Structure and dynamics of a lipid nanodisc by integrating NMR, SAXS and SANS experiments with molecular dynamics simulations

Tone Bengtsen[a], Viktor L. Holm[b], Lisbeth Ravnkilde Kjølbye[c,d], Søren R. Midtgaard[b], Nicolai Tidemand Johansen[b], Sandro Bottaro[a], Birgit Schiøtt[c], Lise Arleth[b,1], and Kresten Lindorff-Larsen[a,1]

[a]Structural Biology and NMR Laboratory & Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Copenhagen, Denmark; [b]Structural Biophysics, X-ray and Neutron Science, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark; [c]Department of Chemistry, Aarhus University, Aarhus C, Denmark; [d]Novo Nordisk A/S, Måløv, Denmark

This manuscript was compiled on February 2, 2020

Nanodiscs are membrane mimetics that consist of a protein belt surrounding a lipid bilayer. Nanodiscs are of broad use for solution-based characterization of membrane proteins, but we lack a full understanding of their structure and dynamics. Recently, NMR/EPR experiments provided a view of the average structure of the nanodisc's protein component, while small-angle X-ray and neutron scattering have provided insight into the global structure of both protein and lipids. Here, we investigate the structure, dynamics and biophysical properties of two small nanodiscs, MSP1D1ΔH5 and ΔH4H5. We combine SEC-SAXS and SEC-SANS to obtain low-resolution structures of the nanodiscs as elliptical discs. These are in apparent contrast to the NMR/EPR structure, which showed a more circular conformation. We reconcile these views using a Bayesian/Maximum Entropy method to combine our molecular dynamics simulations of MSP1D1ΔH5 with the NMR and SAXS experiments. We derive a conformational ensemble that represents the structure and dynamics of the nanodisc, and find that it displays conformational heterogeneity with various elliptical shapes. We find substantial differences in lipid ordering in the centre and rim of the discs, and support these observations using differential scanning calorimetry of nanodiscs of various sizes. We find that the order and phase transition of the lipids is highly dependent on the nanodisc size. Together, our results demonstrate the power of integrative modelling and paves the way for future studies of larger complex systems such as membrane proteins embedded in nanodiscs.

Nanodisc | Molecular Dynamics | SAXS | SANS | NMR | DSC

Nanodiscs are widely used membrane models that facilitate biophysical studies of membrane proteins(1). They are derived from, and very similar to, the human ApoA1 protein from high density lipoproteins (HDL particles) and consists of two amphipatic membrane scaffold proteins (MSPs) that stack and encircle a small patch of lipids in a membrane bilayer to form a discoidal assembly. The popularity of nanodiscs arises from their ability to mimic a membrane while at the same time ensuring a small system of homogeneous composition, the size of which can be controlled and can give diameters in a range from about 7 to 13 nm(2, 3).

Despite the importance of nanodiscs in structural biology research and the medical importance of HDL particles, we still lack detailed structural models of these protein-lipid particles. The nanodisc has so far failed to crystallize, so in order to study its structure and dynamics a range of biophysical methods have been used to provide information about specific characteristics. For example, mass spectrometry experiments have provided insight into lipid-water interactions and heterogeneous lipid compositions(4, 5), solid state NMR have provided insights into lipid phase transition states and lipid order (6, 7) and small angle X-ray scattering (SAXS) and -neutron scattering (SANS) have provided insight into the size and low resolution shape of nanodiscs in solution(2, 8–10). These experiments have been complemented by molecular dynamics (MD) simulations that provided both pioneering insights into the structure (11, 12) as well as a better understanding of the assembly process, lipid-protein interactions and membrane mimicking effects (13–15).

Recently, the first high resolution experimental structure of the MSP protein belt encircling the nanodisc was obtained from the small, helix-5-deleted nanodisc, MSP1D1ΔH5 (henceforth ΔH5), reconstituted with DMPC lipids (ΔH5-DMPC)(16). The structure provided an important step towards a better understanding of the nanodisc and was obtained by combining nuclear magnetic resonance (NMR) spectroscopy, electron paramagnetic resonance (EPR) spectroscopy and transmission electron microscopy (TEM)(16). While these experiments were performed on lipid-loaded nanodiscs, the study focused on the protein components, and on determining a time- and ensemble averaged structure of these, but left open the question of the role of the lipids(7) as well as any structural dynamics of the overall nanodisc. Intriguingly,

Significance Statement

Membrane proteins play central roles in biology, but are difficult to study because they naturally function in an extended lipid membrane. Nanodiscs are membrane mimetics that consist of a protein belt surrounding a lipid bilayer, thus creating a well-defined and homogeneous environment for membrane proteins. Unfortunately, we lack a detailed structural and biophysical description of the protein and lipid components, hampering interpretation of experiments and design of improved nanodisc systems. We combined molecular simulations with small-angle X-ray and neutron scattering experiments and previously measured nuclear magnetic resonance experiments to study the structure and dynamics of the proteins and lipids in a nanodisc. Our results highlight spatial heterogeneity of lipid and protein components, providing new insights into the behaviour of nanodiscs.

[1]To whom correspondence should be addressed.
E-mail: arleth@nbi.ku.dk (L.A.) and lindorff@bio.ku.dk (K.L.-L.)

the resulting structure of the belt proteins corresponded to that of an almost circularly-shaped disc, while our previous SAXS/SANS investigations are clearly consistent with discs with an on-average elliptical cross-section(8, 10).

Here we continue towards the goal of understanding the structure and dynamics of the nanodisc and its membrane mimicking capabilities. We performed SAXS and SANS experiments on the ΔH5-DMPC variant, and integrated these with MD simulations and previously published NMR data(16) through an integrative Bayesian/maximum entropy (BME) approach (17–21). We thereby obtain a model of the conformational ensemble of the ΔH5-DMPC nanodisc that is consistent with the structural information obtained from each method, as well as our molecular simulations, and successfully explains differences in previous structural interpretations. In addition, we study the membrane mimicking capabilities by both simulation studies of lipid order parameters using the reweighted simulation ensemble and by Differential Scanning Calorimetry (DSC) measurements of the melting transition of DMPC in differently sized nanodiscs. The results provide information about the internal dynamics of the encircled lipid bilayer. This study shows how multiple methods and their integration helps simultaneously understand the structure and the dynamics of a complex protein-lipid system, and paves the way for future studies of complex biophysical systems such as membrane proteins embedded in nanodisc systems.

Results and Discussion

Structural investigations of ΔH5-DMPC and ΔH4H5-DMPC nanodiscs by SAXS and SANS. We determined optimal reconstitution ratios between the DMPC lipids and the ΔH5 and ΔH4H5 protein belts to form lipid-saturated nanodiscs based on a size-exclusion chromatography (SEC) analysis (Fig. S1 and Methods). In line with previous studies(3), we found that reconstitution ratios of 1:33 for ΔH4H5:DMPC and 1:50 for ΔH5:DMPC were optimal in order to form single and relatively narrow symmetric peaks. We then performed combined SEC-SAXS and SEC-SANS experiments to determine the size and shape of DMPC loaded ΔH5 and ΔH4H5 nanodiscs (Fig. 1). These experiments were performed at 10 °C, and based on results from previous NMR experiments on nanodiscs(7) as well as a melting temperature $T_M \approx 24$ °C for DMPC, we expect the lipids to be in the gel-phase. Our SAXS and SANS data all exhibit a flat Guinier region at low q and indicate no signs of aggregation (Fig. 1A, B). In both the ΔH5-DMPC and ΔH4H5-DMPC systems, the SAXS data exhibit an oscillation at medium to high q ([0.05:0.2] Å^{-1}) arising from the combination of a negative excess scattering length density of the hydrophobic alkyl-chain bilayer core and positive excess scattering length densities of the hydrophilic lipid PC-headgroups and the amphipathic protein belt. The SANS data decreases monotonically as a function of q in accordance with the homogeneous contrast situation present here. These two different contrast situations, core-shell-contrast for SAXS and bulk-contrast for SANS, are also clearly reflected in the obtained $p(r)$-functions (Fig. 1C, D), which also confirm that the ΔH5-DMPC nanodiscs are slightly larger than the ΔH4H5-DMPC nanodiscs. The data are in qualitative agreement with the SAXS and SANS data obtained for MSP1D1 nanodiscs (2, 8) and similar systems (9, 10), and indicate a discoidal structure.

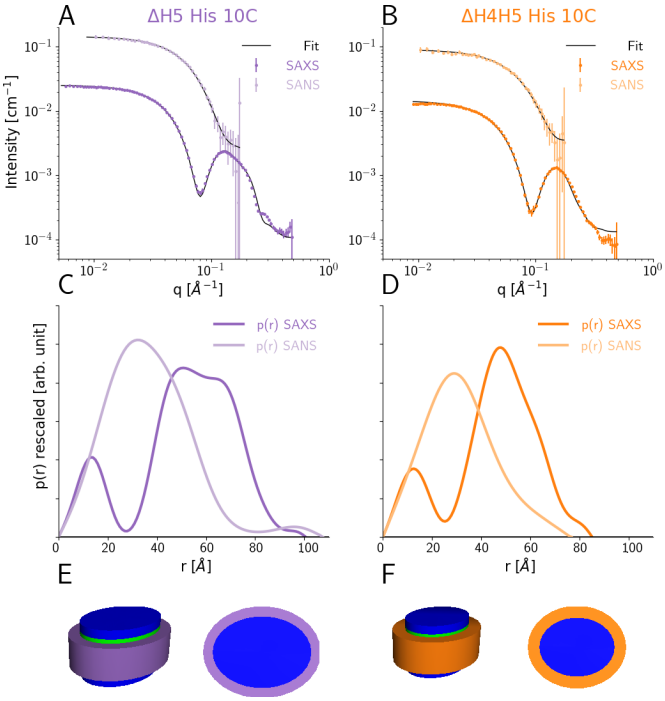


Fig. 1. SEC-SAXS and SEC-SANS analysis of nanodiscs. **A)** SEC-SAXS (dark purple) and SEC-SANS (light purple) for ΔH5-DMPC nanodiscs at 10 °C. The continuous curve show the model fit corresponding to the geometric nanodisc model shown in E. **B)** SEC-SAXS (dark orange) and SEC-SANS (light orange) data for the ΔH4H5-DMPC nanodiscs at 10 °C. **C,D)** Corresponding pair-distance distribution functions. **E, F)** Fitted geometrical models for the respective nanodiscs (drawn to scale relative to one another).

We analyzed the SEC-SAXS and -SANS data simultaneously by global fitting of a previously developed molecular-constrained geometrical model for the nanodiscs (8, 22) using the WillItFit Software developed in our laboratory (23). The nanodisc model (see further details in methods section) describes the interior of the nanodisc as a stack of flat, elliptically-shaped bilayer discs to account for the hydrophobic bilayer that is sandwiched in between the two hydrophilic headgroup layers. The inner lipid nanodisc is encircled by a hollow cylinder with an elliptical cross-section, which models the two protein MSP-belts stacked upon one another (Fig. 1E, F). Using this model, we obtained excellent simultaneous fits to SAXS and SANS data for both the ΔH4H5-DMPC and ΔH5-DMPC nanodiscs (Fig. 1A, B).

From the fitted parameters (Table 1 left) we found the area per headgroup for DMPC for both systems to be close to 55 Å^2. This value is somewhat higher than the A_{head} of gel-phase DMPC which is reported to be 47.2±0.5 Å^2 at 10 °C (24). Our finding of larger average A_{head}-values in the ΔH4H5 and ΔH5 nanodiscs is in agreement with the very broad melting transition observed in our DSC data (see below). The fitting revealed that the number of DMPC molecules in the nanodiscs were 65±13 and 100±14 for ΔH4H5 and ΔH5, respectively. These values are in good agreement with the optimal reconstitution ratios reported above. We note that both the SAXS and SANS data had to be re-scaled by a constant close to unity to fit the data (Table 1). In the case of the ΔH4H5-DMPC SANS data, the scaling constant (1.7±0.5) was larger than expected which is most likely the

Table 1. Parameters of the SAXS and SANS model fit. Left) Parameters for the simultaneous model fits to SEC-SAXS and SEC-SANS of His-tagged nanodiscs (denoted -His) for both ΔH4H5-DMPC and ΔH5-DMPC. Both measurements were obtained at 10 °C. Right) Standard solution SAXS measurements of the ΔH5-DMPC nanodisc without His-tags (denoted -ΔHis) obtained at two different temperatures, in the gel phase at 10 °C and in the liquid phase at 30 °C. * marks parameters kept constant.

	SEC-SAXS+SEC-SANS		SAXS	
	ΔH4H5-His	ΔH5-His	ΔH5-ΔHis	ΔH5-ΔHis
T	10 ^{o}C	10 ^{o}C	10 ^{o}C	30 ^{o}C
$\chi^2_{reduced}$	1.95	5.12	3.76	2.40
	Fitting Parameters			
Axis Ratio	1.3±0.4	1.2±0.2	1.4±0.1	1.3±0.1
A_{Head}	55± 5 Å^2	54± 2 Å^2	52± 2 Å^2	60± 3
H_{Belt}	24* Å^2	24* Å^2	24* Å^2	24* Å^2
N_{Lipid}	65±13	100±14	102± 7	104± 9
CV_{belt}	1*	1*	1*	0.97±0.02
CV_{lipid}	1.00±0.02	1.01±0.01	1.003±0.007	1.044±0.007
Scale$_{x-ray}$	1.13±0.28	1.1±0.2	1.2±0.1	1.2±0.2
Scale$_{neutron}$	1.7±0.5	0.8±0.2	-	-
	Results From Fits			
H_{lipid}	40 Å	41 Å	41 Å	38 Å
H_{tails}	28 Å	28 Å	29 Å	26 Å
R_{major}	27 Å	32 Å	34 Å	36 Å
R_{minor}	21 Å	27 Å	25 Å	28 Å
W_{belt}	10 Å	9 Å	9 Å	9 Å

result of a less accurate protein concentration determination for this system. In addition to the structural parameters reported, small constant corrections to the background were fitted to the SAXS and SANS data. Both the ΔH4H5 and ΔH5 proteins have protruding His$_6$-tags linked to them via a TEV cleavage site, these 23 amino acids were assumed to be disordered and included in the model as a Gaussian random coil of R_G=12.7 Å as estimated from previous studies of disordered proteins(25). Also an interface-roughness of nanodiscs of 2 Å were included in the model. Both the His-tag and roughness inclusion are in accordance with our previously developed methodology(8, 22, 23). We note that the fitted parameters of the geometric models suggest that both nanodiscs were found to be somewhat elliptical, with ratios of the two axes between 1.2 and 1.4. This observation is in apparent contrast to the recently described integrative NMR/EPR structural model of the ΔH5-DMPC nanodisc which was found to be more circular (16).

Temperature dependence probed by SAXS and SANS. We continued to investigate the impact of temperature and the presence of the His-tags on both the SAXS measurements and the resulting geometrical model of ΔH5-DMPC. To do so, we acquired standard solution SAXS data for a new preparation of the ΔH5-DMPC nanodiscs, this time without His-tags and measured at both 10 °C and 30 °C. At these two temperatures the DMPC is expected to be dominantly in the gel and liquid phase, respectively, as they are below and above the melting transition temperature(7) (see also DSC analysis below). We used a standard solution SAXS setup for these measurements, as this at present provides a better control of both the sample temperature and sample concentration than in the SEC-SAXS based measurement. The effect of the DMPC melting transition is clearly reflected in the SAXS data (Fig. S2) where both the position of the first minimum and the shape of the oscillation changes as the DMPC transitions from the gel to the molten state. In addition, we note that the intensity of the forward scattering decreases significantly with increasing temperature. This is a result of the small but significant temperature-dependent change of the partial specific molecular volume of the DMPC and hence a change in its excess scattering length density. To analyze the data, we again applied the molecular constrained geometrical model for the nanodiscs (Table 1, Right). Here, the effect of the DMPC melting transition can clearly be seen on the obtained DMPC area per headgroup which increases significantly as a result of the melting. Qualitatively similar observations of the melting transition of DMPC and DPPC based nanodiscs were previously reported in the MSP1D1 and MSP1E3D1 nanodiscs using DSC, SAXS and fluorescence(26, 27).

As a final control, we used the standard solution SAXS experiments to verify that the above use of His-tags in the SEC-SAXS and SEC-SANS measurements did not affect the overall structural properties. These data and the corresponding model fits show that, as expected, a similar nanodisc structure is obtained with and without His-tags (Fig. S2), and confirm the model derived from the corresponding His-tagged versions of the ΔH5-DMPC, with an overall elliptical shape (axis ratio 1.3 ±0.1).

Molecular Dynamics Simulations. The results described above suggest an apparent discrepancy of the solution structure of the ΔH5-DMPC nanodisc when viewed either by NMR/EPR or SAXS/SANS. In particular, the NMR/EPR structure revealed a circular shape whereas the SAXS/SANS experiments suggested an elliptical shape. The two kinds of experiments, however, differ substantially in the aspects of the structure that they are sensitive to. Further, both sets of models were derived in a way to represent the distribution of conformations in the experiments by a single 'average' structure.

In order to understand the structural discrepancies between the two solution methods better, and to include effects of multi-conformational averaging, we proceeded by investigating the underlying structural dynamics at a detailed atomistic level through a set of MD simulations of the His-tag truncated ΔH5-DMPC nanodisc. In these simulations, we mimicked the experimental conditions of the standard solution SAXS measurements obtained at 30°C and used 100 DMPC lipids in the bilayer as found above. We performed two simulations (total simulation time of 1196 ns) using the CHARMM36m force field(28). We visualized the conformational ensemble of the ΔH5-DMPC nanodisc by clustering the simulations, and found that the three most populated clusters represent 95% of the simulations. Notably, these structures all have elliptical shapes, but differ in the directions of the major axis (Fig. 2A).

We then examined the extent to which the simulations agree with the ensemble-averaged experimental data, focusing on the SAXS experiments and NOE-derived distance information from NMR. We calculated the SAXS intensities from the simulation frames using both FOXS(29, 30) (Fig. 2B) and CRYSOL(31) (Fig. S3) and compared to the corresponding standard solution SAXS experiments obtained at 30 °C. Similarly, we used r^{-3}-weighted averaging to calculate the effective distances in the simulations and compared them to the

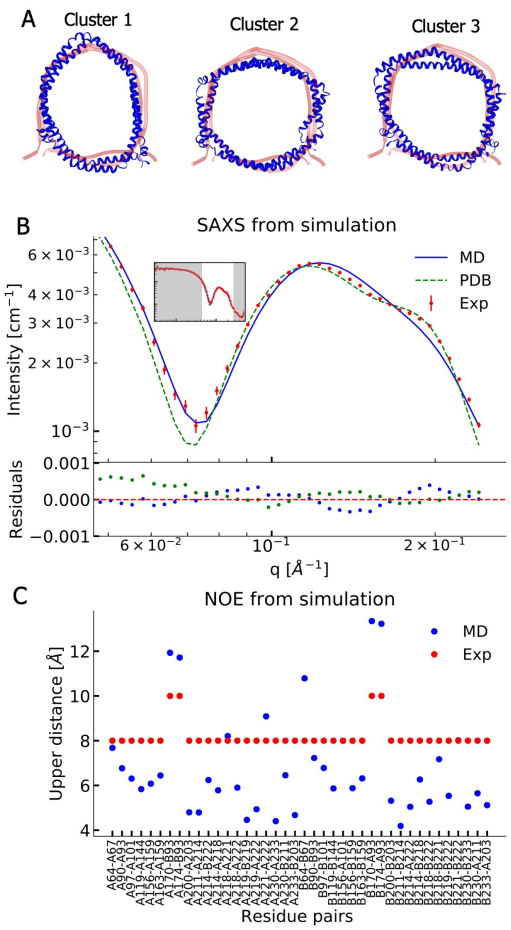


Fig. 2. Comparing MD simulations with experiments. **A)** Visualization of the conformational ensemble from the MD simulation by clustering (blue). Only the protein parts of the nanodisc are visualized while the lipids are left out to emphasize the shape. The top three clusters contain 95% of all frames. The previous NMR/EPR-structure is shown for comparison (red). **B)** Comparison of experimental standard solution SAXS data (red) and SAXS calculated from the simulation (blue). Green dotted line is the back-calculated SAXS from the integrative NMR/EPR-structure (labelled PDB). Residuals for the calculated SAXS curves are shown below. Only the high q-range is shown as the discrepancy between simulation and experiments are mainly located here, for the entire q-range see Fig. S3. **C)** Comparison of average distances from simulations (blue) to upper-bound distance measurements (red) between methyl NOEs. The labels show the residues which the atoms of the NOEs belong to.

previously reported methyl (Fig. 2C) and amide NOEs (Fig. S4)(16). The discrepancy observed between the simulation and the experiments were quantified by calculating χ^2 (Table 2).

The comparison between experiments and simulations reveal an overall good agreement between the two. Interestingly, the simulations agree well with the SAXS data in the q-region where scattering is dominated by the lipid bilayer and where our geometric fitting of the models for SAXS generally are very sensitive. The MD simulation trajectory captures accurately the depth of the SAXS minimum around $q = 0.07$; however, the shoulder observed in the experiments in the range 0.15 Å^{-1} – 0.20 Å^{-1} is not captured accurately.

Direct comparison of the previously determined integrative NMR/EPR structure (16) to the SAXS data is made difficult by the missing lipids in the structure. We thus built a model of the lipidated structure by first adding DMPC lipids to the NMR/EPR solved structure (PDB ID 2N5E), and then equilibrating only the lipids by MD, keeping the protein conformation fixed. When we use this structure to calculate the SAXS data, the back-calculated data overshoots the depth of the SAXS minimum but captures well the shoulder observed in the experimental data (Fig. 2B). Thus, neither the MD trajectory nor the NMR/EPR structure fit perfectly with the measured SAXS data.

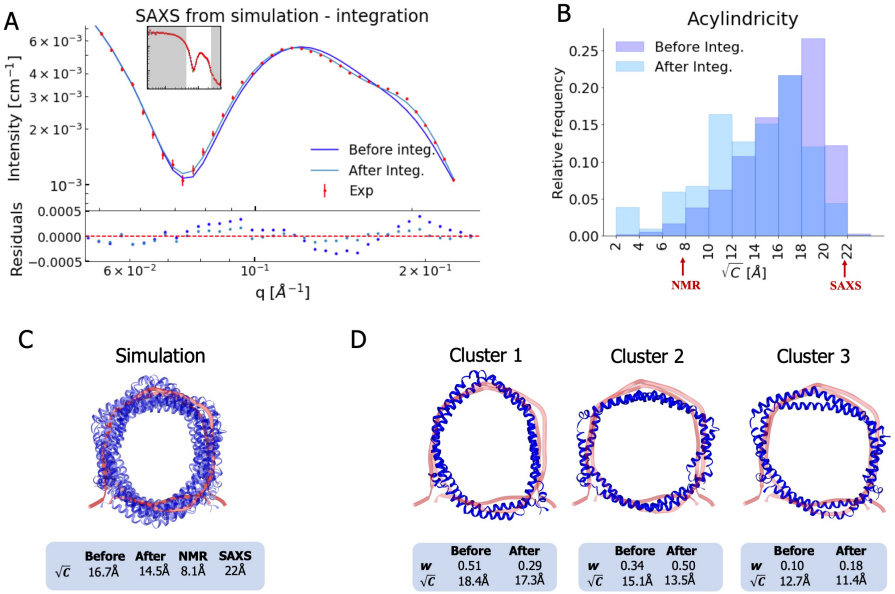
When comparing the simulations to the NMR-derived distances between methyl groups (Fig. 2C), we generally find good agreement, but observe a few distances that exceed the experimental upper bounds. A similar trend is observed in the comparison to amide NOEs (Fig. S4) which shows overall good agreement but with a few NOEs violating at similar positions as for the methyl NOEs.

As a final consistency check, we also compared the simulations to the SANS data for ΔH5-DMPC. The scattering contrast is very different in SAXS and SANS, and the scattering from the lipid bilayer has a relatively higher amplitude in the latter. This gives an independent check that the simulation provides a good description of the structure of the lipid bilayer. As the SEC-SANS data were measured on a His-tagged ΔH5-DMPC nanodisc, we therefore simulated this situation by creating an ensemble of His-tag structures and randomly sampled and attached these to the outer MSP-belts in the simulation frames under the assumption that the His-tags are disordered on the nanodiscs (Fig. S5). As for the SAXS and NOE data we also here find a generally good agreement.

Table 2. Comparing experiments and simulations We quantify agreement between SAXS and NMR NOE experiments by calculating the χ^2. The previously determined NMR structure(16) (PDB ID 2N5E) is labelled PDB, the unbiased MD simulation by MD, and simulations reweighted by experiments are labelled by MD and the experiments used in reweighting. S_{rel} is a measure of the amount of reweighting used to fit the data(19) (see Methods for more details).

Data for integration	S_{rel}	χ^2	
		SAXS	NOE
PDB	–	2.9	9.5
MD	0	10.0	8.2
MD + SAXS	-1.7	1.5	7.9
MD + NOE	-1.9	8.9	4.2
MD + SAXS + NOE	-1.7	1.9	6.0

Integrating experiments and simulations. While the MD simulations are generally in very good agreement with the SAXS and NMR data, there remain discrepancies that could contain information about the conformational ensemble of ΔH5-DMPC in solution. We therefore used a previously described Bayesian/Maximum Entropy (BME) approach(17–20, 32) to integrate the MD simulations with the SAXS and NMR data. Briefly, the BME method refines the simulation against measured experimental averages by adjusting the weights of each frame in the simulation. To do so, BME finds the optimal balance between 1) minimizing the discrepancy between the simulation and the observed experimental averages and 2) ensuring as little perturbation of the original simulation as possible thereby limiting chances of overfitting. The outcome



	Before	After	NMR	SAXS
$\sqrt{C}$	16.7Å	14.5Å	8.1Å	22Å

	Before	After
w	0.51	0.29
$\sqrt{C}$	18.4Å	17.3Å

	Before	After
w	0.34	0.50
$\sqrt{C}$	15.1Å	13.5Å

	Before	After
w	0.10	0.18
$\sqrt{C}$	12.7Å	11.4Å

Fig. 3. Integrating simulations and experiments. **A)** SAXS data calculated from the simulation before and after reweighting the ensemble using experimental data. Only the high q-range is shown as the discrepancy between simulation and experiments are mainly located here, for the entire q-range see Fig. S6. Agreement with the NOEs before and after integration are likewise shown in Fig. S7. **B)**. Histogram of the acylindricity of the simulations ($\sqrt{C}$) both before integration (dark blue) and after integration (light blue). **C)** Visualization of the conformational ensemble showing structures sampled every 100 ns in cartoon representation (blue), the original NMR/EPR structure is shown in rope representation for comparison (red). The table below shows the acylindricity of the entire conformational ensemble before and after integration and compared to the original NMR/EPR (NMR) structure and the SAXS/SANS model fit. **D)** Weights and acylindricity of the three main clusters of the MD simulation (blue) before and after integration.

is a conformational ensemble that is more likely to represent ΔH5-DMPC in solution. In practice, this is achieved by changing the weight of each configuration in the ensemble obtained from the MD simulations, and we therefore call this a 'reweighted ensemble' (20, 33). The amount of reweighting can be quantified by an entropy change (S_{rel}) that reports on how much the weights had to be changed to fit the data(19, 20) (see Methods). Alternatively, the value $\phi_{eff} = \exp(S_{rel})$ reports on the effective ensemble size, that is what fraction of the original frames that were used to derive the final ensemble(21).

We used both the SAXS and NOE data individually, as well as combined, to understand the effects of each source of data on the reweighted conformational ensemble (Table 2). We note that when a specific type of data is used to generate the ensemble, the resulting χ^2 simply reports on how well the simulation has been fitted to the data; because of the maximum-entropy regularization to avoid overfitting, we do not fit the data as accurately as possible. The two types of experimental data complement each other in structural information content. Specifically, the SAXS data report on the overall size and shape, and is sensitive to both the protein and the lipids through atom-atom pair distributions in a range starting from ≈ 10 Å, whereas the NOEs contain local, specific atom-atom distances from the protein belts of the ΔH5 but not any direct information about the lipids.

We find that refining the simulations against the SAXS data primarily improves the agreement with that data, but does not substantially improve agreement with the NOEs. A similar observation was made when refining against the NMR data. Thus, the two data types only improve the MD trajectory with respect to the structural properties they are sensitive to, but do not compensate for other structural inaccuracies in the simulation, again highlighting the orthogonal information in the two sources of information. In addition, we performed reweighting with the methyl NOEs and the amide NOEs separately (Table S1). The already low discrepancy of the amide NOEs barely improves while the discrepancies of both methyl NOEs and SAXS are unaffected by integration with amide NOEs alone, implying that the structural information content contained in the amide NOEs is already correctly captured by the force field. Thus, as expected, reweighting against a source of data where there is already good agreement has little effect on the conformational ensemble.

Because the NOE and SAXS experiments provide independent information we refined the ensemble against both sets of data (Fig. 3). We find that we can fit both sources of data at reasonable accuracy without dramatic changes of the weights away from the Boltzmann distribution of the force field ($\phi_{\text{eff}} = 18\%$). We thus proceed with our analysis of the structural features of ΔH5-DMPC using an ensemble of conformations that is based on integrating the MD simulations with both the SAXS and NOE experiments. For comparison we also present the structural features based only on the MD simulation trajectory i.e. before reweighting the simulation against experiments.

Analysis of the measured SAXS and SANS revealed an elliptical shape of the ΔH5-DMPC upon fitting of a single structure to the data, i.e. an time- and ensemble averaged structure. In contrast, the previously integrative NMR/EPR method gave rise to a more circular configuration upon fitting a single structure. Combining the results of the two studies, we hypothesized that the nanodisc possesses underlying elliptical fluctuations with the major axis changing within the nanodisc. In such a system NMR and EPR measurements, which build on ensemble averaged information of specific atom-atom interactions, will give rise to an on-average circular structure. SAXS and SANS, on the contrary, which build on distributions of global distances rather than specific atom-atom distances,

will not distinguish between the orientation of the major axis within the nanodisc and thus give rise to observations of an elliptical shape. By complementing the experiments with MD simulations we obtained a trajectory with structural features that supports this hypothesis.

To investigate the hypothesis further, we quantified the degree of ellipticity in terms of an acylindricity parameter, C, defined as the difference between the x and y components of the gyration tensor (see Methods for details). C is thus a measure of how far from a perfect circular cylinder the shape is, and $C = 0$ corresponds to a circular shape. We calculated both the average and distribution of the acylindricity from the simulated ensemble both before and after reweighting against the experimental data (Fig. 3B and 3C). In addition, we calculated the acylindricity of both the integrative NMR structure and from the structural model obtained from the SAXS and SANS measurements.

We find that the acylindricity decreased from $\sqrt{C} = 17$ Å in the original MD simulation trajectory to $\sqrt{C} = 15$ Å after integration of the NMR and SAXS data, showing that the experiments indeed affect the structural features. This value is in the middle of that obtained from the analytical geometric model fitted to the SAXS data ($\sqrt{C} = 22$ Å) and that of the integrative integrative NMR/EPR structure(16). Thus, the acylindricity of the final, heterogenous ensemble lies between that of the two conformations that were fitted as single structures to fit either the NMR or SAXS data.

To understand better the elliptical shape of the ΔH5-DMPC nanodisc and the role played by reweighting against experiments, we calculated the average acylindricity for each cluster of conformations of ΔH5-DMPC both before and after integration with experimental data (Fig. 3D). We note that because our reweighting procedure acts on the individual conformations and not at the coarser level of clusters, the average acylindricity changes slightly for each cluster upon reweighting. Clusters 1 and 2, which together constitute about 80% of the conformational ensemble (both before and after reweighting), are both clusters with high acylindricity, but with almost orthogonal directions of the major axis in the elliptical structure. The major change after integration is the exchange in populations of the two clusters resulting in cluster 2 to be weighted highest, underlining the influence and importance of the integration. Thus, our MD simulations and the integration with the experiments support the hypothesis of underlying elliptical fluctuations with the major axis changing direction inside the nanodisc, and we note that the detailed molecular description of this was only possible by combining the MD simulations with both the SAXS and NMR data.

Analyses of lipid properties in small nanodiscs. Nanodiscs are often used as models for extended lipid bilayers, but the presence and interactions with the protein belt — and the observed shape fluctuations — are likely to impact the properties of the lipid molecules in the nanodisc compared to a standard bilayer. We therefore proceeded to use our experimentally-derived ensemble of nanodisc structures to investigate the properties of lipids in the small ΔH5-DMPC nanodisc, and compared them to those in a DMPC bilayer.

Following previous solid-state NMR(6, 7) and simulation(13, 14) work, which showed perturbed lipid dynamics inside nanodiscs, we analyse the thickness of the DMPC bilayer (Fig. 4A) and order parameters, S_{CH}, of the DMPC lipids (Fig. 4B,C).

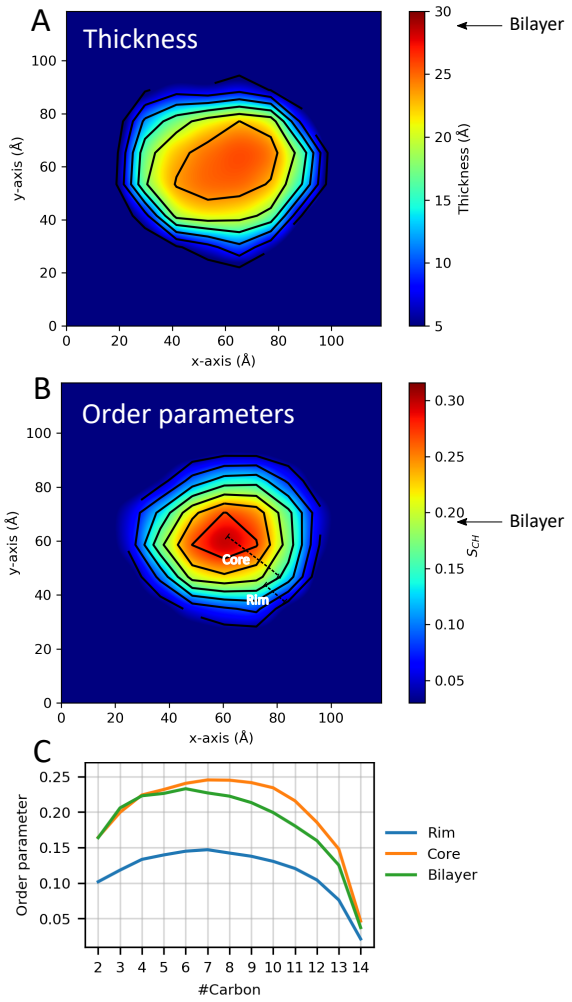


Fig. 4. Lipid properties from simulations. 2D plots of the (**A**) thickness and (**B**) order parameters of the DMPC lipids in the ΔH5 nanodisc, respectively. The core and rim zones are indicated in panel B. Arrows indicate the average value in simulations of a DMPC bilayer. **C)** The order parameters as a function of carbon number in the lipid tails in the ΔH5-DMPC disc. The rim zone is defined as all lipids within 10 Å of the MSPs, while the core zone is all the lipids not within 10 Å of the MSPs.

As done previously(13, 14), we subdivide the lipid area in the nanodisc into zones dependent on the distance from the MSP protein belts (above or below 10 Å). The results of both the thickness and S_{CH} order parameter analyses show the same trend: a clear difference between the lipids close to the protein belt and those more central in the nanodisc. The results illustrates that the DMPC lipid bilayer in small ΔH5 nanodiscs are far from homogeneous but rather thinner and un-ordered near the rim of the protein belt and thicker and more ordered in the core of the nanodisc. As a reference we compare the thickness and lipid dynamics to the results from simulations of a DMPC bilayer, and the results show, as expected, that the more central lipids in the nanodisc behave more similarly to a pure bilayer (Fig. 4). These results are in line with previous simulation studies on the larger DMPC nanodiscs, MSP1, MSP1E1 and MSP1E2(13, 14), albeit performed without experimental reweighting, as well as

with solid state NMR data on the both ΔH5-DMPC and the larger MSP1-DMPC(6, 7). We note that the solid state NMR studies suggest smaller differences in the order parameters between the ΔH5-DMPC nanodisc and a DMPC bilayer (7) compared to the difference between a DMPC bilayer and the larger MSP1-DMPC(6), suggesting that the smaller nanodisc is more similar to a bilayer membrane. However, the study also shows an almost abolished lipid gel-liquid phase transition in the ΔH5-DMPC nanodisc where a DMPC bilayer otherwise show a steep transition.

DSC analysis of lipid melting in nanodiscs. To gain further insight into lipid bilayer properties in nanodiscs — specifically the impact of the differentiated lipid order in the core and rim of the nanodisc — we proceed to study the melting transition of DMPC lipids in nanodiscs of different sizes. Specifically, we performed Differential Scanning Calorimetry (DSC) to study the lipid melting transition of DMPC inside ΔH4H5, ΔH5 and the larger MSP1D1 nanodiscs, and used pure DMPC vesicles as reference.

Our results show that the melting transition peak broadens significantly in all three nanodisc systems compared to that of pure DMPC vesicles (Fig. 5). For all three discs, we note that the melting starts already at around 10 °C below the T_M and that the transition does not start leveling off to a plateau until at around 10 °C above the T_M. Furthermore, the pre-transition visible at 15 °C for the DMPC vesicles, is not visible in any of the nanodisc systems. The broader melting transition is in line with the observed differentiated lipid ordering in nanodiscs from the reweighted simulations, as such differences in how ordered the lipids are necessarily will cause differences in the melting temperature and thus give rise to the broader peaks. Furthermore, the broadened peaks are in line with results observed in previous solid state NMR experiments which found an substantially broadened and diminished lipid gel-liquid phase transition in the ΔH5-DMPC nanodisc in the temperature range 10-28 °C (7).

In addition, our results show that the transition enthalpy per mole of DMPC, i.e. the area under the curves, increases with the nanodisc size. Both the broadening and the size-dependent increase of the transition enthalpy are consistent with previous observations for, respectively, DMPC and DPPC in MSP1D1 and in the larger MSP1E3D1 systems(26). Since the lipid area-to-rim ratio increases with increasing nanodisc size, Denisov *et al.* proposed that these observations were simply due to the absence of a cooperative melting transition of the lipids at the nanodisc rim(26); a suggestion fully supported by our observation of substantially perturbed lipids near the belt proteins.

Interestingly, we observe that the maximum of the melting transition, T_M, depends on the choice of nanodisc belt and can fall both below and above the T_M of plain DMPC vesicles (24 °C). In the smallest ΔH4H5 nanodisc, the DMPC has a $T_M \approx 22.5$ °C. In ΔH5 the DMPC has T_M at 24.5 °C which is close to that of the DMPC vesicles, while the larger MSP1D1 nanodisc has a $T_M \approx 28$ °C. This T_M value is similar to the value of 28.5 °C measured for DMPC melting in MSP1D1 by Denisov *et al.*(26), who in addition measured a T_M value of 27.5 °C in the even larger MSP1E3D1 discs.

Together our results are in line with previous NMR experiments (7), and suggest that the state of the ordering of the lipids in the nanodiscs is inhomogeneous compared to the DMPC vesicle, and that the behaviour of the lipids is modulated by their interaction with the membrane scaffold proteins. Our results point towards a non-trivial effect of the DMPC-MSP interactions. They can both destabilize DMPC in the gel-phase in the smaller nanodiscs (ΔH4H5-DMPC) where the low area-to-rim ratio leads to the lower T_M compared to the DMPC vesicles, but also stabilize the DMPC gel-phase in larger nanodiscs with larger area-to-rim ratios such as MSP1D1-DMPC and MSP1E3D1-DMPC. Thus, when using nanodiscs as membrane mimics it is relevant to keep in mind that the given lipid gel/liquid state might be affected. We also note that even if lipids in larger discs are less perturbed than those in the smallest discs, introduction of membrane proteins into the discs might in itself perturb the lipids in ways similar to the MSPs.

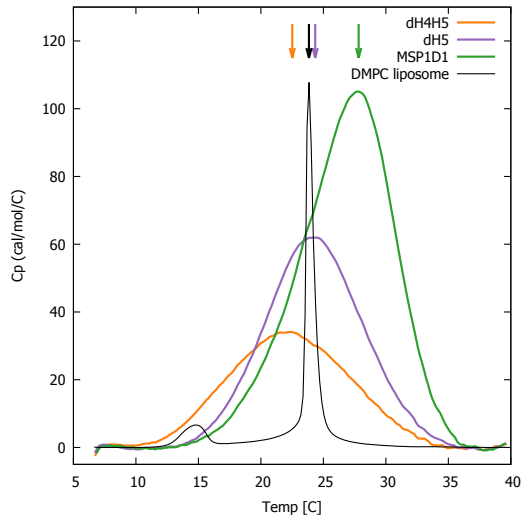


Fig. 5. DSC analysis of lipid melting in nanodiscs. The three DMPC-filled nanodiscs studied, listed by increasing size, are ΔH4H5-DMPC (orange), ΔH5-DMPC (purple) and MSP1D1-DMPC (green). DSC data from plain DMPC-vesicles (black) are shown for comparison. Arrows indicates the temperature with maximal heat capacity. DSC data from the three nanodisc samples are normalized by DMPC concentration, while the data from the DMPC liposome is on an arbitrary scale.

Conclusions

Lipid nanodiscs are versatile membrane mimetics with a wide potential for studies of the structure, function and dynamics of membrane proteins. Despite their widespread use and numerous studies, we still do not have a full and detailed understanding of the structural and dynamic features of nanodiscs. This in turn limits our ability to interpret e.g. solution scattering experiments when membrane proteins are embedded into such nanodiscs. In order to further our understanding of the conformations and structural fluctuations of both the protein and lipid components in nanodiscs, we have performed a series of biophysical experiments on DMPC-loaded ΔH5 and ΔH4H5 nanodiscs.

Using SEC-SAXS and SEC-SANS measurements, we investigated the solution structure of the ΔH4H5-DMPC and ΔH5-DMPC. Model-based analysis of this data showed an 'average' elliptical shape of both nanodiscs. In contrast, a previously determined integrative NMR/EPR (16) method gave rise to a more circular average structure of the ΔH5 nanodisc.

We reconcile these two apparently opposing views and provide a richer and more detailed view of the nanodisc and its dynamics by performing MD simulations. In particular, we used a Bayesian/Maximum Entropy approach to integrate the MD simulations with the SAXS and NMR data to uncover the existence of underlying fluctuations between elliptical shapes with orthogonal major axes in consistency with both sources of data. We note that the NMR/EPR-derived structure, and our MD simulations initiated from this structure, provide very good agreement with the SAXS data even without reweighting. Because our SAXS data are rather precise, however, we were able to detect subtle deviations that enabled us to refine our model.

In addition to studying the overall shape fluctuations, we also analysed the lipid structure and dynamics in the nanodiscs, and find an inhomogeneous distribution. Specifically, we find substantially perturbed lipid properties near the belt proteins, whereas the lipids more central in the disc behaved more similar to those in a pure DMPC bilayer. We used DSC to investigate the lipid melting transition in the small nanodiscs in comparison to the lipid vesicles and found that the melting takes place over a much broader temperature range in the small nanodiscs. The observed correlation between the size of the belt proteins and the lipid melting enthalpy give support to the proposition(26), that the arrangement of the lipids near the nanodisc rim must be substantially perturbed.

Together, our results provide an integrated view of both the protein and lipid components of nanodiscs that provide new details in their structure and dynamics. Approaches such as the one described here takes advantage of the increasing possibilities for accurate NMR and scattering data in solution, strongly improved models for lipid bilayers for certain lipid types as well as new approaches to integrate experiments and simulations. In this way, our study exemplifies how integrating multiple biophysical experiments and simulations may lead to new insight into a complex system and paves the way for future studies of membrane proteins.

Materials and Methods

Expression of Membrane Scaffold Protein (MSP) variants. The constructs for ΔH4H5, ΔH5 and MSP1D1are those reported previously (3, 34). The expression was performed according to established procedures(34) while the purification protocol was slightly modified from that previously described (34). In brief, the cells were opened in lysis buffer containing 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 1% Triton X-100 and 6 M GuHCl by vigorous shaking for 15 min. Insoluble material was subsequently removed by centrifugation at 18,000 rpm for 1 hour using an SS-34 rotor. The supernatant was loaded on Ni-NTA resin pre-equilibrated in lysis buffer and washed extensively with the same buffer. Extensive washes using lysis buffer without GuHCl and subsequently wash buffer containing 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 20 mM Imidazole and 50 mM Cholate was performed in order to remove GuHCl and Triton X-100. Protein was eluded in buffer containing 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 500 mM Imidazole, concentrated, flash frozen and stored at -80 °C until further use. Cleavage of the TEV-site was performed by addition of 1:20 TEV protease, and dialysing at room temperature for 6–12 hours against 20 mM TrisHCl pH 8, 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT. TEV protease and any uncleaved MSP was removed by passing the solution over Ni-NTA resin again.

Reconstitution of ΔH5-DMPC and ΔH4H5-DMPC Nanodiscs. Before assembly, the DMPC lipids were suspended in a buffer containing 100 mM NaCl, 20 mM Tris/HCl pH 7.5, and 100 mM sodium cholate detergent to a final lipid concentration of 50 mM. All lipids used were obtained from Avanti Polar Lipids(35). We determined optimal reconstitution ratios between the DMPC lipids and the ΔH5 and ΔH4H5 by first mixing the lipid and MSP stock solutions at a series of different molar concentration ratios in the range from 1:9 to 1:80 depending on the MSP type (Fig. S1). In all samples, cholate was removed after mixing by addition of an excess amount of Amberlite detergent absorbing beads to start the assembly of the nanodiscs. The samples were left in a thermomixer for 4 hours at 28 °C and the Amberlite was removed by spinning the solution down at 5000 rpm. Afterwards, purification was performed using size exclusion chromatography (SEC) on an Äkta purifier (FPLC) system with a Superdex200 10/300 column from GE Healthcare Life Science (S200)(36). The obtained SEC traces allowed for determining the optimal reconstitution ratios between the DMPC lipids and the ΔH5 and ΔH4H5 and we found that reconstitution ratios of 1:33 for ΔH4H5:DMPC and 1:50 for ΔH5:DMPC were both optimal in order to form single and relatively narrow symmetric peaks. This is in close agreement with the previously reported ratios of 1:20 for ΔH4H5:DMPC and 1:50 for ΔH5:DMPC (3). More narrow and well-defined SEC-peaks were obtained if the reconstitution took place at or above the melting temperature, T_M, of DMPC at 24 °C (data not shown) as recommended in the literature(34).

Differential Scanning Calorimetry (DSC). The measurements were performed on a VP-DSC (MicroCal) using a constant pressure of 1.7 bar (25 psi) and a scan rate of 1 °C/min between 6 °C and 40 °C. All samples had been purified in PBS buffer prior to the measurement. Data was background subtracted and baseline corrected in the Origin instrument software using a "Cubic Connect" baseline correction. Finally, the data was normalized by the lipid concentration of the individual samples.

SEC-SANS. SEC-SANS was performed at the D22 small-angle scattering diffractometer at the ILL, Grenoble, France using a very recently developed SEC-SANS setup (37, 38). Briefly, the setup was as follows: the *in situ* SEC was done using a modular HPLC system (Serlabo) equipped with a Superdex200 10/300 GL gel filtration column (GE) with a void volume of approximately 7.5 ml and a flow rate of 0.25 ml/min. The SmartLine 2600 diode-array spectrophotometer (Knauer) was connected via optic fibers either to an optic cell of 3 mm path length placed at the outlet of the chromatography column, enabling the simultaneous recording of chromatograms at four different wavelengths, including 280 nm which we used for the concentration determination.

All components of the HPLC setup including buffers and the column were placed in a closed cabinet connected to an air-cooling system set to 10 °C to control the temperature of the sample. Before measurements, we equilibrated the column in a D_2O-based buffer, and the buffer in the sample was exchanged to a D_2O-based buffer using a illustra NAP-25 gravity flow column (GE)(36). The D_2O buffer contained 20 mM Tris/DCl pH 7.5 and 100 mM NaCl.

The experiments were carried out with a nominal neutron wavelength, λ, of 6.0 Å and a wavelength distribution, $\Delta\lambda/\lambda = 10\%$ FWHM, a rectangular collimation of 40 mm × 55 mm and a rectangular sample aperture of 7 mm × 10 mm. The distance of the sample-detector used for the characterization of the empty nanodiscs was 5.6 m (with collimation of 5.6 m), covering a momentum transfer range, q, of 0.0087 $Å^{-1}$ to 0.17 $Å^{-1}$, with $q = 4\pi \sin(\theta)/\lambda$, where θ is half the angle between the incoming and the scattered neutrons. Measured intensities were binned into 30 s frames. Sample transmission was approximated by the buffer measured at the sample-detector distance of 11.2 m. The measured intensity was brought to absolute scale in units of scattering cross section per unit volume (cm^{-1}) using direct beam flux measured for each collimation prior to the experiment. Data reduction was performed using the GRASP software (39). The SANS data appropriate for buffer subtraction was identified based on when the 280 nm absorption during the SEC curve showed no trace of protein.

SEC-SAXS. SEC-SAXS was performed at the BioSAXS instrument at BM29 at the ESRF, Grenoble, France(40). Briefly, the setup at BM29 included an HPLC controlled separately from the SAXS measurement, coupled to a UV-Vis array spectrophotometer collecting absorption from 190 nm to 800 nm. Data were collected

with a wavelength of 0.9919 Å using a sample-detector distance of 2.87 m which provided scattering momentum transfers ranging from 0.003 $Å^{-1}$ to 0.49 $Å^{-1}$. The capillary was cooled to 10 °C, however, the HPLC including the SEC-column was placed at ambient temperature. Size exclusion chromatography was performed using the exact same column as for SEC-SANS and equivalent H_2O-based buffer. A flow rate of 0.5 ml/min was used. Data reduction was carried out using the in-house software, and subsequent conversion to absolute units was done with water as calibration standard (41). The 1 s frames recorded were subsequently averaged in 10 s bins.

Standard solution SAXS. Standard solution SAXS data were obtained at the P12 beamline at the PETRA III storage ring in Hamburg, Germany (42) using a wavelength of 1.24 Å, a sample-detector distance of 3 m, providing a momentum transfers covering from 0.0026 $Å^{-1}$ to 0.498 $Å^{-1}$ and a variable temperature in the exposure unit. 20 exposures of 0.045 seconds were averaged, background subtracted and normalized by the available software at the beamline. The measurements were performed at both 10 °C and 30 °C.

SAXS and SANS data analysis.. The output of the SAXS and SANS experiments were small-angle scattering data in terms of absolute intensities $I(q)$. As usual, the scattering vector $q = 4\pi \sin\theta/\lambda$ where θ is the scattering angle and λ is the wavelength of the incoming beam. $I(q)$ was transformed into the pair distance distribution function, $p(r)$, by indirect Fourier transformations (IFT) which was performed at the BayesApp-server (43).

SAXS/SANS modeling was carried out using our previously developed WillItFit software (23) which is available for download https://sourceforge.net/projects/willitfit/. The applied structural models (see futher description below) are an adaptation of similar models previously developed to analyse SAXS and SANS data from MSP1D1 nanodiscs (8, 22).

In brief, the model describes the nanodiscs of a certain ellipticity based on analytical form factors (44) as explained by Skar-Gislinge *et al.* (8, 22). The ellipticity, in terms of the axis ratio is allowed to vary in the fit and can also take the size of unity corresponding to a circular disc. The model is fitted on absolute scale and utilizes information from the molecular composition of the applied protein belt and DMPC hydrophilic head and hydrophobic tail groups, provided via an input card to calculate the excess scattering length densities and total partial volumes of these compounds. Apart from the parameters listed in Table 1, the model also fits a small constant background added to the model, and includes a term accounting for interface roughness, fixed to 2 Å in the present analysis, and where relevant, a Gaussian random coil description of the linked TEV-His-tag with $R_G = 12.7$ Å consistent with the assumption that the 23 amino acids of the tag are in a fully disordered state(25).

MD simulations. The MD simulations were started from the first model in PDB ID 2N5E (16). A total of 50 pre-equilibrated DMPC lipids (45, 46) were inserted into each bilayer inside the protein belt. The number of lipids was chosen from the measured optimal reconstitution ratio, and in accordance with the reconstitution ratio used in the experiments for the NMR structure(16) as well as obtained from our fit of the geometric model to the SAXS and SANS data. The MD simulations were performed using the GROMACS 5.0.7 software package (47, 48) and the CHARMM36m force field(28). The system was solvated in a cubic box and neutralized by addition of Na^+ counter ions followed by a minimization of the solvent. Equilibration was performed in 6 steps following the protocol from CHARMM-GUI(49) with slow decrease in the positional restraint forces on both lipids and protein. The volume of the box was then equilibrated in the NPT ensemble at 303.15 K and 1 bar giving a final box with side lengths 13.2 nm. The production run was performed in the NVT ensemble at 303.15 K (above the phase transition of the DMPC lipids) using the stochastic velocity rescaling thermostat (50), 2 fs time steps and the LINCS algorithm to constrain all bonds. We performed two production runs (lengths 600 ns and 595 ns) starting from the same equilibrated structure. We concatenated these two MD simulations into a single trajectory, which then represents our sample of the dynamics of the system. We clustered the conformations from the simulations (one structure extracted for every nanosecond) with the RMSD based Quality Threshold method(51, 52) using C_α atoms only and with a cluster diameter cutoff of 0.58 nm; this resulted in six clusters. We also performed a 50ns-long simulation of a pure DMPC bilayer. The simulation parameters were the same as for the nanodisc system apart from the using the NPT ensemble and anisotropic pressure control.

Calculating SAXS and SANS from simulations. We performed SAXS calculations using both CRYSOL(31) and FOXS(29, 30) on structures extracted every 1 ns from the simulations using fixed parameters and q-range from 0.0 $Å^{-1}$ to 0.25 $Å^{-1}$. Most of the the overall structural information is contained within this q-range, and the calculations of SAXS intensities from the structures are also less accurate in the wide-angle regime. The SAXS profile of the NMR/EPR structure was calculated by adding DMPC lipids to the first model of the PDB entry and subsequent equilibration by MD (fixing the protein), and then using FOXS to back-calculate the SAXS. SANS calculations were performed using CRYSON(53) with the fraction of D_2O in solution set to 1.0 in accordance with the experimental measurements.

We used standard solution SAXS data experimental recorded at 30 °C on the ΔH5-DMPC (without His-TEV-tags) to compare to our simulations, as this setup is most similar to that used to derive the NMR/EPR structure. For more details on the SAXS/SANS calculations, see Supporting Information Text.

NOE calculations from simulations. We calculated distances corresponding to the experimentally observed NOEs on structures extracted every 1 ns from the simulations. To compare with the experimental distances, available as upper bounds, we averaged the distances, R, between the respective atoms (or the geometric center for pseudo atoms) as $\left\langle R^{-3} \right\rangle^{-1/3}$(54). When calculating χ^2 for validation we only include those distances where this average exceeded the experimentally-determined upper-bounds.

Integrating experiments and simulations. We used a Bayesian/maximum entropy approach (17–19), as implemented in the BME software (20) (github.com/KULL-Centre/BME), to integrate the molecular simulations with the SAXS and NMR experiments. The name originates from the two equivalent approaches, Bayesian and Maximum Entropy ensemble refinement, which are equivalent when the errors are modelled as Gaussians (17, 20, 32). We here provide a brief overview of the approach and refer the reader to recent papers for more details (17, 20, 21, 32).

Given that our MD simulations provide a good, but non-perfect, agreement with experiments the goal is to find an improved description of the nanodisc that simultaneously satisfies two criteria: (i) the new ensemble should match the data better than the original MD ensemble and (ii) the new ensemble should be a minimal perturbation of that obtained in our simulations with the CHARMM36m force field in accordance with the maximum entropy principle. In a Bayesian formulation, the MD simulation is treated as a *prior* distribution and we seek a *posterior* that improves agreement with experiments. This may be achieved by changing the weight, w_j, of each conformation in the MD-derived ensemble by minimizing the negative log-likelihood function(17, 20):

$$\mathcal{L}(w_1 \dots w_n) = \frac{m}{2}\chi_r^2(w_1 \dots w_n) - \theta S_{\mathrm{rel}}(w_1 \dots w_n). \qquad [1]$$

Here, the reduced χ_r^2 quantifies the agreement between the experimental data (F_i^{EXP}) and the corresponding ensemble values, ($F(\mathbf{x})$), calculated from the weighted conformers ($\mathbf{x}$):

$$\chi_r^2(w_1 \dots w_n) = \frac{1}{m}\sum_i^m \frac{(\sum_j^n w_j F_i(\mathbf{x}_j) - F_i^{EXP})^2}{\sigma_i^2}. \qquad [2]$$

The second term contains the relative entropy, S_{rel}, which measures the deviation between the original ensemble (with initial weights w_j^0 that are uniform in the case of a standard MD simulation) and the reweighted ensemble $S_{\mathrm{rel}} = -\sum_j^n w_j \log\left(\frac{w_j}{w_j^0}\right)$. The temperature-like parameter θ tunes the balance between fitting the data accurately (low χ_r^2) and not deviating too much from the prior (low S_{rel}). It is a hyperparameter that needs to be determined (Fig.

S8). In practice it turns out that minimizing $\mathcal{L}$ can be done efficiently by finding Lagrange multipliers in an equivalent Maximum Entropy formalism and we refer the reader to previous papers for a full description and discussion of the approaches including how to determine θ (17, 20, 32).

Acylindricity. In order to quantify how 'elliptical' the different nanodisc conformations are, we calculated the square root of the acylindricity, $\sqrt{C}$, where the acylindricity is defined from the principal components of the gyration tensor as $C := \lambda_x^2 - \lambda_y^2$, where the z-axis is orthogonal to the membrane and has the smallest principal component. In our calculations we included only the protein backbone atoms (excluding also the flexible tails from residues 55-63). This choice also makes it possible to compare with a similar calculation from the geometric model fitted from the SAXS and SANS data where the acylindricity was calculated using the major and minor axes from the geometric fit.

Lipid properties. We calculated the bilayer thickness and lipid order parameters for both the nanodisc and a simulated DMPC lipid bilayer. The values obtained for the nanodisc were from the reweighted ensemble every 1 ns.

We defined the bilayer thickness as the minimum distance along the bilayer normal between two phosphate headgroup pairs in the two leaflets. The headgroup pairs were identified and saved for each leaflet, top and bottom, along with the corresponding thickness and xy-coordinates. The pairs were further distributed unto a 6×6 grid in the xy-plane with each bin corresponding to 22 Å for both the top and bottom leaflet. An averaged grid was then obtained from the two grids of the leaflets. The order parameters S_{CH} are defined as(55):

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

where θ is the angle between the C-H bond and the bilayer normal. The order parameters were calculated for each lipid and each carbon along the two lipid tails every 1 ns. The values were further averaged across the two lipid tails before distributed unto a 6×6 grid. An average across frames and lipids were then obtained for each bin. In order to study the profile of the lipid tails, an average across frames, lipids, and tails were likewise obtained. Parameters were calculated from the simulations of the DMPC bilayer in the same way.

ACKNOWLEDGMENTS. The research described here was supported by a grant from the Lundbeck Foundation to the BRAINSTRUC structural biology initiative (L.A. and K.L.-L.), a Hallas-Møller Stipend from the Novo Nordisk Foundation (K.L-L.), the Velux Foundations (K.L.-L.) and the Danish Council for Independent Research in Technology and Production (B.S.).

1. TH Bayburt, YV Grinkova, SG Sligar, Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2**, 853–856 (2002).
2. IG Denisov, YV Grinkova, AA Lazarides, SG Sligar, Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J. Am. Chem. Soc.* **126**, 3477–3487 (2004).
3. F Hagn, M Etzkorn, T Raschle, G Wagner, Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *J. Am. Chem. Soc.* **135**, 1919–1925 (2013).
4. MT Marty, H Zhang, W Cui, ML Gross, SG Sligar, Interpretation and deconvolution of nanodisc native mass spectra. *J. The Am. Soc. for Mass Spectrom.* **25**, 269–277 (2014).
5. MT Marty, et al., Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
6. K Mörs, et al., Modified lipid and protein dynamics in nanodiscs. *Biochimica et Biophys. Acta (BBA) - Biomembr.* **1828**, 1222 – 1229 (2013).
7. D Martinez, et al., Lipid internal dynamics probed in nanodiscs. *ChemPhysChem* **18**, 2651–2657 (2017).
8. N Skar-Gislinge, et al., Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **132**, 13713–13722 (2010).
9. SR Midtgaard, et al., Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **10**, 738–752 (2014).
10. S Midtgaard, M Pedersen, L Arleth, Small-Angle X-Ray Scattering of the Cholesterol Incorporation into Human ApoA1-POPC Discoidal Particles. *Biophys. J.* **109**, 308–318 (2015).
11. AY Shih, IG Denisov, JC Phillips, SG Sligar, K Schulten, Molecular Dynamics Simulations of Discoidal Bilayers Assembled from Truncated Human Lipoproteins. *Biophys. J.* **88**, 548–556 (2005).
12. AY Shih, A Arkhipov, PL Freddolino, SG Sligar, K Schulten, Assembly of lipids and proteins into lipoprotein particles. *J. Phys. Chem. B* **111**, 11095–11104 (2007).
13. I Siuda, DP Tieleman, Molecular Models of Nanodiscs. *J. Chem. Theory Comput.* **11**, 4923–4932 (2015).
14. A Debnath, LV Schäfer, Structure and Dynamics of Phospholipid Nanodiscs from All-Atom and Coarse-Grained Simulations. *J. Phys. Chem. B* **119**, 6991–7002 (2015).
15. M Vestergaard, JF Kraft, T Vosegaard, L Thøgersen, B Schiøtt, Bicelles and Other Membrane Mimics: Comparison of Structure, Properties, and Dynamics from MD Simulations. *J. Phys. Chem. B* **119**, 15831–15843 (2015).
16. S Bibow, et al., Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. *Nat. Struct. Mol. Biol.* **24**, 187–193 (2017).
17. G Hummer, J Köfinger, Bayesian ensemble refinement by replica simulations and reweighting. *J. Chem. Phys.* **143**, 243150 (2015).
18. B Róycki, YC Kim, G Hummer, SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure* **19**, 109–116 (2011).
19. S Bottaro, G Bussi, SD Kennedy, DH Turner, K Lindorff-Larsen, Conformational ensembles of rna oligonucleotides from integrating nmr and molecular simulations. *Sci. Adv.* **4**, eaar8521 (2018).
20. S Bottaro, T Bengtsen, K Lindorff-Larsen, Integrating molecular simulation and experimental data: A bayesian/maximum entropy reweighting approach. *bioRxiv*, 457952 (2018).
21. S Orioli, AH Larsen, S Bottaro, K Lindorff-Larsen, How to learn from inconsistencies: Integrating molecular simulations with experimental data. *arXiv preprint arXiv:1909.06780* (2019).
22. N Skar-Gislinge, L Arleth, Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. *Phys. Chem. Chem. Phys.* **13**, 3161–3170 (2011).
23. MC Pedersen, L Arleth, K Mortensen, $WillItFit$: a framework for fitting of constrained models to small-angle scattering data. *J. Appl. Crystallogr.* **46**, 1894–1898 (2013).
24. S Tristram-Nagle, Y Liu, J Legleiter, JF Nagle, Structure of gel phase dmpc determined by x-ray diffraction. *Biophys. journal* **83**, 3324–3335 (2002).
25. JE Kohn, et al., Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci.* **101**, 12491–12496 (2004).
26. IG Denisov, MA McLean, AW Shaw, YV Grinkova, SG Sligar, Thermotropic phase transition in soluble nanoscale lipid bilayers. *The J. Phys. Chem. B* **109**, 15580–15588 (2005).
27. V Graziano, L Miller, L Yang, Interpretation of solution scattering data from lipid nanodiscs. *J. Appl. Crystallogr.* **51**, 157–166 (2018).
28. J Huang, et al., CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2016).
29. D Schneidman-Duhovny, M Hammel, JA Tainer, A Sali, Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974 (2013).
30. D Schneidman-Duhovny, M Hammel, JA Tainer, A Sali, FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.* **44**, W424–W429 (2016).
31. D Svergun, C Barberato, MH Koch, CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
32. A Cesari, A Gil-Ley, G Bussi, Combining simulations and solution experiments as a paradigm for rna force field refinement. *J. Chem. Theory Comput.* **12**, 6192–6200 (2016).
33. S Bottaro, K Lindorff-Larsen, Biophysical experiments and biomolecular simulations: A perfect match? *Science* **361**, 355–360 (2018).
34. TK Ritchie, et al., *Chapter 11 Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs.* (Elsevier Masson SAS) Vol. 464, pp. 211–231 (2009).
35. Avanti Lipids Polar Inc, 1,2-dimyristoyl-sn-glycero-3-phosphocholine (https://avantilipids.com/product/850345) (2018) [accessed 23 Sep. 2018].
36. GE Healthcare Life Sciences, Superdex 200 10/300 gl and 5/150 gl (https://www.gelifesciences.com/en/in/shop/chromatography/prepacked-columns/size-exclusion/superdex-200-10-300-gl-and-5-150-gl-p-05902) (2018) [accessed 23 Sep. 2018].
37. A Jordan, et al., SEC-SANS: size exclusion chromatography combined in situ with small-angle neutron scattering . *J. Appl. Crystallogr.* **49**, 2015–2020 (2016).
38. NT Johansen, MC Pedersen, L Porcar, A Martel, L Arleth, Introducing SEC-SANS for studies of complex self-organized biological systems. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 1178–1191 (2018).
39. C Dewhurst, Grasp v.7.15 (https://www.ill.eu/fr/users-en/scientific-groups/large-scale-structures/grasp/) (2017).
40. P Pernot, et al., Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).
41. D Orthaber, A Bergmann, O Glatter, SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **33**, 218–225 (2000).
42. CE Blanchet, et al., Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443 (2015).
43. S Hansen, Update for bayesapp : a web site for analysis of small-angle scattering data. *J. Appl. Crystallogr.* **47**, 1469–1471 (2014).
44. JS Pedersen, Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Adv. Colloid Interface Sci.* **70**, 171–210 (1997).
45. J Domański, PJ Stansfeld, MSP Sansom, O Beckstein, Lipidbook: A public repository for force-field parameters used in membrane simulations. *J. Membr. Biol.* **236**, 255–258 (2010).
46. CJ Dickson, L Rosso, RM Betz, RC Walker, IR Gould, GAFFlipid: a General Amber Force Field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter* **8**, 9617 (2012).
47. S Pronk, et al., GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
48. MJ Abraham, et al., Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
49. J Lee, et al., CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and

CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

50. G Bussi, D Donadio, M Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126** (2007).
51. LJ Heyer, S Kruglyak, S Yooseph, Exploring expression data: identification and analysis of coexpressed genes. *Genome research* **9**, 1106–1115 (1999).
52. RL Melvin, WH Gmeiner, FR Salsbury, All-atom molecular dynamics reveals mechanism of zinc complexation with therapeutic f10. *The journal physical chemistry. B* **120**, 10269–10279 (2016).
53. DI Svergun, et al., Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. United States Am.* **95**, 2267–2272 (1998).
54. J Tropp, Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: The effect of fluctuating internuclear distances. *The J. Chem. Phys.* **72**, 6035–6043 (1980).
55. TJ Piggot, JR Allison, RB Sessions, JW Essex, On the calculation of acyl chain order parameters from lipid simulations. *J. Chem. Theory Comput.* **13**, 5683–5696 (2017).