hexagonal form factor; green: lamellar form factor).

Table 2
Comparison of structural parameters of DOPE with literature values.

a (Å)	V_L (Å^3)	d_{HH} (Å)	R_W (Å)	A_0 (Å^2)	C_0 (Å^{-1})	Reference
76.9 ± 0.2	1142 ± 10	32.4 ± 1.3	22.2 ± 0.7	62.2 ± 6.0	-0.0409 ± 0.0010	This work
71.9	1224	31.8	20.0	–	–	Tate & Gruner (1989)[a]
74.9	–	–	22	–	-0.033^i	Rand et al. (1990)[b]
72.75	–	–	19.1	–	–	Turner & Gruner (1992)[c]
–	–	–	–	–	-0.0367 ± 0.0005	Leikin et al. (1996)[d]
72.9	1220	36.0	20.4	47.4^i	–	Harper et al. (2001)[e]
76	–	–	–	51.5^i	-0.031^i	Di Gregorio & Mariani (2005)[f]
–	–	–	–	–	-0.0399 ± 0.0005	Kollmitzer et al. (2013)[g]
–	–	–	–	–	-0.0365 ± 0.0012	Chen et al. (2015)[h]

(a) $T = 308$ K, dodecane. (b) $T = 295$ K, tetradecane. (c) $T = 303$ K, dodecane. (d) $T = 298$ K. (e) $T = 303$ K. (f) $T = 298$ K. (g) $T = 308$ K, tricosene. (h) $T = 303$ K, tetradecane. (i) Determined in the pivotal plane.

correlations between adjustable parameters, just by looking at the 2D marginal probability density distributions [see e.g. Fig. 5(a)]. Marginal distributions of all other parameters are

Figure 5
Marginal posterior distributions $p(C_0 \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ and $p(d_H \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ of intrinsic curvature C_0 and headgroup width d_H and $p(C_0, d_H \mid \mathbf{I}, \boldsymbol{\sigma}, \mathcal{I})$ (a). The red cross and corresponding lines mark the sample with the lowest χ^2 (MAP solution), and the green circle shows the mean value of the distribution. Panels (b) and (c) show the corresponding fits and SLD profiles.

shown in the supplementary Figs. S1–S3. Most parameter pairs show no correlations and exhibit probability distributions with Gaussian-like behavior, including σ_{fluc}, V_{lipid} and Γ. Significant correlations can be seen for the parameters d_H and d_{BB} with C_0, as well as between $n_{W,lam}$ and A_L. A strong correlation between two parameters suggests the possibility to simplify the model. However, this would be highly specific for a given amphiphile and was consequently not considered. The parameters d_H and d_{BB} exhibit broad probability distributions with no well defined maximum. In turn C_0 has a peaked probability distribution yielding a well defined estimate value and uncertainty.

Here, we discuss for illustration purposes the correlation between C_0 and d_H [Fig. 5(a)]. Solutions along the diagonal line give similar scattering intensities, but lead to significantly different electron density profiles [see Figs. 5(b) and 5(c)].

The correlation between d_H and C_0 may appear counterintuitive. From geometric/ physical arguments it follows that small d_H values represent a bending of the phosphate– ethanolamine director of the lipid headgroup toward the polar/apolar interface, which leads to a shift of C_0 toward more positive values. However, this would lead to significantly different scattered intensities and hence to non-optimal solutions. The mathematical algorithm therefore aims to compensate for this by decreasing C_0 for small d_H. The abrupt drop of the headgroup thickness probability distribution at small d_H is due to the termination criterion [equation (23)].

In the following, we discuss some expectation values $\langle \mathbf{x} \rangle$ and the errors $\sigma_{\mathbf{x}}$ obtained by applying equations (29) and (31). Table 2 compares the obtained structural parameters of DOPE with existing literature values. Our results are in good agreement with previous reports, given the different additives (alkanes
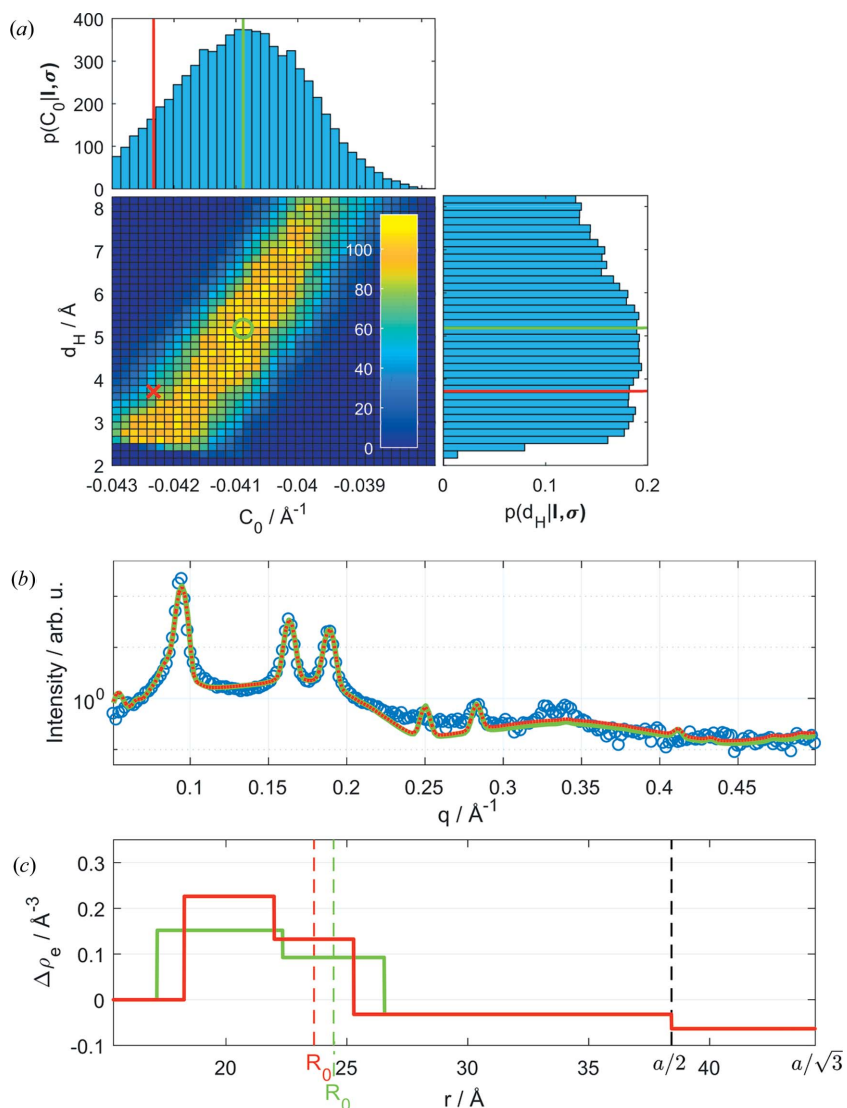
or alkenes, some did not use any filler molecule) and slight variations in temperatures. Note that in some cases A_0 and C_0 have been reported for the pivotal plane. The pivotal plane marks the position within the lipids where the molecular area does not change upon deformation and is usually slightly closer to the hydrocarbon tail than the neutral plane (Leikin *et al.*, 1996; Kollmitzer *et al.*, 2013). This leads to a slight shift of C_0 toward positive values.

The most direct comparison of C_0 can be made with our previous work (Kollmitzer *et al.*, 2013), which was performed

at the same temperature and tricosene content. Here, we find that the global model combined with Bayesian analysis yields an intrinsic curvature, which agrees within experimental uncertainty with our previous result.

5.2. Effect of temperature

Increasing the temperature for DOPE should yield a decrease of lipid chain length and a concomitant significant increase of the area per lipid at the methyl terminus leading to more negative intrinsic curvatures as reported previously
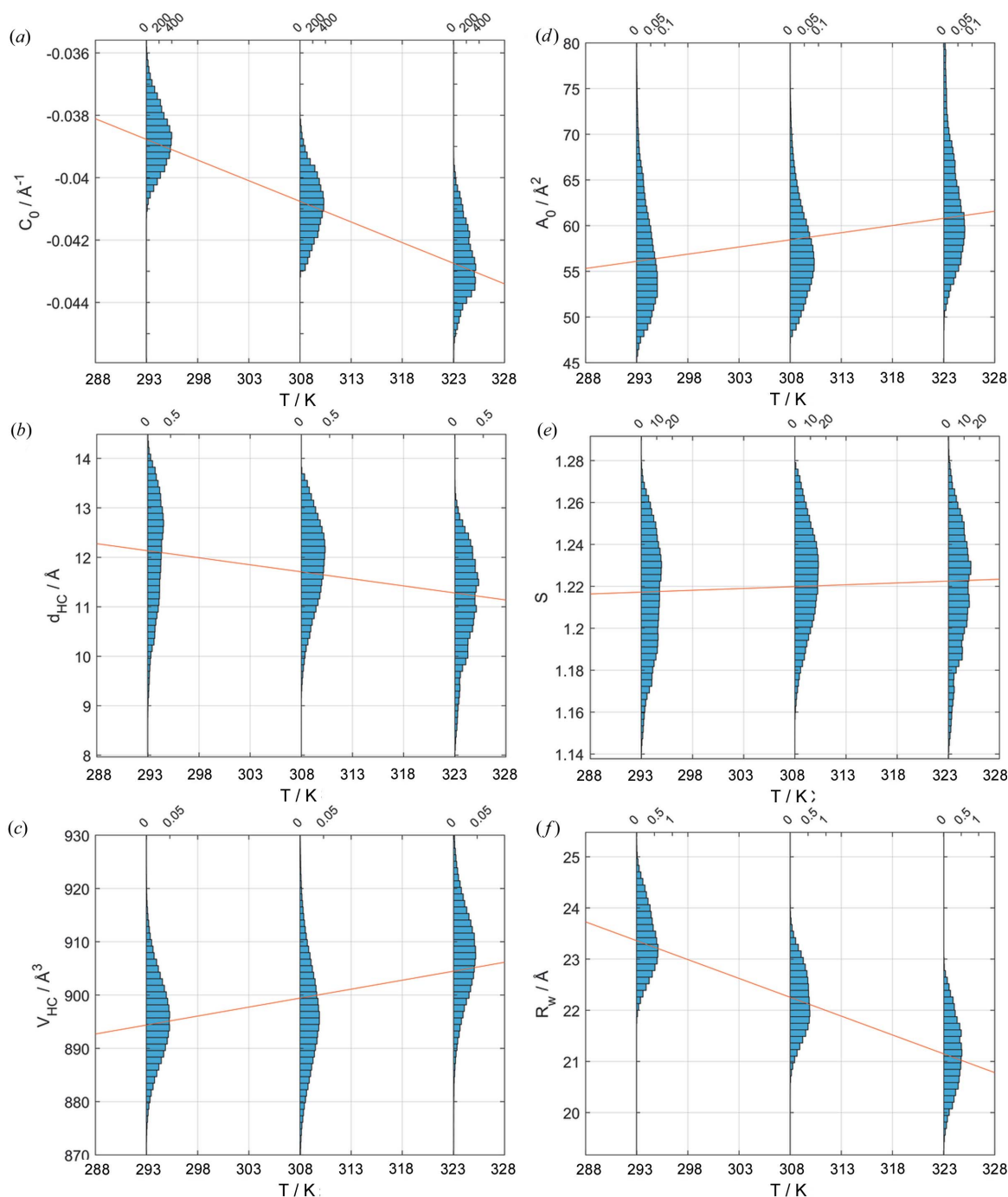

Figure 6
Structural parameters of DOPE H_{II} as a function of temperature resulting from the Bayesian analysis. Probability densities (*a*) of the intrinsic curvatures, (*b*) of the hydrocarbon chain length, (*c*) of the hydrocarbon chain volume, (*d*) of the area per lipid in the neutral plane, (*e*) of the shape parameter and (*f*) of the radius of the water cylinder. Red lines indicate linear regressions of the probability density distributions.

(Turner & Gruner, 1992, 1989; Harper *et al.*, 2001; Kollmitzer *et al.*, 2013). Indeed, our analysis yielded a linear decrease of C_0 and d_{HC} (Fig. 6). The slope $\Delta C_0/\Delta T = (-1.323 \pm 0.001) \times 10^{-4}$ (Å K)$^{-1}$ is identical to our previously reported value (Kollmitzer *et al.*, 2013). The relative change of the chain length is in turn $\Delta d_{HC}/\Delta T = (-0.0188 \pm 0.0001)$ Å K^{-1}. Interestingly, the shape parameter $\hat{S}$ shows only a modest increase of $\Delta \hat{S}/\Delta T = (1.75 \pm 0.03) \times 10^{-4}$ K^{-1}, despite the more negative C_0 values at higher temperatures and despite the decrease of d_{HC} and the increase of V_{HC} [$\Delta V_{HC}/\Delta T = (0.3364 \pm 0.0010)$ Å^3 K^{-1}]. This results from a concomitant increase of headgroup area [$\Delta A_0/\Delta T = (0.1567 \pm 0.0006)$ Å^2 K^{-1}] in the neutral plane – and analogously also at the position of the polar/apolar interface ($R_0 + d_{BB}/2$) – which compensates for the changes of d_{HC} and V_{HC}. The radius of the water core decreases [$\Delta R_W/\Delta T = (-7.364 \pm 0.007) \times 10^{-2}$ Å K^{-1}].

6. Effect of hydrocarbon chain composition

Finally, we tested the applicability of the analysis technique to PEs with differing hydrocarbon chain composition. In particular, we studied the H$_{II}$ phases of POPE, which has a palmitoyl and an oleoyl chain, DMPE, which has two myristoyl chains, and diC16:1PE, with two palmitoleoyl hydrocarbons. Note that pure POPE forms an H$_{II}$ phase only above 344 K, while the lamellar-to-H$_{II}$ phase-transition temperature T_H for pure di16:1PE was reported to be 316.6 K and $T_H > 373$ K for DMPE (Koynova & Caffrey, 1994). The addition of alkanes or alkenes to inverted hexagonal phases is known to reduce stress resulting from interstitial space between the individual rods (Vacklin *et al.*, 2000; Chen & Rand, 1998; Kirk & Gruner, 1985). We previously demonstrated that tricosene sufficiently lowers T_H for POPE to perform an H$_{II}$ phase analysis at physiological temperature (Kollmitzer *et al.*, 2013). Similarly, di16:1PE formed a neat H$_{II}$ phase at 308 K upon addition of 12 wt% tricosene (see below). In the case of DMPE we found a pure H$_{II}$ scattering pattern only for $T \geq$

353 K, indicating a significantly less negative C_0. For this reason we performed the global analysis at 308 K for POPE and di16:1PE and at 353 K for DMPE.

In contrast to the results for DOPE, the 21 peak was clearly present in the scattering data of all three lipids, a feature which helped us to obtain a better agreement of the model with experimental data (Fig. 7). The corresponding probability density distributions for C_0 clearly show that monounsaturated hydrocarbons induce significantly more negative intrinsic curvature than saturated hydrocarbons, which is due to the kink induced by the *cis* double bond. The proximity of values for POPE and DMPE is attributed to the temperature difference and the associated decrease of C_0 (Fig. 8). Assuming a similar temperature dependence to that observed for DOPE yields C_0 close to zero for DMPE at 308 K, which agrees with the well established observation that DMPE prefers to form bilayers at ambient temperatures. Besides the difference between saturated and unsaturated hydrocarbons, our analysis also clearly shows that $C_0^{DOPE} < C_0^{diC16:1PE}$. That is, increasing the chain length of monounsaturated acyl chains also leads to a more negative C_0 value. This signifies that the kink induced by the *cis* double bond leads to a progressive increase of hydrocarbon splay upon acyl chain extension.
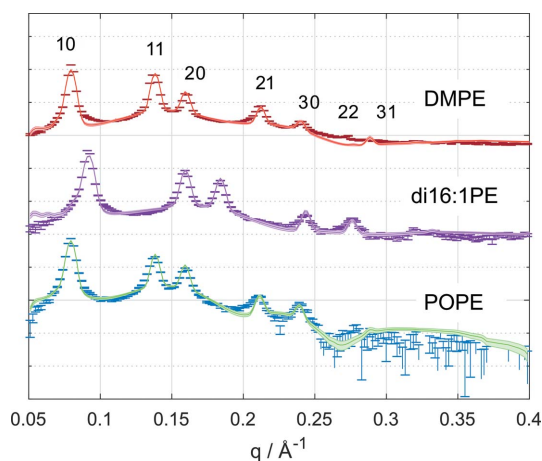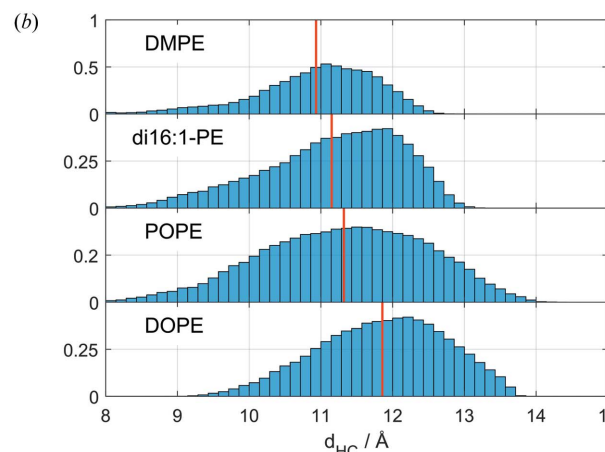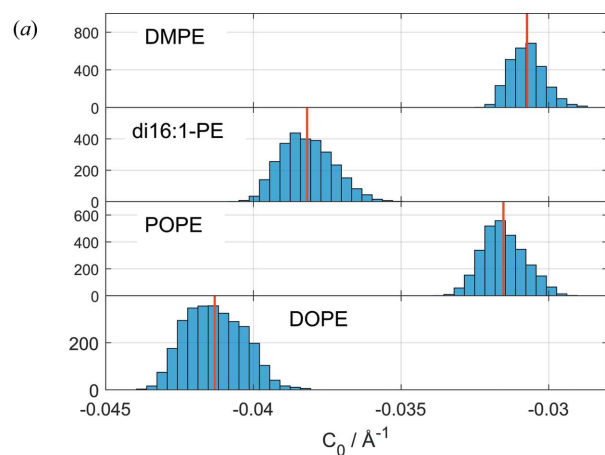

Figure 7
Data points (including error bars) and expectation values of the intensity (with error bands) of SAXS patterns of POPE (308 K), di16:1PE (308 K) and DMPE (353 K).

Figure 8
Intrinsic curvature (*a*) and hydrocarbon chain length (*b*) probability densities and mean values (red lines) for various lipids at 308 K (except DMPE: 353 K).

Table 3
Comparison of structural parameters of different phosphatidylethanolamines.

	C_0 (Å^{-1})	d_{HC} (Å)	R_W (Å)	A_0 (Å^2)	V_{HC} (Å^3)	$\tilde{S}$
DMPE[a]	-0.0314 ± 0.0006	10.9 ± 0.9	30.6 ± 0.7	59 ± 6	737 ± 7	1.16 ± 0.02
di16:1PE[b]	-0.0382 ± 0.0009	11.1 ± 1.0	24.0 ± 0.7	60 ± 7	797 ± 8	1.20 ± 0.03
POPE[b]	-0.0317 ± 0.0007	11.3 ± 1.2	29.2 ± 0.9	68 ± 9	884 ± 9	1.17 ± 0.03
DOPE[b]	-0.0409 ± 0.0010	11.9 ± 0.9	22.2 ± 0.7	62 ± 6	897 ± 10	1.22 ± 0.03

(a) T = 353 K. (b) T = 308 K.

The mean values of hydrocarbon chain length show a trend in the expected direction, *i.e.* they increase with the number of hydrocarbons (Fig. 8), but all cases exhibit a broad distribution as a result of the not well defined backbone width d_{BB}.

The expectation values for C_0 and d_{HC} for the different lipids are listed in Table 3, including resulting structural parameters for R_W, A_0, V_{HC} and $\tilde{S}$. Previously, we reported $C_0 = -0.0316$ Å^{-1} for POPE at 310 K (Kollmitzer *et al.*, 2013), which is in excellent agreement with our present value. Regarding other structural parameters, in particular we found that R_W and A_0 decrease as C_0 becomes more negative, which is mainly attributed to the geometry of the H$_{II}$ phase. The hydrocarbon chain volumes are in agreement with the chemical compositions. That is, DMPE with two C14:0 chains has the smallest and DOPE with two C18:1 chains has the largest V_{HC} value, and the volumes of POPE and diC16:1PE take intermediate values. Our hydrocarbon volume of POPE is about 4% lower than the value reported for POPE in the absence of tricosene at the same temperature, where it forms a fluid lamellar phase (Kučerka *et al.*, 2015). This indicates a slightly tighter hydrocarbon chain packing in fully relaxed monolayers. The shape parameter (DMPE $\simeq$ POPE < diC16:1PE < DOPE) clearly shows that among all lipids presently studied DOPE has the highest propensity to form an H$_{II}$ phase, which is consistent with its low T_H (Koynova & Caffrey, 1994).

7. Conclusions

We have introduced a global scattering model for fully hydrated, unoriented H$_{II}$ phases. Compared with previous models for H$_I$ phases (Freiberger & Glatter, 2006; Sundblom *et al.*, 2009), the H$_{II}$ phase analysis required the addition of diffuse scattering not originating from hexagonal structures. While the exact origin of this additional contribution remains unclear, we successfully modeled the measured SAXS pattern upon including a lamellar form factor. The SLD of the lipid unit cell was constrained by compositional modeling using complementary information on lipid volume and structure. This description is generic and entails the analysis of SAXS and small-angle neutron scattering (SANS) data. In particular a joint analysis of SAXS and differently contrasted SANS data [see *e.g.* Pabst *et al.* (2010) and Heberle & Pabst (2017)] might be beneficial for increased structural resolution regarding the lipid head and backbone groups.

Here, we analyzed SAXS data using Bayesian probability theory combined with MCMC simulations. This was necessary because of the weakly defined global minimum of the optimization cost function. The full probabilistic approach provides the probability density distributions of the involved parameters, leading to reliable parameter estimates including errors.

The obtained estimates are in good agreement with previously reported structural data of DOPE and POPE. We further provided details for lipid structures of DMPE and di16:1PE in the H$_{II}$ phase, clearly demonstrating that out of all of the presently studied lipids DMPE is the least prone to form an H$_{II}$ phase. The developed technique will be easily transferred to other H$_{II}$ phase amphiphiles using appropriate compositional modeling. In particular, we envisage a high potential for applications in drug-delivery formulations involving H$_{II}$ structures that exhibit only weak Bragg peaks but significant contributions from diffuse scattering, such as hexosomes [see *e.g.* Yaghmur & Glatter (2009)]. Another potential application is the determination of intrinsic lipid curvatures of lamellar-phase-forming lipids using mixtures with DOPE (Kollmitzer *et al.*, 2013), which is particularly encouraged by the high robustness of the retrieved C_0 estimates. Such approaches are currently being explored in our laboratory.

References

Alley, S. H., Ces, O., Barahona, M. & Templer, R. H. (2008). *Chem. Phys. Lipids*, **154**, 64–67.
Bubolz, J. T. & Feigenson, G. W. (1999). *Biochim. Biophys. Acta*, **1417**, 232–245.
Chen, Y.-F., Tsang, K.-Y., Chang, W.-F. & Fan, Z.-A. (2015). *Soft Matter*, **11**, 4041–4053.
Chen, Z. & Rand, R. P. (1998). *Biophys. J.* **74**, 944–952.
Dan, N. & Safran, S. A. (1998). *Biophys. J.* **75**, 1410–1414.
Di Gregorio, G. M. & Mariani, P. (2005). *Eur. Biophys. J.* **34**, 67–81.
Ding, L., Liu, W., Wang, W., Glinka, C. J., Worcester, D. L., Yang, L. & Huang, H. W. (2004). *Langmuir*, **20**, 9262–9269.
Förster, S., Timmann, A., Konrad, M., Schellbach, C., Meyer, A., Funari, S. S., Mulvaney, P. & Knott, R. (2005). *J. Phys. Chem. B*, **109**, 1347–1360.
Franks, N. P., Melchior, V., Kirschner, D. A. & Caspar, D. (1982). *J. Mol. Biol.* **155**, 133–153.
Freiberger, N. & Glatter, O. (2006). *J. Phys. Chem. B*, **110**, 14719–14727.
Frewein, M., Kollmitzer, B., Heftberger, P. & Pabst, G. (2016). *Soft Matter*, **12**, 3189–3195.
Frolov, V. A., Shnyrova, A. V. & Zimmerberg, J. (2011). *Cold Spring Harb. Perspect. Biol.* **3**, a004747.

Glatter, O. (2018). *Scattering Methods and Their Application in Colloid and Interface Science.* Amsterdam: Elsevier.

Gregory, P. (2005). *Bayesian Logical Data Analysis for the Physical Sciences.* Cambridge University Press.

Harper, P. E., Mannock, D. A., Lewis, R. N., McElhaney, R. N. & Gruner, S. M. (2001). *Biophys. J.* **81**, 2693–2706.

Heberle, F. A. & Pabst, G. (2017). *Biophys. Rev.* **9**, 353–373.

Israelachvili, J. N. (2011). *Intermolecular and Surface Forces.* Oxford: Academic Press.

Jaynes, E. T. & Bretthorst, G. L. (2003). *Probability Theory, The Logic of Science.* Cambridge University Press..

Jeffreys, H. (1946). *Proc. R. Soc. London Ser. A*, **186**, 453–461.

Kirk, G. L. & Gruner, S. M. (1985). *J. Phys. Fr.* **46**, 761–769.

Kollmitzer, B., Heftberger, P., Rappolt, M. & Pabst, G. (2013). *Soft Matter*, **9**, 10877–10884.

Koltover, I., Salditt, T., Rädler, J. O. & Safinya, C. R. (1998). *Science*, **281**, 78–81.

Koynova, R. & Caffrey, M. (1994). *Chem. Phys. Lipids*, **69**, 1–34.

Kozlov, M. M. & Winterhalter, M. (1991). *J. Phys. II*, **1**, 1077–1084.

Kučerka, N., van Oosten, B., Pan, J., Heberle, F. A., Harroun, T. A. & Katsaras, J. (2015). *J. Phys. Chem. B*, **119**, 1947–1956.

Leber, R., Pachler, M., Kabelka, I., Svoboda, I., Enkoller, D., Vácha, R., Lohner, K. & Pabst, G. (2018). *Biophys. J.* **114**, 1945–1954.

Leikin, S., Kozlov, M. M., Fuller, N. L. & Rand, R. P. (1996). *Biophys. J.* **71**, 2623–2632.

Linden, W. von der, Dose, V. & von Toussaint, U. (2014). *Bayesian Probability Theory.* Cambridge University Press.

Marchal, D. & Demé, B. (2003). *J. Appl. Cryst.* **36**, 713–717.

Marsh, D. (2006). *Chem. Phys. Lipids*, **144**, 146–159.

Oster, G. & Riley, D. P. (1952). *Acta Cryst.* **5**, 272–276.

Pabst, G., Kucerka, N., Nieh, M.-P., Rheinstädter, M. C. & Katsaras, J. (2010). *Chem. Phys. Lipids*, **163**, 460–479.

Pabst, G., Rappolt, M., Amenitsch, H. & Laggner, P. (2000). *Phys. Rev. E*, **62**, 4000–4009.

Rand, R. P., Fuller, N. L., Gruner, S. M. & Parsegian, V. A. (1990). *Biochemistry*, **29**, 76–87.

Rieder, A. A., Koller, D., Lohner, K. & Pabst, G. (2015). *Chem. Phys. Lipids*, **186**, 39–44.

Sivia, D. S. & Skillings, J. H. (2012). *Data Analysis, A Bayesian Tutorial*, 2nd ed. Oxford Science Publications.

Sundblom, A., Oliveira, C. L. P., Palmqvist, A. E. C. & Pedersen, J. S. (2009). *J. Phys. Chem. C*, **113**, 7706–7713.

Székely, P., Ginsburg, A., Ben-Nun, T. & Raviv, U. (2010). *Langmuir*, **26**, 13110–13129.

Tate, M. W. & Gruner, S. M. (1989). *Biochemistry*, **28**, 4245–4253.

Turner, D. C. & Gruner, S. M. (1992). *Biochemistry*, **31**, 1340–1355.

Vacklin, H., Khoo, B. J., Madan, K. H., Seddon, J. M. & Templer, R. H. (2000). *Langmuir*, **16**, 4741–4748.

Yaghmur, A. & Glatter, O. (2009). *Adv. Colloid Interface Sci.* **147–148**, 333–342.