



Information gain from isotopic contrast variation in neutron reflectometry on protein–membrane complex structures

Frank Heinrich,^{a,b*} Paul A. Kienzle,^b David P. Hoogerheide^b and Mathias Lösche^{a,b,c}

Received 19 February 2020

Accepted 21 April 2020

Edited by F. Meilleur, Oak Ridge National Laboratory, USA, and North Carolina State University, USA

Keywords: neutron reflectometry; protein–membrane complex; information content; experimental optimization; marginal posterior entropy.

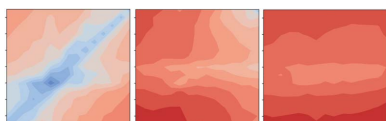
Supporting information: this article has supporting information at journals.iucr.org/j

^aDepartment of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA, ^bCenter for Neutron Research, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, Maryland 20899-6102, USA, and ^cDepartment of Biomedical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA. *Correspondence e-mail: fheinrich@cmu.edu

A framework is applied to quantify information gain from neutron or X-ray reflectometry experiments [Treece, Kienzle, Hoogerheide, Majkrzak, Lösche & Heinrich (2019). *J. Appl. Cryst.* **52**, 47–59], in an in-depth investigation into the design of scattering contrast in biological and soft-matter surface architectures. To focus the experimental design on regions of interest, the marginalization of the information gain with respect to a subset of model parameters describing the structure is implemented. Surface architectures of increasing complexity from a simple model system to a protein–lipid membrane complex are simulated. The information gain from virtual surface scattering experiments is quantified as a function of the scattering length density of molecular components of the architecture and the surrounding aqueous bulk solvent. It is concluded that the information gain is mostly determined by the local scattering contrast of a feature of interest with its immediate molecular environment, and experimental design should primarily focus on this region. The overall signal-to-noise ratio of the measured reflectivity modulates the information gain globally and is a second factor to be taken into consideration.

1. Introduction

In the past decade, specular neutron reflection has emerged as a powerful technique to determine the structure of membrane-associated proteins at fluid lipid bilayer membranes under conditions that closely resemble the thermodynamic nature of membrane–protein interactions in the cell (Heinrich & Lösche, 2014). This novel application of a well established technique (Russell, 1990) thus addresses an important gap in structural biology where most traditional methods such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and electron microscopy require sample environments that are very different from physiological conditions. Neutron scattering is the method of choice in such experiments because of the absence of beam damage and the ability to change the scattering properties of specific structures within the sample by isomorphous isotopic substitution (Heinrich, 2016). In addition, the function of a protein–membrane complex can be monitored by complementary surface-sensitive characterization techniques. For example, ion transfer across a channel-reconstituted membrane can be quantified via electrochemical impedance spectroscopy and compared with that observed in other functional studies (McGillivray *et al.*, 2009). A functional protein–membrane complex can further be exposed to tertiary reaction partners,



such as biological ligands, and the ensuing protein reorganization can be assessed in sequential measurements (Datta *et al.*, 2011).

Although reflectometry only yields one-dimensional structural profiles at low resolution, integrative structural modeling can provide atomistic three-dimensional models of a protein–membrane complex (Akgun *et al.*, 2013; Shenoy *et al.*, 2012). In such an approach, (partial) high-resolution structural information is provided by X-ray crystallography or NMR, while molecular dynamics (MD) simulation is used to construct a structural model of the membrane-associated protein that is consistent with all experimental techniques. Information from neutron reflectometry (NR) can strongly confine the conformational space available to such a structural model, for example, when converted into a steering potential for the MD simulation (Treece *et al.*, 2020). Having such advanced modeling strategies in mind, we recently developed a framework based on information theory and Bayesian statistics to optimize the design of reflectometry experiments with respect to the information gain (Treece *et al.*, 2019).

The aim of any reflectometry experiment is to improve on prior knowledge about a structure, which can be expressed as a probability density function (PDF) $p(\theta)$ of the parameter vector θ of an underlying structural model. The analysis of a measured neutron or X-ray reflection data set provides $p(\theta|y)$, the posterior PDF of θ given the data y . The measure of the information gain ΔH within the model description is defined as the difference in entropy between these two PDFs, $\Delta H = H(\Theta) - H(\Theta|y)$, with the Shannon entropy $H(\Theta) = -\int p(\theta) \log p(\theta) d\theta$. Information gain is computed for virtual experiments in which we systematically vary experimental design variables for the measurement of representative interfacial structures (Treece *et al.*, 2019).

The major experimental design strategy for increasing the information content in biological NR is optimizing the scattering contrast between different parts of the interfacial surface architecture. In most cases, this is achieved by selective deuteration of a subset of molecular components. In studies of lipid membranes and membrane-associated proteins, the neutron scattering length density (nSLD) is typically varied on three components: (i) the lipid molecules constituting the bilayer membrane, (ii) the protein and (iii) the aqueous bulk solvent that partitions into all regions of the surface structure that are not completely filled with organic material (Heinrich, 2016). Using information theory we investigate optimal deuteration strategies in such situations. In virtual experiments we analyze a simplified one-layer system with features similar to a lipid bilayer that floats in close proximity to a solid support structure (Section 3.1). In Section 3.2, we then discuss more realistic lipid bilayer structures and finally proceed to bilayers with associated protein (Section 3.3). In particular, we investigate empirically whether contrast matching between the bulk solvent and parts of the interfacial architecture that are not of interest to the experimenter, such as the substrate, is efficient. This question is of particular practical importance, as (somewhat simplified) two different contrast strategies are in use: (i) maximizing the overall contrast between all compo-

nents of the sample structure, and (ii) contrast matching parts of the sample that are not of interest and, thus, highlighting only the feature of interest.

2. Methods

2.1. Neutron reflectometry and information gain

A large body of experimental work established the general structure of substrate-supported bilayers and their physical properties (Budvytyte *et al.*, 2013; Fragneto, 2012; Knoll *et al.*, 2008; Shenoy *et al.*, 2010; Wacklin, 2010). Because it is generally not possible to directly invert the reflectivity from interfaces into a structural profile, such data are often analyzed in terms of a parameterized real-space model $\hat{y} = X_{S,E}(\theta)$. The model X reflects both structure (S) and experimental setup (E) and relates the expected (noise-free) reflectivity \hat{y} to a particular parameter vector $\theta \in \Theta$. Because neutron scattering does not damage the sample, it is common practice to perform a series of related measurements under a sequence of experimental conditions, for example, by measuring a lipid membrane before and after incubation with a protein. The obtained series of measurement results y (with statistical noise) can then be evaluated in a global context by sharing values of parameters conserved between the structural representations of individual measurements while refining other parameter values separately for each representation (Heinrich & Lösche, 2014).

In general, such data analysis seeks to find the posterior PDF $p(\theta|y)$, which quantifies the probability by which a parameter vector θ is realized given the data y and model $X_{S,E}$. A sample of the posterior PDF is obtained via differential evolution Monte Carlo Markov chain (MCMC) model fitting of the experimental data (Maranville *et al.*, 2016). MCMC requires a prior PDF $p(\theta)$ that represents the knowledge about the model parameters before the experiment. We assume that the prior PDF is uniform over parameter intervals of width $\Delta\theta_i$, with $p(\theta_i) = 1/\Delta\theta_i$ within $\Delta\theta_i$ and $p(\theta_i) = 0$ elsewhere, and that parameters are uncorrelated. Thereby, $p(\theta) = \prod p(\theta_i) = \prod 1/\Delta\theta_i$.

The information gain $\Delta\hat{H}$ from an experiment is defined as the difference between the entropies of the prior and posterior PDFs, taken as a weighted average over all realizations of the experimental data $y \in Y$ according to the probability of each experimental realization $p(y)$. This is equivalent to the mutual information $I(Y, \Theta)$ between Y and Θ ,

$$I(Y, \Theta) = \Delta\hat{H} = H(\Theta) - H(\Theta|Y). \quad (1)$$

In practical terms, $I(Y, \Theta)$ is approximated by averaging 10–20 computations of ΔH for independently simulated data y (Treece *et al.*, 2019):

$$\Delta H = H(\Theta) - H(\Theta|y). \quad (2)$$

To optimize an experimental design, the information gain ΔH is computed for simulated experiments from structures of interest under systematic variation of instrument and sample configurations. As discussed in detail by Treece *et al.* (2019),

experiments are simulated for one representative sample vector θ , instead of drawing a random vector from the prior PDF for each simulation. We therefore focus on optimizing the experimental design for a relatively well known sample structure and not for a sample about which little prior information exists. Statistical noise in the simulated experimental data is assumed to originate from counting statistics only. The examples presented here were computed to represent the current generation of neutron reflectometers at the NIST Center for Neutron Research (NCNR), specifically, the Magik reflectometer (Dura *et al.*, 2006). However, the methodology is straightforward to adapt to other neutron or X-ray instruments, and even future instrumentation whose design parameters are known.

While $H(\Theta)$ can be calculated analytically, the computation of $H(\Theta|y)$ requires numerical methods. An unnormalized sample of the posterior PDF $p(\Theta|y)$ is obtained from an MCMC simulation implemented in *ReflID* (Kirby *et al.*, 2012) with the simulated data y and the model $X_{S,E}(\theta)$ as inputs. $H(\Theta|y)$ is obtained from Monte Carlo sampling over the unnormalized MCMC output (sample size: 5000), for which the normalization factor is determined by integration over a Gaussian mixture model (GMM) (Hastie *et al.*, 2009). This approach is equivalent to the KDN approximation (Silverman, 1986) used earlier (Treece *et al.*, 2019), except that the Gaussian kernel density estimate of the posterior PDF has been replaced by a GMM estimate. Details on determining $H(\Theta|y)$ are provided in the supporting information.

2.2. Marginal entropy of the posterior PDF

In previous work, we computed ΔH associated with the entire parameter vector Θ of a model (Treece *et al.*, 2019) and quantified its dependency on experimental design variables such as maximum accessible momentum transfer, counting time and substrate surface structure for simplified model surface architectures. Here, this framework is applied to more realistic architectures of surface-supported lipid membranes and protein–membrane complexes. This implies that we now

must focus on selected parameters relevant to features of interest while marginalizing over other parameters that are required to build a valid model but are not of practical value to the experimenter. Thereby, the vector $\theta = (\lambda, \delta)$ is partitioned into components describing the parameters of interest λ and nuisance parameters δ . As shown in the supporting information, the marginal entropy $H_\lambda(\Theta|y)$ with respect to the parameters of interest can be obtained via Monte Carlo sampling of $\log p(\lambda|y)$ over the sample of the posterior PDF $p(\theta|y)$ obtained by the MCMC optimizer:

$$H_\lambda(\Theta|y) = - \int p(\theta|y) [\log p(\lambda|y)] d\theta. \quad (3)$$

The marginal PDF $p(\lambda|y)$ is approximated by a GMM estimate from a sample of size $n = 5000$.

3. Results

3.1. A simplified test structure

We first demonstrate entropy marginalization on a test structure (Treece *et al.*, 2019) (Fig. 1) that serves as an idealized prelude to the bilayer structures discussed later. A porous layer (5 vol.% porosity) is suspended above an Si wafer and surrounded by aqueous solvent. Distinct from our earlier analysis, the nSLD of the porous layer is systematically varied alongside the nSLD of the aqueous solvent. We varied the solvent and porous layer nSLDs between that of H_2O ($\rho_n \simeq -0.5 \times 10^{-6} \text{ \AA}^{-2}$) and D_2O ($\rho_n \simeq 6.5 \times 10^{-6} \text{ \AA}^{-2}$) in steps of $\Delta\rho_n = 0.5 \times 10^{-6} \text{ \AA}^{-2}$ to identify favorable solvent contrasts as a function of the nSLD of the porous layer. A typical NR experiment protocol for the determination of membrane structures is to perform a series of up to three individual measurements with aqueous solvents of different isotopic compositions, each with a counting time of ~ 6 h and a maximum momentum transfer of $Q_z = 0.25 \text{ \AA}^{-1}$. Accordingly, we quantified the expected information gain in virtual experiments that either attributed 18 h to a single measurement of the structure in a particular solvent or distributed the same time equally between two or three measurements with

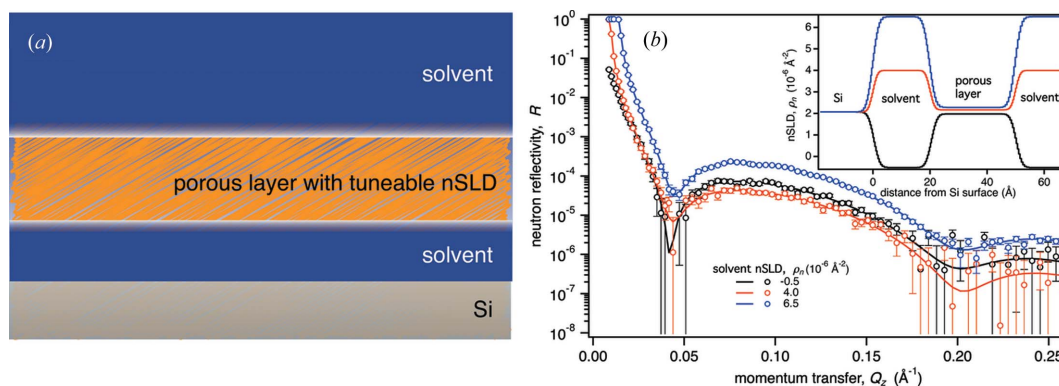


Figure 1

Structural model of a simple test system (a) and outcomes of virtual NR measurements of this structure with isotopically different aqueous solvents (b). A 30 Å-thick porous layer (5% volume porosity) of tuneable nSLD is surrounded by aqueous solvent and suspended at a distance of 20 Å from an Si surface. Pores in the layer are solvent filled. Calculated reflectivities with simulated noise for the NCNR's Magik reflectometer in a configuration where the nSLD of the porous layer matches that of the Si support are shown in panel (b) for three different solvent nSLDs. Also shown are fitted reflectivity curves and their associated nSLD profiles (inset). Figure adapted from the work of Treece *et al.* (2019).

different solvents. The parameterization of the simulated structure is reported in Table 1 together with limits on the model parameters that define the prior PDF. Series of measurements with different solvents were simultaneously analyzed, sharing values of invariant parameters across measurements.

The idealized sample shown in Fig. 1(a) captures the essential features of a lipid bilayer on a solid support without the complications of headgroup structures of a realistic membrane. Here, we investigate how different nSLDs of the surrounding medium and the number of measured reflectivities with distinct isotopic solvent compositions (solvent

contrasts) affect the information gain about the suspended porous layer. The marginal information gain from all parameters that describe this layer – its thickness, material nSLD and completeness – is shown in the upper row of Fig. 2. The lower rows report 68% confidence limits on the parameters of interest, which provide an intuitive insight into individual contributions to the information gain (while neglecting parameter correlations).

The first column in Fig. 2 provides insights into the optimal experimental design for measuring such a minimal system using only one measurement (*i.e.* solvent contrast). The largest information gain of up to 13 bits is obtained for D₂O-based

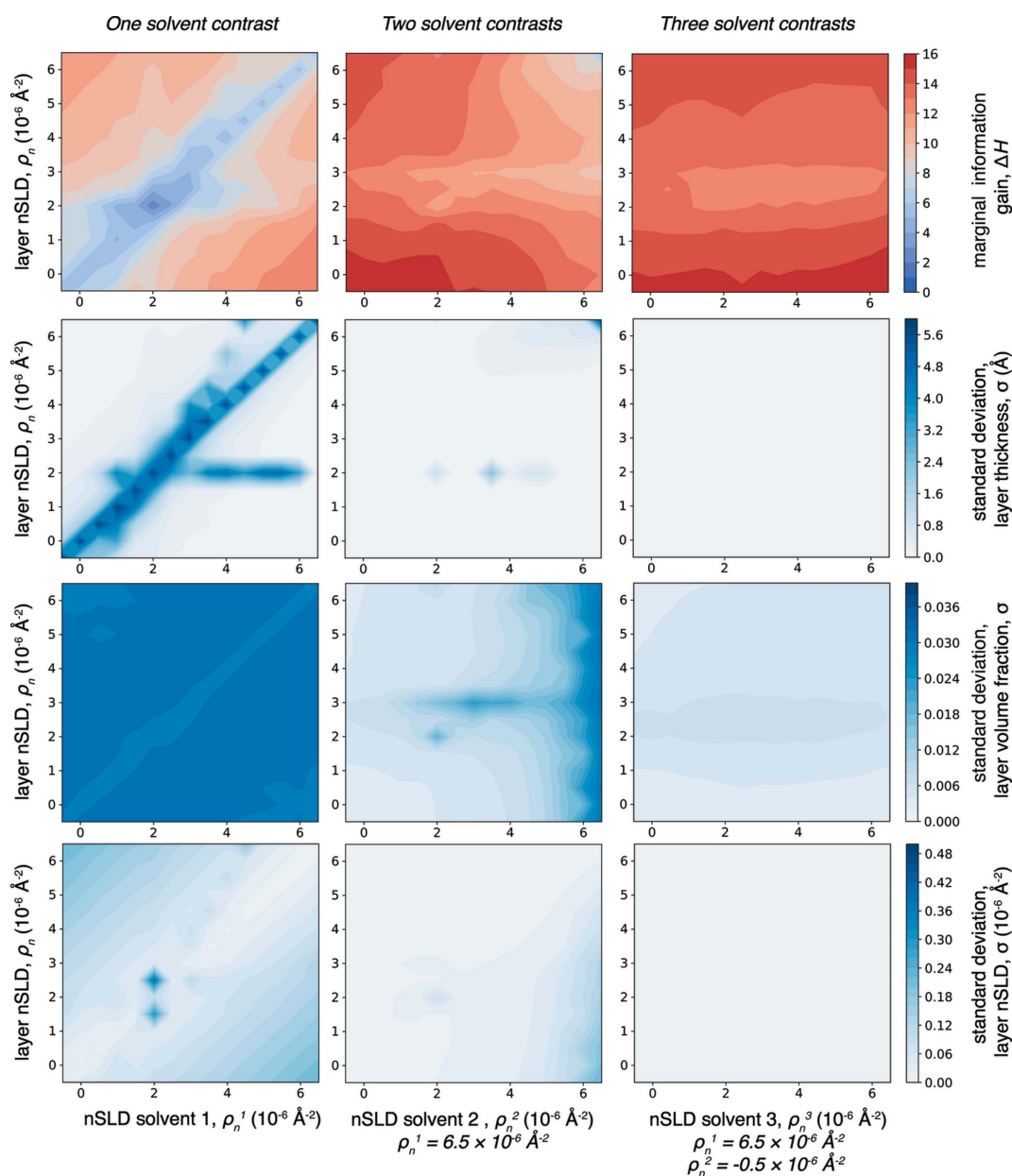


Figure 2

Marginal information gain (top row), and 68% confidence limits on the thickness (second row), volume fraction (third row) and nSLD (fourth row) of the porous layer (Fig. 1). Virtual NR experiments comprise one to three individual measurements, with one bulk solvent (first column), two bulk solvents (second column, one bulk solvent nSLD fixed at $\rho_n = 6.5 \times 10^{-6} \text{ \AA}^{-2}$) and three bulk solvents (third column, a second bulk solvent nSLD fixed at $\rho_n = -0.5 \times 10^{-6} \text{ \AA}^{-2}$). Individual reflectivity curves for experiments with more than one measurement were analyzed simultaneously. The total simulated measurement time is 18 h in all experiments.

Table 1

Simulation parameters for virtual NR experiments of the test system shown in Fig. 1.

Where ranges are given in the first data column, parameters were systematically varied in the optimization. The nSLD of bulk Si is fixed at $\rho_n = 2.07 \times 10^{-6} \text{ \AA}^{-2}$. When the structure was characterized using more than one solvent contrast, each additional measurement carried its own parameter for the solvent nSLD and scattering background.

Model parameter	Parameterized sample representation, θ	Fit boundaries, prior PDF limits
Thickness of interstitial water (\AA)	20	± 10
Thickness of porous layer (\AA)	30	± 10
SLD of porous layer (10^{-6} \AA^{-2})	$[-0.5, 6.5]$	± 1
Volume fraction of porous layer	0.95	± 0.05
SLD of solvent, ρ_n (10^{-6} \AA^{-2})	$[-0.5, 6.5]$	± 0.5
Interfacial roughness (\AA)	3	± 2
Log_{10} of background [†]	-8	± 1

[†] A constant background is routinely fitted as a free parameter to each experimental NR curve, and the same procedure is adopted here. This accounts for insufficient background subtraction during data reduction and is typically 10% or less of the total background.

solvent ($\rho_n \simeq 6.5 \times 10^{-6} \text{ \AA}^{-2}$), except when the layer nSLD approaches that same value. Across the diagonal where the solvent nSLD matches that of the layer material, the information gain is reduced to a few bits. In addition, there is a streak of low information content when the layer nSLD matches that of the silicon substrate ($\rho_n \simeq 2 \times 10^{-6} \text{ \AA}^{-2}$). The lower panels in the first column of Fig. 2 show that the uncertainties on individual structural parameters do not always correlate with information gain: as expected, the uncertainty on the layer thickness is particularly high in regions of low information gain.¹ However, the uncertainty on the layer material nSLD is lower along the diagonal trough for ΔH . This reflects the fact that, near the match points for the layer nSLD with its surrounding, the ensuing lack of scattering carries information. But overall this information gain is dwarfed by losses for other structural parameters.

The general picture changes if one performs a set of two virtual measurements with independent bulk solvents. We assume that one of them was performed with D₂O, in response to the outcome of the single-measurement results. The central column of Fig. 2 shows that almost any choice of solvent nSLD for the second measurement increases the information gain significantly, with the exception of D₂O, which effectively reproduces the single-measurement predictions. ΔH increases from right to left, approaching 16 bits when the difference between the two solvent nSLDs becomes largest. If the layer material has an nSLD of $\rho_n \simeq 3 \times 10^{-6} \text{ \AA}^{-2}$, ΔH is lower than for other values due to a strong correlation between the parameters for the porous layer thickness and nSLD, and nuisance parameters such as the thickness of the interstitial water layer and the global surface roughness (data not shown). A reason for this emergent behavior is difficult to isolate, even in this simple test example. The parameter standard deviations

¹ Note that the diamond structures, in particular along steep ridges or troughs, in the heat plots are artifacts of the two-dimensional interpolation near very sharp features of the map. Their center values reflect accurate computation results but regions surrounding the centers are shifted towards values of the surrounding computed data points.

in the lower panels reveal that the increase in ΔH from performing a two-measurement experiment mainly results from a higher accuracy in determining layer completeness. This result follows the expectation that the solvent content of the porous layer can be precisely determined by measuring two distinct solvents, whereas it remains less precise if one only uses one solvent.

Adding a third measurement with independent solvent (with fixed choices of D₂O and H₂O for the first two) does not further increase the information gain. The shallow horizontal trough in the center of the top panel of the right column in Fig. 2 is carried over from the two-solvent experiment and just better recognized because the left-to-right slope is now equalized. Overall, it is evident that measuring a third contrast is rather ineffective. However, there is a slight advantage in spending 2/3 of the measurement time on H₂O and 1/3 on D₂O.

Under no conditions did we observe that contrast matching of the solid substrate with the bulk solvent was beneficial for information gain. Contrast matching of the Si substrate yielded slightly lower ΔH than the unmatched situations in an experiment using a single solvent nSLD (see vertical ΔH trough in top-left panel of Fig. 2), and virtual experiments using more than one bulk solvent contrast proved rather indifferent towards including a silicon-matched contrast. The common practice of measuring silicon-matched solvent contrasts is therefore not recommended for systems similar to the idealized structure investigated here. As discussed above, Fig. 2 also reveals that not all parameters are best resolved under the same condition. It is therefore helpful to assess confidence limits on parameters of interest simultaneously with marginal entropies.

To further support the conclusion that two measurements are sufficient to maximize information gain, an experiment of three measurements with independently varied bulk solvent nSLDs was performed. The nSLD of the porous layer was fixed to that of the interior of a bilayer ($\rho_n = -0.5 \times 10^{-6} \text{ \AA}^{-2}$). The result shown in Fig. S4 confirms that two bulk solvent contrasts are sufficient to maximize information gain, while measuring a single bulk solvent is insufficient.

3.2. Solid-supported bilayers

Increasing the complexity of the interfacial structure, we now consider a 1,2-dipalmitoyl-2-*sn*-glycero-3-phosphocholine (DPPC) lipid membrane on a planar Si wafer ($\rho_n = 2.07 \times 10^{-6} \text{ \AA}^{-2}$) with a thin silicon oxide surface layer ($\rho_n = 3.55 \times 10^{-6} \text{ \AA}^{-2}$) as a model for a solid-supported bilayer of any lipid composition (Fig. 3). The bilayer is separated from its support by a 2.5 \AA -thick aqueous layer and 95% complete, with membrane-spanning defects lined with lipid headgroups and filled with solvent. Lipid volumes and hydrocarbon thicknesses were taken from X-ray diffraction studies of stacked lipid membranes (Kučerka *et al.*, 2006). We employ a composition-space model of the membrane (Shekhar *et al.*, 2011) that is based on component volume occupancy (CVO) profiles of suitably parsed molecular components along z

[Fig. 3(c)]. This parameterization gives rise to nSLD distributions [inset in Fig. 3(b)] that define the simulated reflectivities [main panel in Fig. 3(b)] (Heinrich & Lösche, 2014; Shekhar *et al.*, 2011). The major structural difference from the test structure in Fig. 1 is that the hydrocarbon core of the lipid bilayer is lined by a solvent-containing layer of headgroups ($\rho_n = 1.86 \times 10^{-6} \text{ \AA}^{-2}$) on either side. These headgroup layers

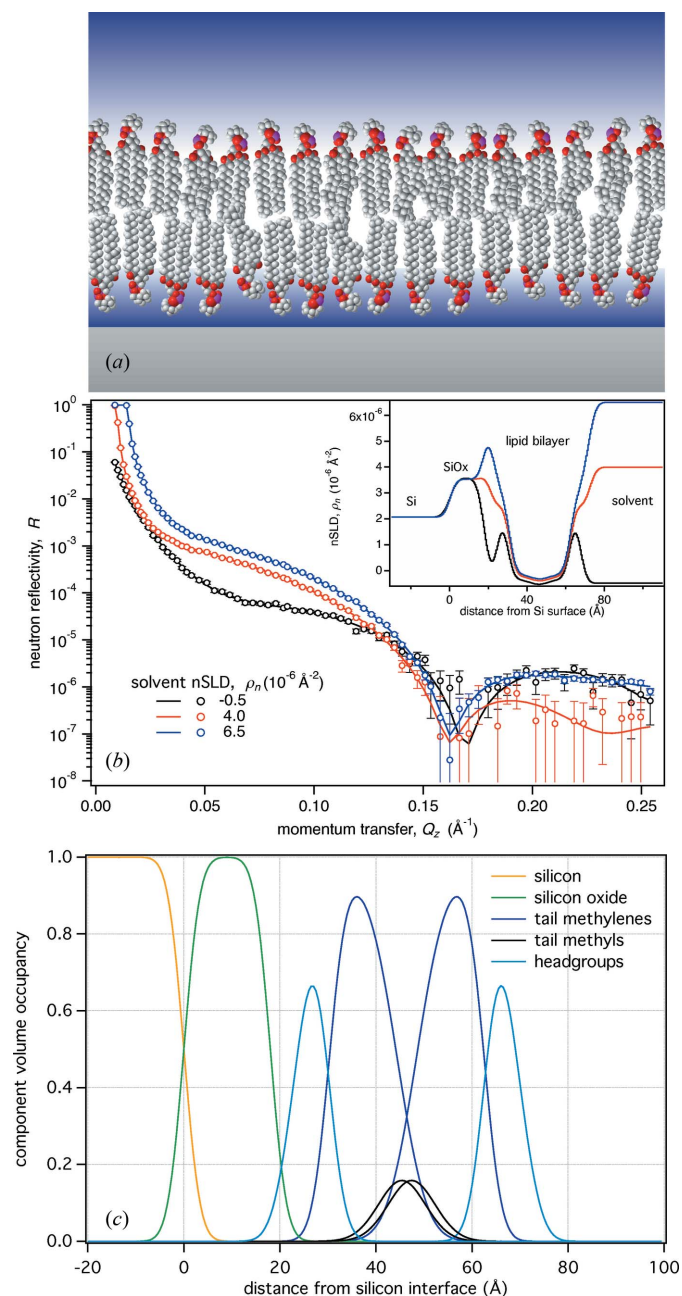


Figure 3
(a) Schematic structure of a DPPC membrane on an oxidized silicon support. (b) Calculated reflectivities with simulated noise for this structure for a virtual experiment on NCNR's Magik reflectometer using tail-protiated DPPC. The experiment consists of three measurements with isotopically different aqueous solvents (see Table 2 for model parameters). Shown also are fitted reflectivity curves and their associated nSLD profiles (inset). (c) Component volume occupancies for the substrate and molecular groups that constitute the lipid bilayer from the fit to the data using a composition-space model.

Table 2

Simulation parameters for virtual NR measurements of the solid-supported lipid bilayer shown in Fig. 3(a), in which the isotopic composition of the bilayer and one bulk solvent nSLD were systematically varied.

Where ranges are given in the first data column, the parameter was systematically varied within these boundaries during the optimization. The nSLD of bulk Si is fixed at $\rho_n = 2.07 \times 10^{-6} \text{ \AA}^{-2}$. If the bilayer was characterized using more than one solvent, each additional measurement carried its own parameter for the solvent nSLD and the scattering background.

Model parameter	Parameterized sample representation, θ	Fit boundaries, prior PDF limits
Thickness of Si oxide (\AA)	18	[10, 30]
nSLD of Si oxide (10^{-6} \AA^{-2})	3.55	[3.2, 3.8]
Thickness of sub-membrane water layer (\AA)	2.5	[1, 10]
Thickness of hydrocarbon chains per bilayer leaflet (\AA)	14.1	[10, 20]
Molar fraction of DPPC-d ₆₂ in the membrane	[0.0, 1.0]	[0.0, 1.0]
Volume fraction of lipid bilayer	0.95	[0.9, 1.0]
nSLD of bulk solvent (10^{-6} \AA^{-2})	[-0.5, 6.5]	± 0.5
Interfacial roughness, σ (\AA)	3	± 2
Bilayer roughness, σ (\AA)	2.5	[2, 5]
Log ₁₀ of background	-8	± 1

provide additional scattering contrast with the hydrocarbon chains and the bulk solvent.

We sought to identify the number and combination of solvent contrasts that maximize the information gain about the bilayer structure from the experiment. The isotopic composition of the bilayer core was continuously varied between tail-protiated DPPC (chain nSLD, $\rho_n \simeq -0.4 \times 10^{-6} \text{ \AA}^{-2}$) and tail-deuterated DPPC-d₆₂ (chain nSLD, $\rho_n \simeq 7.8 \times 10^{-6} \text{ \AA}^{-2}$). Marginal entropy and confidence limits of the membrane-relevant model parameters – thickness, completeness and molar fraction of DPPC-d₆₂ – were computed for one, two and three measurements under different solvent conditions with the same approach as for the porous layer structure discussed above. Here, this translates to the simultaneous variation of the DPPC-d₆₂ fraction and the nSLD of one solvent while keeping the optional second and third solvent constant. All measurements with different solvent contrasts were simultaneously analyzed and their model parameters shared, except for the bulk solvent nSLD and the scattering background (Table 2). The total simulated counting time per experiment was kept constant at 18 h, either attributed to a single measurement with a particular solvent or distributed equally between two or three individual measurements with different solvents.

The marginal information gain from virtual NR experiments with up to three simulated solvent contrasts is shown in Fig. 4, and the results exhibit analogy to those from the previously analyzed test structure (see Fig. 2). The best experimental choice for a single measurement (first column in Fig. 4) is a solvent with high nSLD such as D₂O in combination with either highly deuterated or protiated lipid chains. However, a combination of D₂O with a 1:1 mix of protiated and deuterated chains works poorly, and solvent mixtures with

$\rho_n \simeq 4 \times 10^{-6} \text{ \AA}^{-2}$ or even pure H_2O provide more information. This is due to a well developed trough in ΔH that coincides with a match of the nSLDs of hydrated lipid headgroups and chains. Because the average nSLD of the hydrated lipid headgroup region is sensitive to solvent composition, the matching points shift linearly between 20 mol% DPPC- d_{62} in pure H_2O and 50 mol% DPPC- d_{62} in pure D_2O . A more shallow trough in ΔH is oriented vertically along $\rho_n = 2.07 \times 10^{-6} \text{ \AA}^{-2}$, which corresponds to matching of the Si substrate by the solvent. Inspection of parameter standard deviations shows that ΔH largely follows the ability of the experiment to

determine the leaflet thickness. Membrane volume fraction is ill-resolved by a single measurement, except for a combination of high nSLDs of solvent and lipid chains, which creates a situation in which the headgroups assume a large scattering contrast with their environment. Similarly, the fraction of DPPC- d_{62} in the membrane is determined with reasonable accuracy under those conditions.

The center column of Fig. 4 provides results for an experiment of two measurements with independent solvent nSLDs. As observed for the porous layer model, ΔH increases as the difference in solvent nSLD between the two measurements

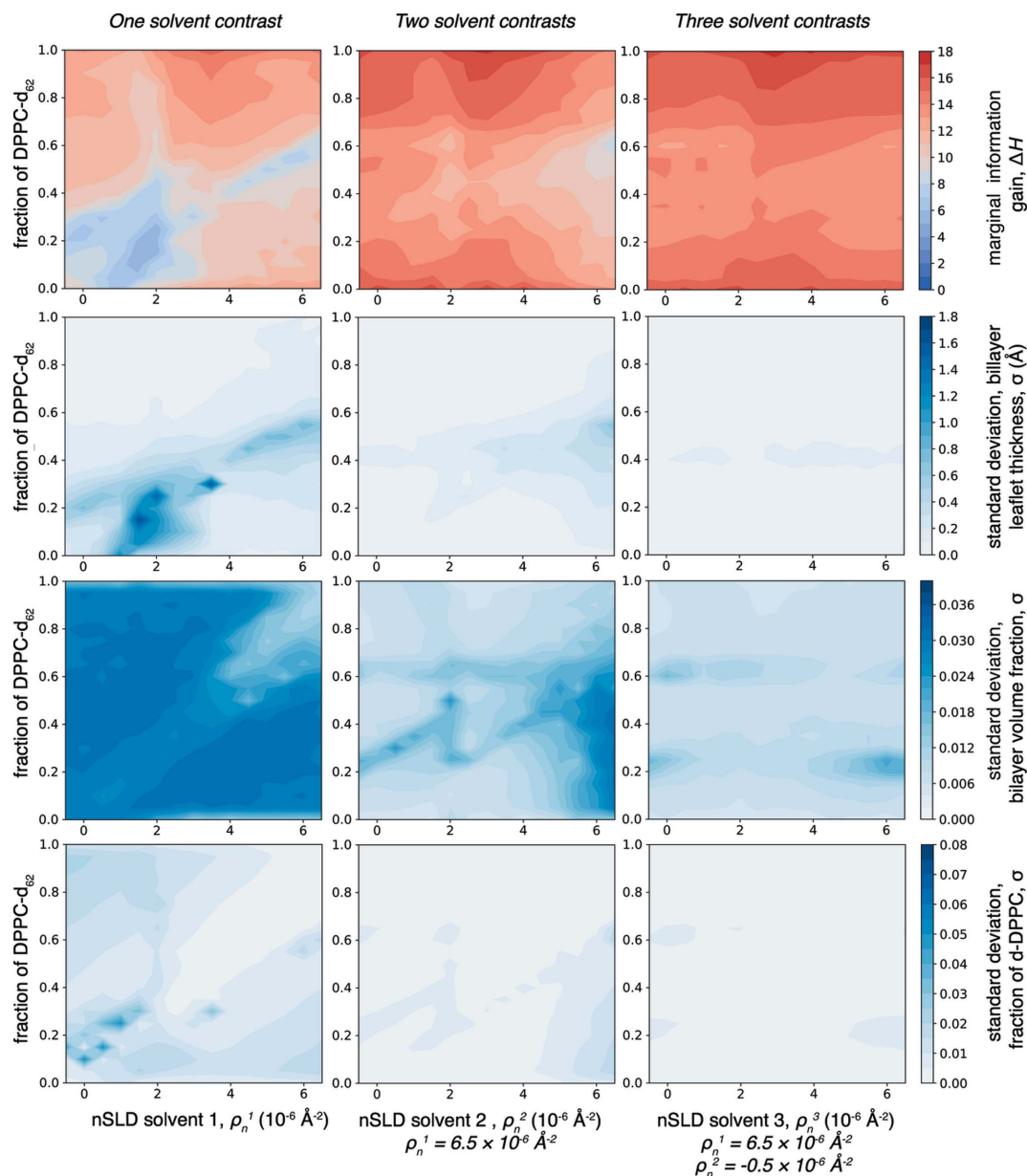


Figure 4

Marginal information gain (top row), and 68% confidence limits on the thickness (second row), volume fraction (third row) and molar fraction of DPPC- d_{62} (fourth row) of a solid-supported DPPC bilayer membrane (Fig. 3). Shown are optimization results for virtual NR experiments employing one bulk solvent nSLD (first column), two bulk solvent contrasts (second column, one bulk solvent nSLD fixed at $\rho_n = 6.5 \times 10^{-6} \text{ \AA}^{-2}$) and three bulk solvent contrasts (third column, a second bulk solvent nSLD fixed at $\rho_n = -0.5 \times 10^{-6} \text{ \AA}^{-2}$). Reflectivity curves for experiments with more than one solvent contrast were analyzed simultaneously. The total simulated measurement time was 18 h for all three types of experiments. nSLD values of the hydrocarbon chains of the DPPC bilayer scale linearly with the molar fraction of DPPC- d_{62} from $\rho_n = -0.4 \times 10^{-6} \text{ \AA}^{-2}$ for 0 mol% DPPC- d_{62} and $\rho_n = 7.8 \times 10^{-6} \text{ \AA}^{-2}$ for 100 mol% DPPC- d_{62} . Thereby, the hydrocarbon nSLD matches that of D_2O at approximately 75 mol% DPPC- d_{62} .

grows. Matching the solvent nSLD and that of the Si substrate ($\rho_n = 2.07 \times 10^{-6} \text{ \AA}^{-2}$) affects ΔH negatively. Performing a third measurement (right column in Fig. 4) increases ΔH by $\sim 1/2$ bit for a bulk solvent nSLD in the range of $\rho_n = (3 \pm 1) \times 10^{-6} \text{ \AA}^{-2}$. Similarly to the porous layer model, there is a trough of low ΔH at 45 mol% DPPC- d_{62} for experiments with two and three measurements, which corresponds to a chain nSLD of $\rho_n \simeq 3 \times 10^{-6} \text{ \AA}^{-2}$. Additional regions of low ΔH are visible in the standard deviation plots for the bilayer volume fraction at 20 and 60 mol% DPPC- d_{62} . As discussed for the porous layer, it is difficult to identify the cause for these increased uncertainties. Overall, Fig. 4 leads to the conclusion that, for maximizing ΔH on the bilayer structure, two independent measurements at high and low solvent nSLDs are sufficient. A slight increase in information content can be achieved by adding a third measurement with an intermediate solvent nSLD. Matching of the substrate and solvent nSLDs is not advantageous for information gain.

A representative experiment that successively uses three bulk solvent contrasts was simulated under systematic variation of all three solvent nSLDs (Fig. S5) for a lipid bilayer with protiated lipid ($\rho_n \simeq -0.5 \times 10^{-6} \text{ \AA}^{-2}$). In agreement with the results in Fig. 4, measurements with two bulk solvent nSLDs ($\rho_n = -0.5 \times 10^{-6} \text{ \AA}^{-2}$ and $6.5 \times 10^{-6} \text{ \AA}^{-2}$) are sufficient to gain most of the available information (~ 16 bits), while measuring a single bulk solvent contrast is insufficient. An additional $1/2$ bit in information is gained by adding a third bulk solvent nSLD of $\rho_n = 3.0 \times 10^{-6} \text{ \AA}^{-2}$.

3.3. Protein structure at membranes and the utility of protein deuteration

There is a growing body of NR studies determining the structure of protein–membrane complexes (Akgun *et al.*, 2013; Heinrich & Lösche, 2014; Hoogerheide *et al.*, 2017; Rondelli *et al.*, 2018; Sani *et al.*, 2020; Shenoy *et al.*, 2012; Wacklin *et al.*, 2016; Yap *et al.*, 2015). Due to large variations in protein size, conformational flexibility and mode of membrane binding, general rules concerning the expected information gain from such measurements are difficult to devise. For a compact, medium-sized protein ($\sim 40 \text{ \AA}$ diameter), we determine how isotopic variation of the solvent, the bilayer and the protein affects information gain for different penetration depths of the protein. In Section 3.2 we concluded that the major part of recoverable information on the structure of the naked bilayer is obtained in a combination of measurements at high and low solvent nSLD. Since a protein in its membrane-associated state typically occupies not more than 20% of the available in-plane area, the reflectivity from the protein–membrane complex is dominated by the bilayer structure and only perturbed by the protein. Consequently, we simulate NR measurements that interrogate the structures of the naked and the protein-associated bilayer successively in D_2O and H_2O bulk solvents. In addition, the nSLDs of the lipid hydrocarbon chains and the protein are systematically varied. The DPPC bilayer supported by an Si substrate discussed in the previous section (Fig. 3) was used as a representative membrane

structure. The protein was placed at three different penetration depths with respect to the membrane: peripheral, inserted into the lipid headgroup region and deeply inserted into the lipid hydrocarbon core, as shown in Fig. 5.

Virtual NR experiments of the as-prepared and protein-decorated bilayer were simulated, comprising two measurements per condition with bulk solvent nSLDs of $\rho_n = -0.5 \times 10^{-6}$ and $6.5 \times 10^{-6} \text{ \AA}^{-2}$. The nSLD of the hydrocarbon membrane core was tuned by changing the fraction of DPPC- d_{62} in the bilayer between 0 and 100 mol%. The nSLD of protein in H_2O solvent was varied between $\rho_n = 2.0 \times 10^{-6}$ and $6.0 \times 10^{-6} \text{ \AA}^{-2}$, *i.e.* the range between fully protiated and fully deuterated protein. The exchange of labile hydrogens in D_2O -containing solvents was taken into account.

Fig. 6 shows that NR measurements are most informative for membrane-peripheral proteins and that proper contrast engineering becomes ever more important as the protein protrudes deeper into the bilayer. Not only is the marginal information gain ΔH largest for peripherally localized protein, but the ΔH landscape is also rather flat, with a difference between maximum and minimum information gain of only 5 bits for any combination of protein and bilayer nSLD (top-right panel of Fig. 6). In comparison, this difference increases to 10 bits if the protein penetrates the membrane substantially (top-left and center panels of Fig. 6). A deep trough is observed diagonally along the locations where the nSLDs between the membrane interior and the protein match, for example at 30 mol% DPPC- d_{62} in the bilayer for an experiment with fully protiated protein ($\rho_n = 2.0 \times 10^{-6} \text{ \AA}^{-2}$). The cause of this trough is rather easily identified upon inspection of the standard deviations associated with the

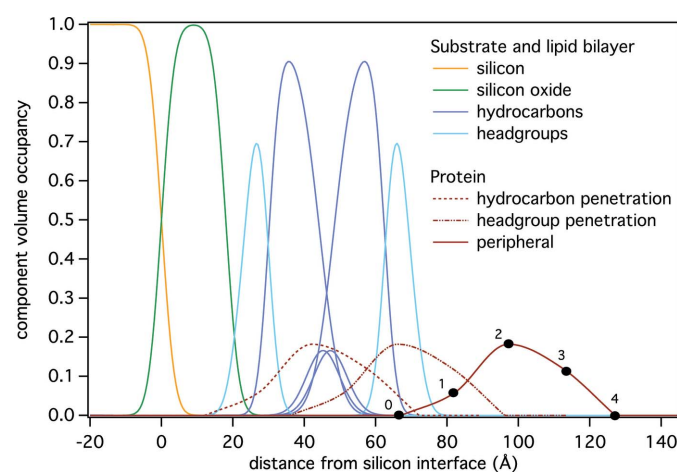


Figure 5

Model structures of a protein–membrane complex in which the protein penetrates the bilayer at different depths. In three separate optimizations, a generic protein CVO profile is parameterized by a Hermite spline (control points: numbered black dots) and is placed at three different membrane penetration depths: membrane-peripheral, penetrating the substrate-distal lipid headgroups and penetrating deeply into the hydrocarbon core of the bilayer. The bilayer structure adapts seamlessly to the protein, such that overfilling of the space does not occur. The submolecular components of the lipid bilayer (lipid headgroups, chain polymethylene and chain-end methyl groups) are shown for the protein-free membrane.

protein volume occupancies at each control point (second to fourth rows in Fig. 6). To a first approximation, every control point contributes independently to ΔH . If a control point is located in the hydrocarbon region of the lipid bilayer, the uncertainty on the volume occupancy is large when the protein and hydrocarbon nSLDs match, which affects ΔH negatively. If a control point is located in the highly hydrated headgroup region, this correlation is much weaker, but matching protein, hydrocarbon and headgroup nSLDs ($\rho_n = 1.9 \times 10^{-6} \text{ \AA}^{-2}$) is highly unfavorable. Volume occupancies of control points outside the lipid bilayer are well resolved independent of protein or hydrocarbon nSLD, generally showing only a weak advantage of high protein and hydrocarbon nSLDs.

In summary, for a protein embedded in the lipid bilayer, high nSLD contrast to its immediate surrounding (hydrocarbon or headgroup) is beneficial. This can be achieved by fully deuterating either the lipid chains or the protein, but not both. At the same time, sufficient contrast with lipid chains and headgroups should be established, so it is generally advantageous to deuterate the protein, thereby increasing the contrast with the lipid headgroups, rather than the lipid chains. Only for extra-membranous protein, where a well defined hydrated layer provides contrast between both the lipid chains and the protein, should the highest nSLD of protein and hydrocarbon chains available be chosen to maximize ΔH . However, deviating from those rules sacrifices only a

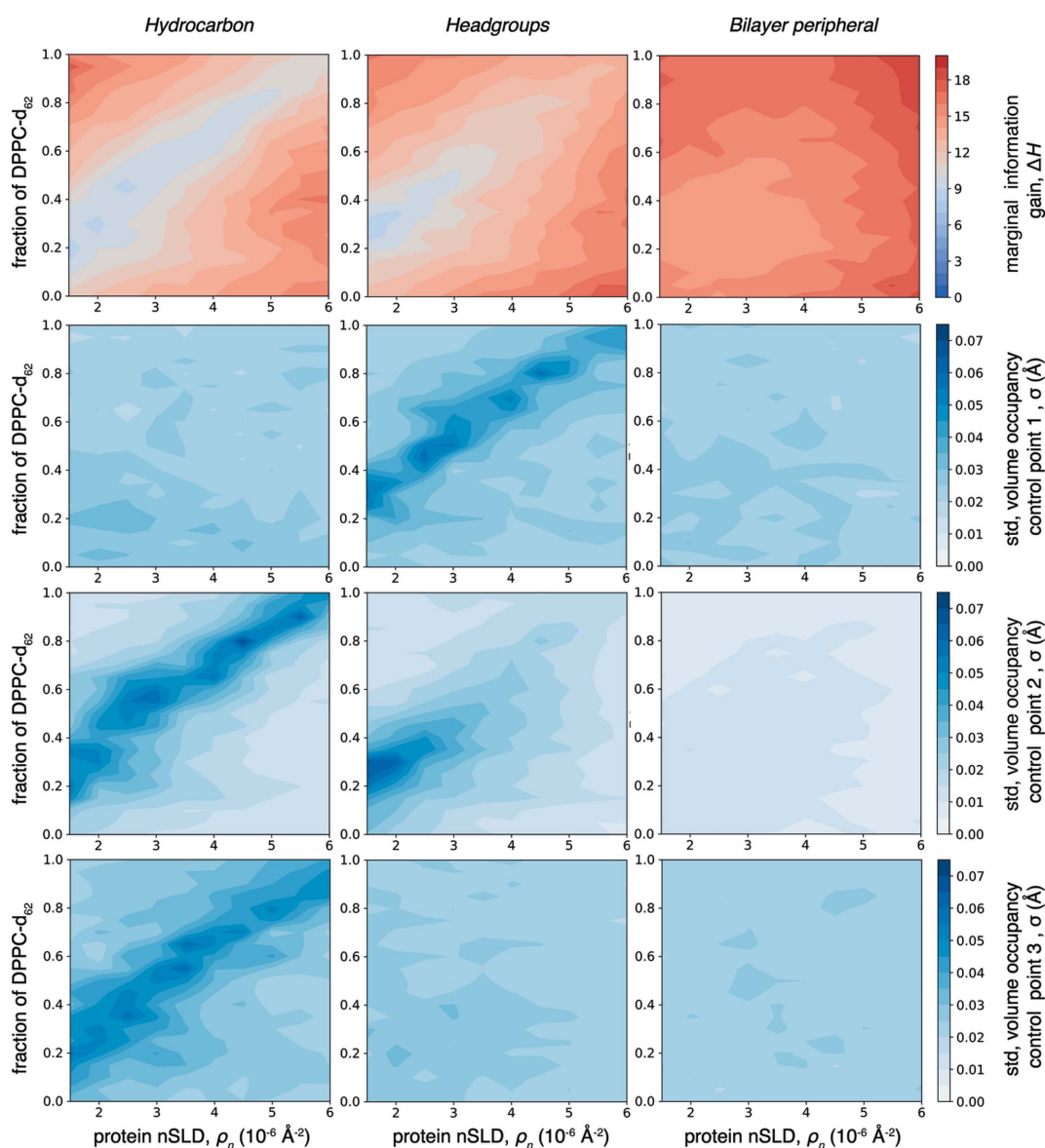


Figure 6

Marginal information gain (top row), and 68% confidence limits on the volume occupancies of the protein associated with control points 1–3 of the protein Hermite spline (second to fourth rows). Shown are optimization results for virtual NR experiments using two bulk solvent nSLDs ($\rho_n = 6.5 \times 10^{-6}$ and $-0.5 \times 10^{-6} \text{ \AA}^{-2}$). The total simulated measurement time is 18 h. nSLD values of the hydrocarbon chains of the DPPC bilayer scale linearly with the molar fraction of DPPC-d₆₂ from $\rho_n = -0.44 \times 10^{-6} \text{ \AA}^{-2}$ for 0 mol% DPPC-d₆₂ to $\rho_n = 7.77 \times 10^{-6} \text{ \AA}^{-2}$ for 100 mol% DPPC-d₆₂; therefore the nSLD of the hydrocarbon chains matches that of D₂O ($\rho_n \simeq 6.5 \times 10^{-6} \text{ \AA}^{-2}$) at approximately 75 mol% DPPC-d₆₂.

moderate amount of information as long as match points of protein and bilayer material are avoided by at least $\Delta\rho_n = 1 \times 10^{-6} \text{ \AA}^{-2}$.

4. Discussion and conclusion

A framework to quantify the information gain from surface-sensitive scattering of neutrons or X-rays (Treece *et al.*, 2019) has been applied to assess the utility of contrast matching in neutron reflectometry from biological model interfaces. Neutron reflectometry is routinely harnessed to address problems in structural biology that are difficult to resolve by other techniques. While we focused our discussion on the characterization of lipid bilayers and bilayer-associated proteins, the results of this study are applicable to a large number of related problems with impact in membrane biology, the physics of self-organizing molecular systems, the design of inert surfaces in biomedical engineering, or the optimization of detergent systems in chemical engineering.

The characterization of interfacial molecular structures by surface-sensitive neutron scattering has entirely been driven by heuristic rules of thumb since its inception about three decades ago (Johnson *et al.*, 1991; Russell, 1990; Vaknin *et al.*, 1991). Our analysis confirms some of those while it rejects others as misconceived. Specifically, matching a substrate nSLD with the solvent (Johnson *et al.*, 1991; Penfold *et al.*, 1997), probably derived from the practice of contrast matching in small-angle neutron scattering, does not universally contribute to information gain in NR experiments.

Condensing the results for the specific cases investigated in this study leads to a rather simple proposition for a set of rules for optimal experimental design with respect to the nSLD contrast between molecular components. Similar rules will apply to comparable soft-matter interfaces that concern, for example, polymers or detergents, as from a technical standpoint there is nothing peculiar to membrane-associated proteins except a specific range of nSLD values that proteins and membranes can attain depending on their isotopic makeup.

(i) To resolve structural features in a surface architecture, the local scattering contrast with neighboring structures is the determining factor. According to the results described above for current NR instrumentation, the difference in nSLD between adjacent molecular species within the sample should exceed $\Delta\rho_n \simeq 1 \times 10^{-6} \text{ \AA}^{-2}$ in at least one of the sets of measurements within an NR experiment, with better resolution achieved for larger differences. We have not investigated how $\Delta\rho_n$ depends on the thickness of the structural features that create the nSLD contrast. All investigated features, however, have sizes above the canonical resolution limit of the measurement and are typically on the order of 15–30 Å.

(ii) For solvent-immersed surface architectures, NR experiments comprising two measurements with different bulk solvents are sufficient to gain most information. The nSLD values of the two solvents should be maximally different, which is typically achieved by using H₂O- and D₂O-based solvents. The use of a third measurement with an intermediate

bulk solvent nSLD only incrementally increases information gain, making it an effective strategy only for specific problems.

(iii) Matching the substrate nSLD in one of the NR measurements that constitute the overall experimental design is not beneficial, not even when it constitutes just an additional third measurement.

Rules (ii) and (iii) can be regarded as derivatives from rule (i), which reduces the main proposition of this work to a rule of locality that states that the local nSLD environment mostly determines the ability of a measurement to resolve a structural feature. In addition, an overall high signal-to-noise ratio of the measured reflectivity is favorable, and configurations with low reflectivity such as those in which the difference between the nSLDs of the substrate and the bulk solvent is small should be avoided. Since the above set of rules has been deduced from a systematic set of simulations, they lack the general applicability of a purely theoretical approach and are ultimately tied to the bounds of the simulated structures. Future work that combines information theory and scattering theory might achieve a more general insight into this problem.

5. Related literature

The following additional references are cited in the supporting information: Chen *et al.* (2016), Kramer *et al.* (2010), Seabold & Perktold (2010).

Acknowledgements

We thank Drs Markus Deserno and David L. Worcester for critical discussions. Certain commercial materials, equipment, and instruments are identified in this work to describe the experimental procedure as completely as possible. In no case does such an identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials, equipment, or instruments identified are necessarily the best available for the purpose.

Funding information

This work was supported by the US Department of Commerce (grant Nos. 70NANB13H009 and 70NANB17H299). Will Golding implemented the initial entropy calculations within the MCMC optimization software and was supported by the Center for High Resolution Neutron Scattering (CHRNS), a partnership between the National Institute of Standards and Technology and the National Science Foundation under agreement No. DMR-1508249. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant No. ACI-1053575. Specifically, it used the Bridges system, which is supported by NSF award No. ACI-1445606, at the Pittsburgh Supercomputing Center (PSC) (Townes *et al.*, 2014).

References

- Akgun, B., Satija, S., Nanda, H., Pirrone, G. F., Shi, X., Engen, J. R. & Kent, M. S. (2013). *Structure*, **21**, 1822–1833.

- Budvytyte, R., Mickevicius, M., Vanderah, D. J., Heinrich, F. & Valincius, G. (2013). *Langmuir*, **29**, 4320–4327.
- Chen, B., Wang, J., Zhao, H. & Principe, J. C. (2016). *Entropy*, **18**, 196.
- Datta, S. A. K., Heinrich, F., Raghunandan, S., Krueger, S., Curtis, J. E., Rein, A. & Nanda, H. (2011). *J. Mol. Biol.* **406**, 205–214.
- Dura, J. A., Pierce, D. J., Majkrzak, C. F., Maliszewskyj, N. C., McGillivray, D. J., Lösche, M., O'Donovan, K. V., Mihailescu, M., Perez-Salas, U., Worcester, D. L. & White, S. H. (2006). *Rev. Sci. Instrum.* **77**, 074301.
- Fragneto, G. (2012). *Eur. Phys. J. Spec. Top.* **213**, 327–342.
- Hastie, T., Tibshirani, R. & Friedman, J. H. J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Heinrich, F. (2016). *Methods Enzymol.* **566**, 211–230.
- Heinrich, F. & Lösche, M. (2014). *Biochim. Biophys. Acta*, **1838**, 2341–2349.
- Hoogerheide, D. P., Noskov, S. Y., Jacobs, D., Bergdoll, L., Silin, V., Worcester, D. L., Abramson, J., Nanda, H., Rostovtseva, T. K. & Bezrukov, S. M. (2017). *Proc. Natl Acad. Sci. USA*, **114**, E3622–E3631.
- Johnson, S. J., Bayerl, T. M., McDermott, D. C., Adam, G. W., Rennie, A. R., Thomas, R. K. & Sackmann, E. (1991). *Biophys. J.* **59**, 289–294.
- Kirby, B. J., Kienzle, P. A., Maranville, B. B., Berk, N. F., Krycka, J., Heinrich, F. & Majkrzak, C. F. (2012). *Curr. Opin. Colloid Interface Sci.* **17**, 44–53.
- Knoll, W., Naumann, R., Friedrich, M., Robertson, J. W. F., Lösche, M., Heinrich, F., McGillivray, D. J., Schuster, B., Gufler, P. C., Pum, D. & Sleytr, U. B. (2008). *Biointerphases*, **3**, FA125.
- Kramer, A., Hasenauer, J., Allgöwer, F. & Radde, N. (2010). *2010 IEEE International Conference on Control Applications*, pp. 493–498. Piscataway: IEEE.
- Kučerka, N., Tristram-Nagle, S. A. & Nagle, J. F. (2006). *Biophys. J.* **90**, L83–L85.
- Maranville, B. B., Kirby, B. J., Grutter, A. J., Kienzle, P. A., Majkrzak, C. F., Liu, Y. & Dennis, C. L. (2016). *J. Appl. Cryst.* **49**, 1121–1129.
- McGillivray, D. J., Valincius, G., Heinrich, F., Robertson, J. W. F., Vanderah, D. J., Febo-Ayala, W., Ignatjev, I., Lösche, M. & Kasianowicz, J. J. (2009). *Biophys. J.* **96**, 1547–1553.
- Penfold, J., Richardson, R. M., Zarbakhsh, A., Webster, J., Bucknall, D. G., Rennie, A. R., Jones, R. A. L., Cosgrove, T., Thomas, R. K., Higgins, J. S., Fletcher, P. D. I., Dickinson, E., Roser, S. J., McLure, I. A., Hillman, A. R., Richards, R. W., Staples, E. J., Burgess, A. N., Simister, E. A. & White, J. W. (1997). *Faraday Trans.* **93**, 3899–3917.
- Rondelli, V., Del Favero, E., Brocca, P., Fragneto, G., Trapp, M., Mauri, L., Ciampa, M. G., Romani, G., Braun, C. J., Winterstein, L., Schroeder, I., Thiel, G., Moroni, A. & Cantu', L. (2018). *Biochim. Biophys. Acta*, **1862**, 1742–1750.
- Russell, T. P. (1990). *Mater. Sci. Rep.* **5**, 171–271.
- Sani, M.-A., Le Brun, A. P. & Separovic, F. (2020). *Biochim. Biophys. Acta*, **1862**, 183204.
- Seabold, S. & Perktold, J. (2010). *Proceedings of the 9th Python in Science Conference*, pp. 57–61. SciPy Conferences.
- Shekhar, P., Nanda, H., Lösche, M. & Heinrich, F. (2011). *J. Appl. Phys.* **110**, 102216.
- Shenoy, S., Moldovan, R., Fitzpatrick, J., Vanderah, D. J., Deserno, M. & Lösche, M. (2010). *Soft Matter*, **6**, 1263–1274.
- Shenoy, S. S., Nanda, H. & Lösche, M. (2012). *J. Struct. Biol.* **180**, 394–408.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London, New York: Chapman and Hall.
- Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G. D., Roskies, R., Scott, J. R. & Wilkins-Diehr, N. (2014). *Comput. Sci. Eng.* **16**, 62–74.
- Treece, B. W., Heinrich, F., Ramanathan, A. & Lösche, M. (2020). *J. Chem. Theory Comput.* <http://doi.org/10.1021/acs.jctc.0c00136>.
- Treece, B. W., Kienzle, P. A., Hoogerheide, D. P., Majkrzak, C. F., Lösche, M. & Heinrich, F. (2019). *J. Appl. Cryst.* **52**, 47–59.
- Vaknin, D., Als-Nielsen, J., Piepenstock, M. & Lösche, M. (1991). *Biophys. J.* **60**, 1545–1552.
- Wacklin, H. P. (2010). *Curr. Opin. Colloid Interface Sci.* **15**, 445–454.
- Wacklin, H. P., Bremec, B. B., Moulin, M., Rojko, N., Haertlein, M., Forsyth, T., Anderluh, G. & Norton, R. S. (2016). *Biochim. Biophys. Acta*, **1858**, 640–652.
- Yap, T. L., Jiang, Z., Heinrich, F., Gruschus, J. M., Pfefferkorn, C. M., Barros, M., Curtis, J. E., Sidransky, E. & Lee, J. C. (2015). *J. Biol. Chem.* **290**, 744–754.