



Introducing SEC–SANS for studies of complex self-organized biological systems

Nicolai Tidemand Johansen,^a Martin Cramer Pedersen,^a Lionel Porcar,^b Anne Martel^{b*} and Lise Arleth^{a*}

^aNiels Bohr Institute, University of Copenhagen, Copenhagen, Denmark, and ^bInstitut Laue–Langevin, Grenoble, France.
*Correspondence e-mail: martela@ill.fr, arleth@nbi.ku.dk

Received 31 January 2018

Accepted 14 May 2018

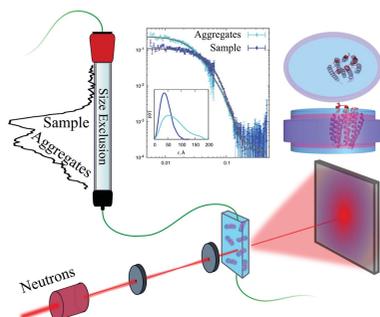
Edited by V. T. Forsyth, Institut Laue–Langevin, France, and Keele University, UK

Keywords: size-exclusion chromatography; small-angle neutron scattering; SEC–SANS; small-angle X-ray scattering; membrane proteins; phospholipid nanodiscs.

Small-angle neutron scattering (SANS) is maturing as a method for studying complex biological structures. Owing to the intrinsic ability of the technique to discern between ¹H- and ²H-labelled particles, it is especially useful for contrast-variation studies of biological systems containing multiple components. SANS is complementary to small-angle X-ray scattering (SAXS), in which similar contrast variation is not easily performed but in which data with superior counting statistics are more easily obtained. Obtaining small-angle scattering (SAS) data on monodisperse complex biological structures is often challenging owing to sample degradation and/or aggregation. This problem is enhanced in the D₂O-based buffers that are typically used in SANS. In SAXS, such problems are solved using an online size-exclusion chromatography (SEC) setup. In the present work, the feasibility of SEC–SANS was investigated using a series of complex and difficult samples of membrane proteins embedded in nanodisc particles that consist of both phospholipid and protein components. It is demonstrated that SEC–SANS provides data of sufficient signal-to-noise ratio for these systems, while at the same time circumventing aggregation. By combining SEC–SANS and SEC–SAXS data, an optimized basis for refining structural models of the investigated structures is obtained.

1. Introduction

In structural biology, X-ray crystallography is still the dominating source of information. However, many proteins, for example membrane proteins, are notoriously difficult to crystallize, and intrinsically disordered proteins or proteins with large intrinsically disordered regions may not be crystallizable or accessible by high-resolution cryo-electron microscopy (cryo-EM). Structural information on these types of systems is sparse, although the potential impact from new knowledge is high: approximately one quarter of the genome in eukaryotes encodes membrane proteins (Fagerberg *et al.*, 2010), which are the target of an estimated 60% of all drugs (Overington *et al.*, 2006). Similarly, approximately one third of the genome in eukaryotes codes for intrinsically disordered proteins (Oldfield *et al.*, 2005), which mediate important regulatory interactions (Oldfield & Dunker, 2014). The term ‘integrative structural biology’ (Ward *et al.*, 2013) was recently coined to describe an approach in which a high-resolution structural model is refined against information from several complementary techniques, both computational and experimental, with the aim of obtaining a more complete description of the structure and dynamics of a system. For this purpose, high-resolution techniques such as crystallography, nuclear



magnetic resonance spectroscopy (NMR) and cryo-EM are optimally combined with more easily accessible information from more low-resolution techniques such as negative-stain EM and small-angle scattering (SAS), as well as various computational approaches for describing the relevant structure or ensemble of structures.

With no inherent size limitation, as is the case for NMR spectroscopy and cryo-EM, SAS is a widely applied method for obtaining low-resolution structural information on biological systems in solution (Svergun & Koch, 2003). Small-angle X-ray scattering (SAXS) has gained popularity in biology over the past few decades, facilitated by excellent synchrotron sources and the development of easily accessible software (Schneidman-Duhovny *et al.*, 2010; Förster *et al.*, 2010; Graewert & Svergun, 2013; Skou *et al.*, 2014; Hopkins *et al.*, 2017; *SasView*, <http://www.sasview.org/>). As part of this development, SEC-SAXS has been introduced and is now a standard method to circumvent problems of aggregating samples (Mathew *et al.*, 2004; David & Pérez, 2009; Pernot *et al.*, 2013; Graewert *et al.*, 2015; Jeffries *et al.*, 2016; Ryan *et al.*, 2018). SEC-SAXS combines size-exclusion chromatography (SEC) and SAXS in such a way that the SEC purification is coupled directly to the capillary in the X-ray beam, such that sample eluting from the column is illuminated immediately.

During this development, small-angle neutron scattering (SANS) on biological systems has remained a more exclusive technique, mainly owing to the more limited access to neutrons. However, SANS is highly complementary to SAXS due to the possibility for contrast variation through $^1\text{H}/^2\text{H}$ substitution. This has facilitated structural studies of multi-component complexes containing one or more components of protein, lipid and nucleic acids, such as DNA-protein complexes, membrane proteins in detergent micelles and quarternary protein complexes (Whitten *et al.*, 2007; Niemann *et al.*, 2008; Appolaire *et al.*, 2014; Gabel, 2015; Zaccai *et al.*, 2016; Midtgaard *et al.*, 2018). An additional advantage is that radiation damage, which is often an issue in synchrotron SAXS, is not a problem in SANS. Despite these advantages, SANS is more challenging than SAXS. The required sample concentration and volume are higher, the exposure times are longer and, to obtain the best signal-to-noise ratio, samples are usually measured in 100% D_2O . Although chemically very similar to H_2O , this slightly different solvent often promotes protein aggregation (Larsson, 1988; Gripon *et al.*, 1997).

A membrane protein in solution is inherently a multi-component system. To remain soluble, the membrane protein must be stabilized by an amphiphilic entity such as a detergent micelle or a lipid bilayer that shields the hydrophobic trans-membrane region of the membrane protein. Detergents provide the most simple reconstitution system in terms of sample preparation, but the polydispersity of the resulting micelle around the membrane protein and the existence of excess micelles decrease the theoretical structural resolution that is obtainable with this approach, although several methods for clever sample preparation and data treatment have been developed (Gabel *et al.*, 2014; Oliver *et al.*, 2017; Koutsioubas, 2017; Midtgaard *et al.*, 2018).

Nanodiscs can be utilized to obtain more monodisperse samples containing membrane proteins embedded in native-like environments (Bayburt *et al.*, 2002). A traditional nanodisc consists of two membrane-scaffold proteins (MSPs) encircling a disc-shaped patch of phospholipid bilayer that can serve as a carrier for a membrane protein. Attempts at crystallizing nanodiscs or membrane proteins in nanodiscs have so far been unsuccessful, but nanodiscs have recently been used to facilitate membrane-protein crystallization (Nikolaev *et al.*, 2017). The diameter of the nanodisc is directly correlated to the length of the MSP (Denisov *et al.*, 2004; Hagn *et al.*, 2013), and NMR and spin-labelling techniques have recently been combined to describe the first high-resolution structure of an 8 nm nanodisc based on a short MSP (Bibow *et al.*, 2017).

SAXS and SANS have previously been combined to study bacteriorhodopsin inside a nanodisc (Kynde *et al.*, 2014), where a combined analysis, *i.e.* global fitting, of SAXS and SANS data was desired because of the high complementarity of the obtained information. The protein component of the nanodisc had higher excess scattering-length density in the X-ray contrast, whereas the lipid component had higher excess scattering-length density in the SANS contrast in 100% D_2O . The simultaneous fitting of the model to the two contrasts allowed a more detailed low-resolution structure of the system to be refined and gave information about both the membrane protein and the surrounding lipid membrane structure.

Because of the structural complexity of the nanodisc system, the analysis of SAS data is not straightforward and generally relies on data from structurally homogeneous/monodisperse samples. However, sample preparation of such homogeneous nanodisc samples is challenging, and until now, it has been our experience that the success rate of obtaining good SANS data from these systems is relatively low owing to combined issues with membrane-protein reconstitution and the above-mentioned increased aggregation in D_2O -based buffers. Despite the higher sample requirements than for SAXS, the feasibility of performing online SEC-SANS with relative short exposure times has recently been demonstrated on a set of water-soluble proteins (Jordan *et al.*, 2016). Overall, the expected outcome for SEC-SANS (and SEC-SAXS) is higher sample homogeneity, which leads to a scattering signal that is not smeared because of structural dispersity. In turn, structural models can be refined to better resolution.

In the present study, we aimed to test and benchmark the new SEC-SANS setup in combination with complex self-assembled systems: nanodiscs without membrane proteins (empty nanodiscs) and membrane-protein-loaded nanodiscs (loaded nanodiscs). Using SANS data from monodisperse samples with good signal-to-noise ratios in an investigated q -range up to $q \approx 0.15 \text{ \AA}^{-1}$, we show that the SEC-SANS method is applicable to and relevant for studying these types of samples, which are usually difficult to produce in a sufficiently pure and homogenous form in the required amounts and concentrations. The SEC-SANS data are complemented by SEC-SAXS data, and we model empty nanodiscs and the membrane protein proteorhodopsin (PR) inside a nanodisc, respectively, with global fits to the SANS and SAXS data sets.

2. Sample preparation

2.1. Materials

All chemicals were obtained from Sigma–Aldrich. The salts used in the SEC–SAS buffers were obtained in the anhydrous forms. 99.9% D₂O was obtained from Cambridge Isotope Laboratories.

2.2. Protein expression and purification

The plasmids encoding the genes for MSPs were obtained from Professor Steven Sligar (University of Illinois, USA), the plasmid encoding PR was obtained from Professor Richard Neutze (University of Gothenburg, Sweden) and the plasmid encoding *Thermotoga maritima* CorA (TmCorA) was obtained from Professor Mikaela Rapp (University of Stockholm, Sweden). MSPs were expressed and purified essentially as described elsewhere (Ritchie *et al.*, 2009), but under denaturing conditions. TmCorA was expressed and purified using a slightly modified protocol to that reported elsewhere (Palombo *et al.*, 2012), and PR was expressed and purified essentially as described elsewhere (Gourdon *et al.*, 2008). A detailed description of the preparation of all proteins can be found in the supporting information. The MSP used for loaded nanodiscs had the N-terminal His₆ tag cleaved off with *Tobacco etch virus* (TEV) protease before the reconstitution procedure described below. The N-terminal His₆ tag was not cleaved off from the MSPs used for empty nanodiscs.

2.3. Preparation of empty and loaded nanodiscs

Empty nanodisc samples were prepared off-site one week in advance, although without the final SEC step. Briefly, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC; Avanti Polar Lipids) lipids were dissolved to a final concentration of 50 mM in gel-filtration buffer (20 mM Tris–HCl pH 7.5, 100 mM NaCl) containing 100 mM sodium cholate. The lipid mixture was mixed with MSP and diluted to 10 mM lipid concentration with gel-filtration buffer and incubated briefly at room temperature. Damp Amberlite XAD-2 detergent-absorbing beads (Sigma–Aldrich) were added to approximately 15% (w/v) and the sample was incubated at 5°C for 4 h under fast agitation. The detergent-absorbing beads were removed and the sample was purified by SEC on a Superdex 200 10/300 GL column (GE). A 1 ml peak fraction was collected, flash-frozen and stored at –80°C.

For loaded nanodiscs, the membrane protein of interest was added to the reconstitution mixture, displacing some volume of gel-filtration buffer, and allowed to equilibrate for 1 h before the addition of detergent-absorbing beads. To avoid excess membrane protein and reduce the amount of oligomeric membrane protein, PR was mixed with MSP1D1 in a 1:16 ratio with a POPC:MSP1D1 ratio of 70:1, whereas TmCorA was mixed with MSP1E3D1 in a 1:20 ratio with a POPC:MSP1E3D1 ratio of 130:1. Empty nanodiscs were separated from the loaded nanodiscs on a 1 ml HisTrap column (GE), utilizing the His₆ tag on the membrane protein. Empty nanodiscs flowing through the column were discarded,

Table 1

Sample volumes and associated absorptions measured at 280 nm (A_{280}) for the samples investigated in this study.

| Sample | A_{280} | Volume (μ l) | |
|--------------------------------|-----------|-------------------|------|
| | | SANS | SAXS |
| MSP1D1–POPC nanodiscs | 1.30 | 200 | 200 |
| MSP1D1–POPC–PR nanodiscs | 3.34 | 250 | 160 |
| MSP1E3D1–POPC–TmCorA nanodiscs | 3.64 | 160 | 140 |

and the bound loaded nanodiscs could be eluted from the column with gel-filtration buffer containing 250 mM imidazole. The samples were diluted twofold in gel-filtration buffer and concentrated to approximately 1 ml in a 10 kDa molecular-weight cutoff (MWCO) spin filter. Each sample was split into two for SANS and SAXS measurements, flash-frozen and stored at –80°C. Samples were transported to the Institut Laue–Langevin (ILL) or the ESRF on dry ice and were stored at –80°C until use. Before measurement, the samples were thawed and centrifuged for 10 min at 14 000g. The final volumes and concentrations of the injected samples for SEC–SANS and SEC–SAXS are shown in Table 1.

3. Methods and instrumentation

3.1. Dedicated SEC–SANS setup on D22

SANS data were recorded on the D22 small-angle scattering diffractometer at the ILL, Grenoble, France. The online SEC (Fig. 1) was performed using a modular HPLC system (Serlabo), including two BlueShadow pumps (Knauer), an active mixing valve, an automated sample loader (ALIAS), a Foxy R1 fraction collector (Teledyne ISCO) and a SmartLine 2600 diode-array spectrophotometer (Knauer). The spectrophotometer was connected *via* optic fibres either to an optic cell (3 mm path length, used in measurements of empty nanodiscs) placed at the outlet of the chromatography column or directly to the middle of the SANS measurement cell, perpendicularly to the neutron beam (10 mm path length, used in measurements of loaded nanodiscs), enabling the simultaneous recording of chromatograms at four different wavelengths, in this case 220, 280, 535 and 600 nm.

The setup was equipped with a Superdex 200 10/300 GL column (GE) with a void volume of approximately 7.5 ml. The data were obtained during two separate beam times. Empty nanodisc data were obtained in the first session and loaded nanodisc samples in the second session. For the empty nanodiscs a flow rate of 0.3 ml min^{–1} was used, whereas loaded nanodisc samples were pumped at 0.5 ml min^{–1} in most of the column void volume; this was reduced to 0.2 ml min^{–1} during sample elution. Furthermore, all components of the HPLC setup, including the buffers and the column, were placed in a closed cabinet connected to an air-cooling system set to 10°C to control the temperature. For empty nanodiscs, the buffer consisted of 20 mM Tris–DCI pH 7.5, 100 mM NaCl and for PR-loaded nanodiscs the buffer consisted of 20 mM Tris–DCI pH 7.5, 300 mM NaCl. For TmCorA-loaded nanodiscs, 100 mM MgCl₂ was added to the latter buffer.

The experiments were carried out with a nominal neutron wavelength (λ) of 6.0 Å, a wavelength distribution ($\Delta\lambda/\lambda$) of 10% full-width at half-maximum, a rectangular collimation of $40 \times 55 \text{ mm}^2$ and a rectangular sample aperture of $7 \times 10 \text{ mm}^2$. Three sample-to-detector distances (11.2, 5.6 and 1.4 m, with collimation lengths of 11.2, 5.6 and 2.8 m, respectively) were used for characterization of the empty nanodiscs, covering a momentum-transfer range q of 0.0044 \AA^{-1} to 0.61 \AA^{-1} ; $q = 4\pi\sin(\theta)/\lambda$, where θ is half of the angle between the incoming and the scattered neutrons. For the measurement of loaded nanodiscs, sample availability limited the number of possible sample-to-detector distances to two: 11.2 and 2.0 m

(with collimation lengths of 11.2 and 2.8 m, respectively), giving a q -range of $0.0044\text{--}0.46 \text{ \AA}^{-1}$. The measured intensities were binned into 30 s frames.

Sample transmission was approximated by the buffer measured at a sample-to-detector distance of 11.2 m. The measured intensity, $I(q)$, was brought to an absolute scale in units of scattering cross-section per unit volume (cm^{-1}) using the direct beam flux measured for each collimation length prior to the experiment. In this way, problems with mismatch between intensities measured at different detector settings were avoided. Data reduction was performed using the *GRASP* software (Dewhurst, 2017).

3.2. SEC-SAXS data acquisition

SEC-SAXS was performed using the BioSAXS instrument on BM29 at the ESRF, Grenoble, France (Pernot *et al.*, 2013). This setup includes an HPLC controlled separately from the SAXS measurement, coupled to a UV-Vis array spectrophotometer measuring absorption from 190 to 800 nm. Data were collected with an X-ray wavelength of 0.9919 \AA at momentum transfers q ranging from 0.003 to 0.49 \AA^{-1} . The capillary was cooled to 10°C , but with the HPLC including the SEC column placed at ambient temperature. Experiments were run on the same column and in H_2O -based solvents containing the same components as in SEC-SANS at flow rates of 0.5 ml min^{-1} for empty nanodiscs and 0.4 ml min^{-1} for loaded nanodisc. Data reduction was carried out using the in-house software at the beamline, and subsequent conversion to absolute units was performed with water as a calibration standard. The 1 s frames recorded were averaged in 10 s bins; this is the typical measurement time for static SAXS samples on BM29.

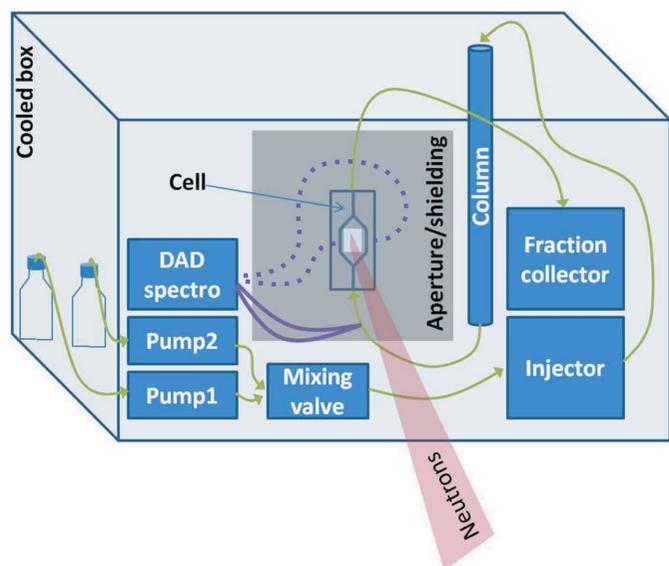
3.3. SEC-SAS data processing

3.3.1. Selection of sample frames. Scattergrams, *i.e.* scattering intensity as a function of elution volume, were obtained by normalizing the total intensities of the individual frames. Chromatograms of absorption at 280 nm were normalized, and for the setups with the UV cell located before the sample cell, the chromatograms were shifted to align with the scattergrams. Where measured, absorptions at wavelengths other than 280 nm were normalized to the absorption at 280 nm. The radius of gyration (R_g) of background-subtracted data was calculated for individual frames with *AutoRg* (Petoukhov *et al.*, 2007). The interval used for extracting the optimal scattering data was chosen to have a converged R_g within error bars in a window covering approximately the right-side half of the desired SEC peak, corresponding to a typical fraction collected for other types of experiments to avoid the overlap of larger particles. The selected frames were averaged to produce a single scattering curve. For SAXS, the averaged data were logarithmically rebinned.

Pair distance [$p(r)$] distributions were calculated by inverse Fourier transformations (IFTs; Glatter, 1977) using the *BayesApp* software available online (Hansen, 2014). For empty nanodiscs, $p(r)$ distributions were calculated for the



(a)



(b)

Figure 1

(a) Photograph of the setup implemented at D22. (b) Schematic describing the SEC-SANS setup, with the different components placed identically to those in the photograph in (a). The scheme highlights the two possible options for UV-visible light absorbance measurement: either through a specific cell (optic fibres drawn as continuous purple lines) or directly across the SANS measurement cell (optic fibres drawn as purple dotted lines). The green lines represent the solution flow.

individual frames of the selected interval, whereas they were only calculated for averaged data for loaded nanodiscs. The maximum pair distances (D_{\max}) and R_g values presented in Figs. 2(g), 3(e) and 4(e) were obtained from this IFT analysis.

Sample concentrations were estimated from the online absorption measurements at 280 nm divided by sample-specific molar extinction coefficients calculated using *ProtParam* on the ExPASy server (Gasteiger *et al.*, 2005). For the empty nanodiscs, however, better estimates of the concentration were obtained from the values of the fractional area of the normalized absorption measured at 280 nm, the volume measured and the total injected amount of sample. This estimate was only possible for the empty nanodiscs, which were already monodisperse before applying them to the column, and thus not for the loaded nanodiscs.

3.3.2. Modelling of empty nanodiscs. The recorded data were analyzed by refining the analytical model for the nanodisc presented in the literature (Skar-Gislinge *et al.*, 2010, 2011) using the data (Fig. 5a). The model was simultaneously refined using the SAXS and SANS data by minimizing the combined χ^2 of both data sets. The source code for the models and descriptions thereof, as well as the optimization routines, can be found in the literature (Pedersen *et al.*, 2013).

In this model, the nanodisc is described by a combination of analytical form factors (Pedersen, 1997), each of which are endowed with the scattering properties of the molecules they are meant to represent. Mathematically, the model has the form

$$I(q) = nP(q), \quad (1)$$

where $P(q)$ is the form-factor intensity for a single nanodisc and n is the number density. $P(q)$ describes the scattering properties of an orientationally averaged (over the two angles α and β) nanodisc described by the complex form-factor amplitude $F(\mathbf{q})$,

$$P(q) = \int_0^\pi \sin \alpha \, d\alpha \int_0^{2\pi} F(\mathbf{q})F^*(\mathbf{q}) \, d\beta, \quad (2)$$

where the vector \mathbf{q} can be expressed in spherical coordinates (q, α, β), with q being the norm. The asterisk denotes complex conjugation.

The form-factor amplitude of the (oriented) nanodisc is described by a sum of contributions accounting for the different components in the nanodisc,

$$\begin{aligned} F(\mathbf{q}) = & \Delta\rho_{\text{protein belt}} V_{\text{protein belt}} A(\mathbf{q})_{\text{protein belt}} \\ & + \Delta\rho_{\text{histidine tag}} V_{\text{histidine tag}} A(\mathbf{q})_{\text{histidine tag}} \\ & + \Delta\rho_{\text{lipid headgroups}} V_{\text{lipid headgroups}} A(\mathbf{q})_{\text{lipid headgroups}} \\ & + \Delta\rho_{\text{lipid alkyl groups}} V_{\text{lipid alkyl groups}} A(\mathbf{q})_{\text{lipid alkyl groups}} \\ & + \Delta\rho_{\text{lipid methyl groups}} V_{\text{lipid methyl groups}} A(\mathbf{q})_{\text{lipid methyl groups}}, \end{aligned} \quad (3)$$

where the $\Delta\rho$ terms are the excess scattering-length densities of the given component, the V terms are the (partial specific) volumes of the components and lastly the $A(\mathbf{q})$ terms are the (normalized) form factors describing the shape of the given

component. As an example, the form factor for the lipid methyl groups, $A(\mathbf{q})_{\text{lipid methyl groups}}$ is that of a flat cylinder with elliptical cross-section (see, for example, Pedersen, 1997). The parameters in the model describe these shapes, their volumes and their scattering properties. The dimensions of the aforementioned cylinder depends on the area per lipid headgroup, the number of lipids in the disc, the volume of a POPC molecule and the axis ratio of the disc, while the relevant excess scattering-length density, $\Delta\rho$, may be calculated from the chemical composition and molecular volume of the constituents (see, for example, Skar-Gislinge *et al.*, 2010). In this way, by systematically using molecular constraints, the number of fit parameters to describe the model could be kept at a minimum (see Fig. 4e and Section 4).

In the refinement process, the height of the MSP belt was fixed to 25.78 Å to match the recently published high-resolution structure of a nanodisc (Bibow *et al.*, 2017). Similarly, the R_g of the protruding His₆ tags was fixed to 12.70 Å in line with previous studies of disordered proteins (Kohn *et al.*, 2004). It was assumed that 50% of the exchangeable H atoms in the MSP1D1 proteins had exchanged with the solvent (100% D₂O) in the SANS experiments, in line with the literature (Morgan *et al.*, 2011). The effects of instrumental smearing was accounted for in the fits in accordance with the approaches outlined in the literature (Pedersen *et al.*, 1990).

The minimization was performed using the Levenberg–Marquardt (Levenberg, 1944; Marquardt, 1963) algorithm in *WillItFit* (Pedersen *et al.*, 2013), which was also used to compute the estimated refinement errors.

3.3.3. Modelling of PR-loaded nanodiscs. A model analogous to the model of bacteriorhodopsin in a nanodisc presented in the literature (Kynde *et al.*, 2014) was refined from the data obtained from monomeric PR-loaded nanodiscs (Fig. 5b). Apart from the analytical description of the nanodisc described above, the model was based on the published solution NMR structure of PR (Reckel *et al.*, 2011; PDB entry 2l6x) in the Protein Data Bank (Berman *et al.*, 2000), acquired from the Orientations of Proteins in Membranes (OPM) database (Lomize *et al.*, 2006). For reference, the OPM database lists the hydrophobic height of the protein as 27.8 ± 1.3 Å.

H atoms were added to the structure using the built-in algorithms in the *PHENIX* software suite (Adams *et al.*, 2010) and solvent exchange was accounted for using the same tool. Based on some of the few available data (Hebling *et al.*, 2010; Mehmood *et al.*, 2012), it was assumed that no solvent exchange had occurred in the transmembrane part of the protein, whereas it was assumed that 100% of the exchangeable H atoms in the flexible intracellular and extracellular loops of the protein had exchanged. In the calculations of scattering intensities, the PR structure was coarse-grained to amino-acid residue level for computational efficiency. As explained in the previous section, the height of the MSP belt was fixed during the refinement process.

The model of PR in a nanodisc is similar to that presented in (1)–(3). We simply add another contribution from the PR to (3) as described previously (Kynde *et al.*, 2014).

4. Results

4.1. Empty nanodiscs: feasibility of the SEC–SANS setup

To test and benchmark the new dedicated SEC–SANS setup at D22 (Fig. 1), empty nanodiscs were first measured. For all three detector settings, the scattergrams in terms of the normalized total scattering intensity are plotted as a function of retention volume along with the normalized chromatogram measured at 280 nm and the calculated R_g values (Fig. 2*a*). The three scattergrams and the chromatogram all feature a well defined single peak, indicating a monodisperse sample, as well as a flat region corresponding to the buffer, which was used for background subtraction. Fig. 2*b*) shows a complementary SEC–SAXS data set measured on the same sample as used for the SEC–SANS data, although in an H₂O-based buffer. The SAXS data in Supplementary Fig. S1 show that the structure of the empty nanodisc is preserved after exchange from H₂O to D₂O.

The normalized and scaled scattering data of the individual frames for SANS and SAXS, respectively, are shown in Fig. 2*d*), with the top data set corresponding to the first frame of the selected interval and the rest in the order of elution. As they were very noisy at $q > 0.1 \text{ \AA}^{-1}$, the data were plotted without error bars to make the individual data sets easily separable for the eye. All scattering data exhibit a well defined Guinier region (a flat low- q region on a double-logarithmic scale) down to the minimum q value, indicating no aggregation or presence of larger particles. Guinier plots for the averaged SANS and SAXS data (Fig. 5*a*), respectively, are shown in Supplementary Figs. S2*a*) and S2*b*) and confirm this observation.

$p(r)$ distributions for the individual frames are shown in Fig. 2*f*). Overall, the $p(r)$ distributions for the different frames are rather similar and terminate at approximately the same D_{max} in SANS and SAXS, respectively. This shows that the samples are rather monodisperse within the resolution of

the SEC and that the SANS and SAXS samples indeed have the same underlying structure, forming an optimal basis for a subsequent simultaneous analysis. The values for the concentrations calculated from the absorption at 280 nm, as well as the D_{max} and R_g values derived from the $p(r)$ distributions, are summarized in Fig. 2*g*).

A more careful inspection of the data reveals that while the individual SANS data sets appear to be highly similar, the SAXS data sets have a small but clear systematic variation as a function of retention volume (Fig. 2*d*). Inspection of the $p(r)$ distributions (Fig. 2*f*) and the characteristic sizes (Fig. 2*g*) confirms this apparent change in both SANS and SAXS as seen by the R_g decreasing systematically from 31.2 to 27.8 Å for the SANS data and from 45.4 to 40.1 Å for the SAXS data. Hence, the SEC–SANS and SEC–SAXS analyses provide additional information by revealing a small but significant structural dispersion for the empty nanodiscs. Despite these small variations of the particle shape over the SEC peaks, averaged data sets were used for modelling as described in Section 4.4 in order to improve the statistics.

4.2. PR-loaded nanodiscs: separation of different species

Next, PR-loaded nanodiscs were measured by SEC–SAS. As for the empty nanodiscs, scattergrams and chromatograms (280 and 535 nm) were plotted for both SEC–SANS and SEC–SAXS, yielding two dominating but overlapping peaks eluting after the column void volume and with approximately 3 ml separation between their peaks (Fig. 3*b*). PR has a light red color owing to the absorption of green light, and the chromatogram measured at 535 nm confirmed the presence of PR in both peaks. Clearly, the two peaks observed in the scattergrams and chromatograms were not separated sufficiently to extract a monodisperse sample from the relatively broad first peak (peak 1). The selected frames were picked from the middle of the peak and processed to obtain an idea of

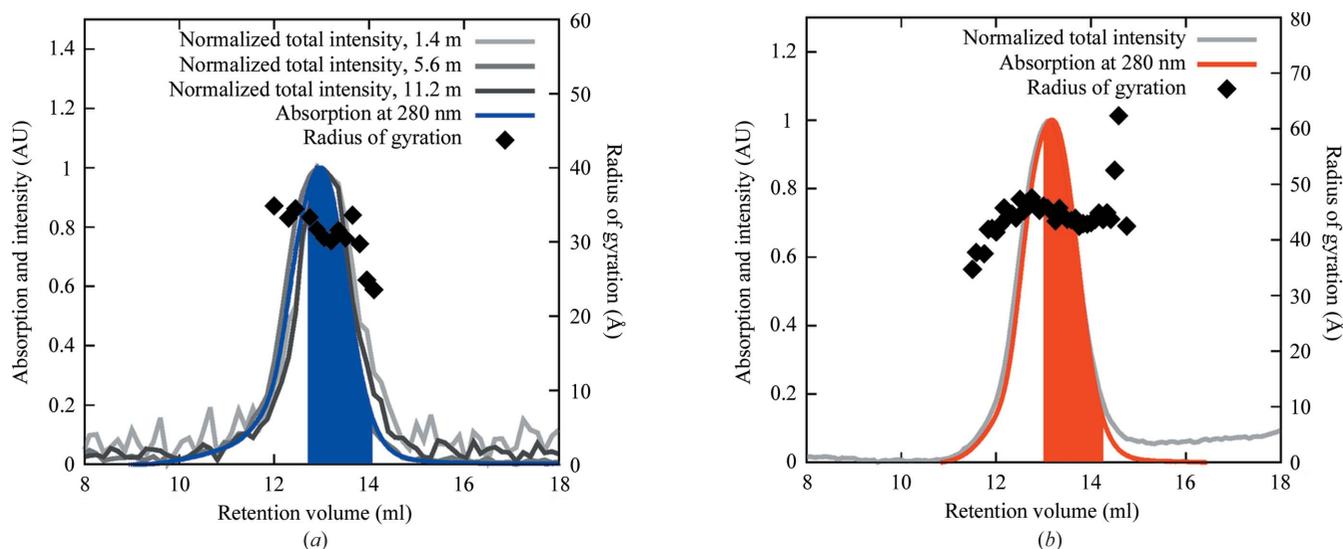


Figure 2

(*a*, *b*) SEC–SAS data for the empty nanodiscs, including scattergrams and the chromatograms measured at 280 nm as well as calculated R_g values. The colored areas indicate the intervals from which individual frames were taken.

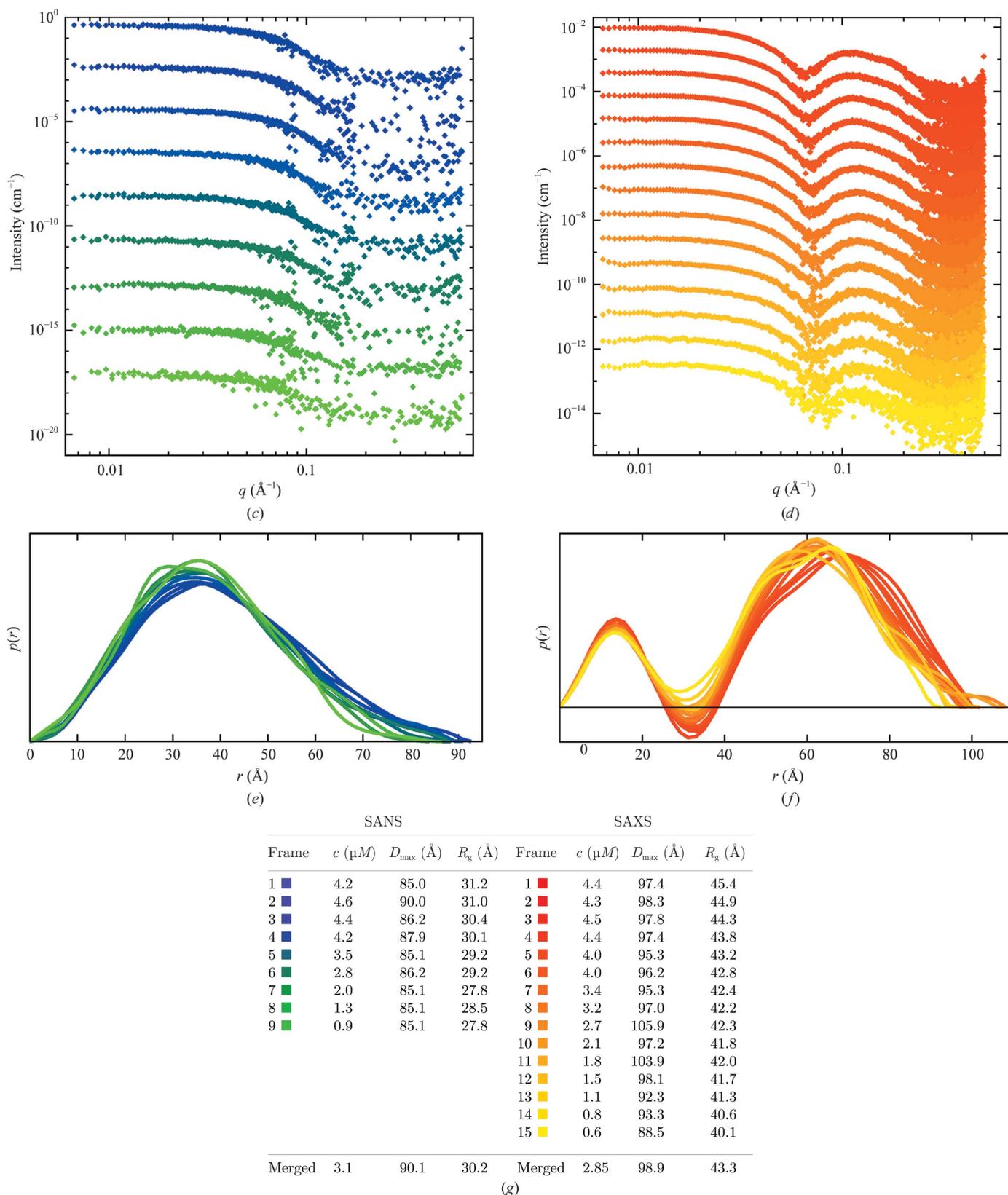


Figure 2 (continued)

(c, d) Reduced and background-subtracted SAS data for the individual frames taken from the indicated intervals in (a) and (b) shown without the associated error bars. The topmost data sets are presented on an absolute scale, whereas the others have been rescaled by a factor of 100 (SANS data) and a factor of five (SAXS data) relative to the data set above. (e, f) p(r) distributions refined from each of the data sets in (c) and (d), respectively. The averaged data are shown in Fig. 5(a). (g) Parameters from *BayesApp* refined from the IFT of the data.

the overall size of the particles that were present, but without further interpretation. To minimize interference from the larger particles present in peak 1, the frames from the second peak (peak 2) were chosen from the right side of this peak.

The obtained averaged data sets for both peaks are plotted along with the fits of the IFT analysis in Fig. 3(d), with the derived $p(r)$ distributions shown in the insets. The parameters refined from the IFT analysis are summarized in Fig. 3(e). Owing to the low sample concentration, the SANS data are

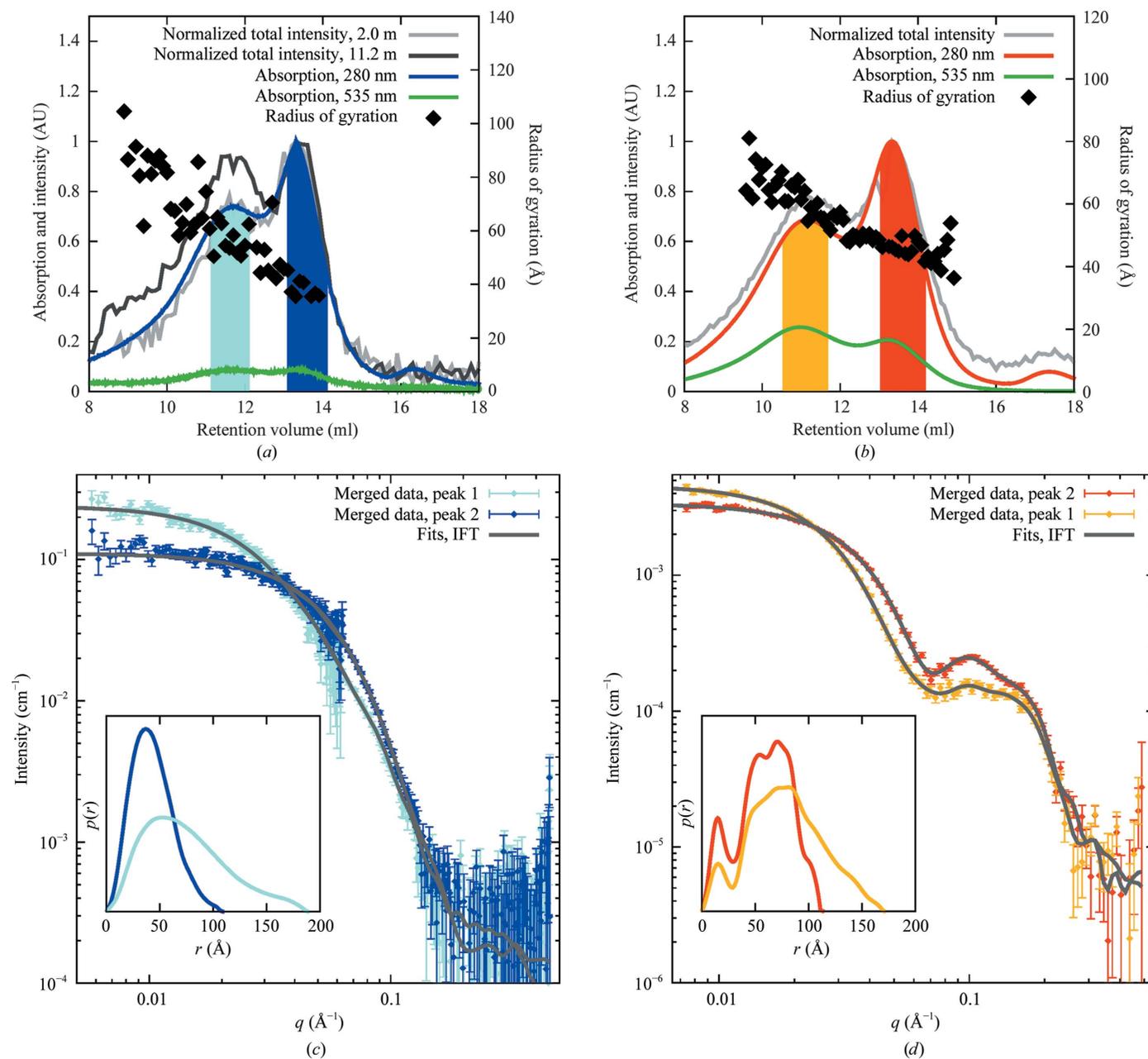


Figure 3 (a, b) SEC-SAS data from the experiment on PR-loaded nanodiscs as presented for empty nanodiscs in Figs. 2(a) and 2(b). (c, d) Averaged data for the frames recorded across the highlighted peaks in (a) and (b), respectively. Insets: $p(r)$ distributions refined from the data. The data sets collected from peak 2 were further analyzed as explained in the text and shown in Fig. 5(b). (e) Parameters from *BayesApp* refined from the IFT of the data.

slightly noisy in both the low- q and high- q regions, but nevertheless feature well determined Guinier regions along with significant structural differences between the SANS data obtained from the two peak populations (Fig. 3c and Guinier plots in Fig. 2b). The SAXS data also feature well defined Guinier regions for both samples as well as oscillating features at around $q \simeq 0.1 \text{ \AA}^{-1}$ (Fig. 3d). For the sample in peak 2, the D_{max} and R_g values are only slightly larger than for the empty nanodisc (Fig. 2g), whereas for the sample in peak 1 they are

much larger. These sizes indicate that the sample from peak 2 is a nanodisc, whereas the sample from peak 1 is likely to be a polydisperse and larger aggregate structure. Further modelling and data interpretation was only carried out for the sample in peak 2, as described in Section 4.4.

4.3. TmCorA-loaded nanodiscs: removal of aggregates

Finally, to test the SEC-SANS method on a larger membrane protein with large intracellular domains, TmCorA-

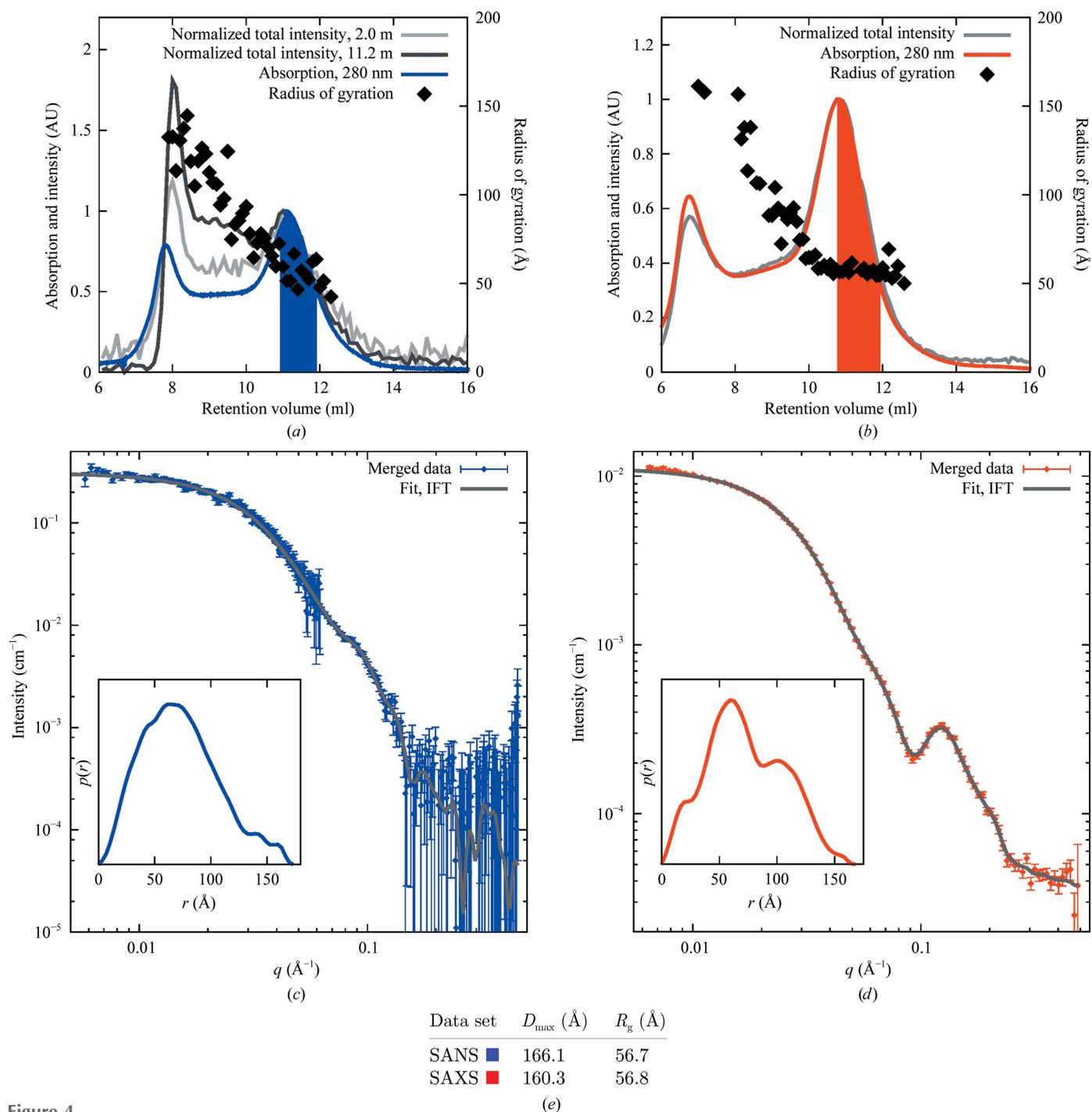


Figure 4 (a, b) SEC-SANS data from the experiments on TmCorA-loaded nanodiscs are shown in the same manner as in Figs. 2 and 3. (c, d) Averaged scattering data. Insets: $p(r)$ distributions refined from the averaged data. (e) Parameters from *BayesApp* refined from the IFT of the data.

loaded nanodiscs were measured. Scattergrams and chromatograms are plotted for both SEC-SANS and SEC-SAXS together with the calculated R_g values as shown in Fig. 4(b), yielding a dominating peak at 11.5 ml with a trailing signal into a peak in the void volume of the column.

Averaged data sets calculated from the frames in the selected interval are plotted together with the fit of the IFT analysis in Fig. 4(d) along the refined $p(r)$ distributions shown in the insets. The SANS data have reasonable statistics in the low- q region and display a well defined Guinier region (Guinier plots are shown in Supplementary Figs. S2a and S2b). Furthermore, a small distinct feature is observed at $q \approx 0.08 \text{ \AA}^{-1}$. The SAXS data are also of good quality, with a well defined Guinier region and several features in the q -region above 0.04 \AA^{-1} .

TmCorA has been proposed to undergo large conformational changes upon the binding of magnesium ions (Matthies

et al., 2016), making it an ideal system to probe with SANS and SAXS. In this work, we stay with the observation that the samples are suited for SEC-SAS and leave out further modelling. This will form part of future work focusing on TmCorA alone.

4.4. Modelling: joint refinement from SEC-SANS and SEC-SAXS data

The nanodisc model was refined from the averaged data of the empty nanodisc (Fig. 5a). The refined model is sketched as an inset in Fig. 5(a) and the refined parameters are shown in Table 2. The model resulted in a combined χ^2 of 8.76 and the presented confidence intervals are based on this value (Pedersen *et al.*, 2014). Apart from the listed parameters, a constant background (close to zero) and a term accounting for interface roughness (Als-Nielsen & McMorrow, 2011; typically

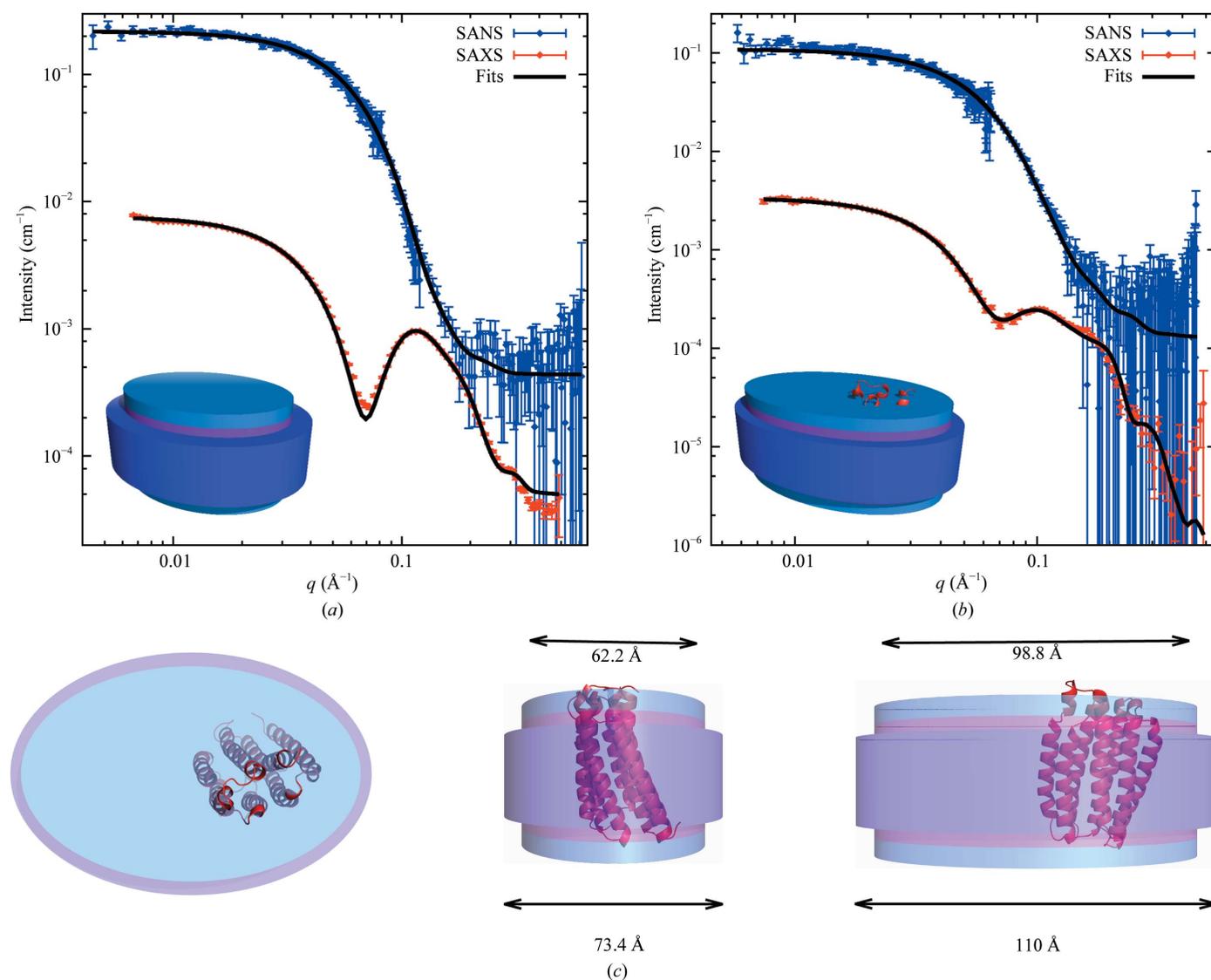


Figure 5
 (a) Averaged data from empty nanodiscs (Fig. 2d) along with the fits of the nanodisc model shown in the inset. The SAXS data have been logarithmically rebinned. The His₆ tags on the MSP1D1 proteins, modelled as random coils attached to the protein belt, are not shown in the rendering. (b) Data from PR-loaded nanodiscs (Fig. 3d) with the fits of a model of PR embedded in a nanodisc as shown in the inset. (c) The refined model of PR in the nanodisc shown from different perspectives.

up to 7 Å) were refined for each of the data sets, giving a total of four fit parameters in addition to those listed in Table 2.

The PR nanodisc model was refined from the data for the PR-loaded nanodisc eluting in peak 2 (Fig. 3*b*). The data with the associated fits, together with a rendering of the refined model, are presented in Fig. 5(*c*), and the associated parameters are shown in Table 2. The combined χ^2 of the presented model is 1.56 and, as before, the presented parameter confidences are based on this value. As for the empty nanodisc, the models include four fit parameters that are not listed in Table 2.

By requiring our models to reproduce the trends of the SANS as well as the SAXS data, we constrain the number of feasible models in our search by effectively enabling the rejection of models with greater confidence owing to the increased amount of data. This effect is amplified by the different contrasts in SAXS and SANS data, models refined from each of which predict some parameters better than others. By combining sets of data, we harvest the predictive power of both.

As outlined in the literature (Pedersen *et al.*, 2014), the effect of refining models simultaneously from several sets of data is improved confidence intervals of the refined parameters. As an example, in our case, owing to the different contrasts of the constituents of the nanodisc, adding SANS data to the refinement improves our confidence on the number of lipids in the nanodisc in particular: refining the nanodisc model from only the presented SAXS data results in an estimate of 106 ± 8.0 lipids per nanodisc, whereas refining including the SANS data in the model refinement yields an estimate of 117 ± 3.1 for the same parameter.

5. Discussion

5.1. Benchmarking the SEC–SANS setup on empty nanodiscs

The nanodisc is a popular platform for reconstituting membrane proteins in a native-like environment and studying their structure and function (Denisov & Sligar, 2016). However, samples of membrane proteins inside nanodiscs are demanding to produce in large amounts, and monodisperse samples are difficult to achieve, especially in D₂O-based solvents, because of issues of increased aggregation. This makes them difficult to study by SANS. Using a newly developed SEC–SANS setup on D22 at the ILL (Jordan *et al.*, 2016), we obtained SANS data with a sufficiently high signal-to-noise ratio from monodisperse samples of nanodiscs, both empty (Fig. 2*a*) and loaded with the membrane proteins PR (Fig. 3) and TmCorA (Fig. 4), respectively, in 100% D₂O. This shows that the method is practically applicable to such complex biological systems.

Compared with static SANS, sample handling is easier in SEC–SANS as the final purification step is carried out online. The success rate of the SEC–SANS experiments presented here, in terms of the proportion of good data sets, was markedly higher than for standard solution SANS experiments performed previously by our group. The method has the

Table 2

Parameters describing the model refined from the data shown in Fig. 5.

We note that the empty nanodisc sample has not had the His₆ tag cleaved off and that the refined volume of MSP1D1 reflects this. Also note that the model for a membrane protein in a nanodisc is parametrized differently to the model presented in the previous section.

| Parameter | Empty nanodiscs | PR-loaded nanodiscs |
|--|-----------------|---------------------|
| Axis ratio of bilayer patch | 1.37 ± 0.11 | 1.59 ± 0.51 |
| Average area per lipid headgroup (Å ²) | 57.1 ± 2.1 | 60.0† |
| Hydrophobic height of bilayer (Å) | 33.0† | 33.8 ± 4.0 |
| No. of lipids per nanodisc | 117 ± 3.1 | 141 ± 11 |
| In-plane displacement of PR from center (Å) | — | 18.8 ± 9.6 |
| Volume of MSP1D1 (Å ³) | 28200 ± 420 | 19500 ± 6300 |
| Volume of POPC (Å ³) | 1270 ± 7.4 | 1315 ± 86 |
| Thickness of protein belt (Å) | 8.40† | 5.59† |
| Minor radius of bilayer patch (Å) | 33.0† | 31.1† |
| Major radius of bilayer patch (Å) | 44.3† | 49.4† |

† Parameters derived from the refined parameters.

additional advantage that both buffer and sample are passed through the same flow cell, making background subtraction more accurate than in standard cuvette-based static SANS experiments. Non-optimal background subtraction from SANS data might result in incorrect values for calculated sizes (Rubinson *et al.*, 2008).

SEC–SAS is especially advantageous for the study of aggregation-prone systems or systems in an equilibrium between structural or oligomeric states, whereas highly stable and monodisperse systems would preferably be measured with longer static measurements to obtain better counting statistics. The nanodisc is usually described as a very stable and monodisperse system, but its scattering pattern varied slightly but significantly as a function of retention volume as observed by SEC–SAS (Figs. 2*c* and 2*d*). This variation is ascribed to a small structural dispersity of the nanodiscs, with a variation in the lipid aggregation number as a function of retention volume, in line with other results from our group (Skar-Gislinge *et al.*, 2018) and consistent with recently reported findings for MSP1E3D1 nanodiscs (Graziano *et al.*, 2018). This structural effect would not have been visible without SEC–SAS.

The empty nanodiscs were found to have a slightly smaller D_{\max} in SANS compared with SAXS (Fig. 2*e*), which is in agreement with previous studies (Kynde *et al.*, 2014; Skar-Gislinge *et al.*, 2010) and is most likely to be a result of the nanodiscs having a higher peripheral excess scattering-length density in the SAXS contrast than in the SANS contrast. Interestingly, however, the D_{\max} values determined in this study from both SAXS and SANS are approximately 10–20 Å smaller than those reported in these studies. For a static SAXS or SANS measurement the nanodiscs are SEC-purified in advance, which may allow sufficient time for a monodisperse population of nanodiscs to equilibrate to a broader distribution of nanodiscs with different lipid:MSP ratios or even for the sample to start forming small populations of slightly larger particles. Our data indicate that in the SEC–SAS setups such issues are under better control as the measured samples of the individual frames are more monodisperse.

5.2. Feasibility for demanding complex biological samples

The SEC–SAS experiments were carried out with sample volumes in the range of a few hundred microlitres and concentrations of the reconstituted membrane protein of a few mg ml⁻¹ (Table 1), and yielded surprisingly good-quality SANS and SAXS data in terms of signal-to-noise ratio.

Firstly, PR-loaded nanodiscs were measured. A typical purification of PR yields several species, including monomers, trimers and hexamers (Gourdon *et al.*, 2008; Stone *et al.*, 2013), and two major species were indeed observed in the PR-loaded nanodisc samples (Fig. 3): a polydisperse population with characteristic sizes that were too large for a monomeric species and a population with the expected sizes for a monomeric PR-loaded nanodisc. The D_{\max} values for the monomeric PR-loaded nanodisc sample from SANS and SAXS (Fig. 3*e*), respectively, were close to those refined for the empty nanodisc (Fig. 2*g*), which is expected given that PR has almost all of its structure embedded in the lipid bilayer of the nanodisc. Despite the low concentration of the sample in the frames extracted from peak 2, the resulting data were of sufficient quality and allowed refinement of the model of PR embedded in a nanodisc (Fig. 5*c*). The D_{\max} was approximately 10 Å smaller than that reported for the similar membrane protein bacteriorhodopsin in MSP1D1-based nanodiscs as measured by standard static SANS and SAXS (Kynde *et al.*, 2014). This could indicate that a more monodisperse sample, and hence data of better quality, could be obtained using the combined SEC–SANS and SEC–SAXS setup.

The concentration of the TmCorA-loaded nanodisc sample was lower than that of the PR-loaded nanodisc, but the larger size of the system led to an increased scattering intensity and the obtained SANS and SAXS data (Fig. 4*d*) had a slightly better signal-to-noise ratio than the data from the PR-loaded nanodisc sample (Fig. 4). Our results might serve as a guideline for the required amount of sample for similar types of membrane proteins in nanodiscs, including the highly interesting class of G-protein-coupled receptors, which possess a similar overall structure in the membrane to that of PR.

5.3. Discussion of model fit results

A good model fit was obtained to the data from the empty nanodiscs (Fig. 5*a*). However, the model did not quite manage to reproduce the trend of the SAXS data around the minimum at $q \approx 0.08 \text{ \AA}^{-1}$, a property that is shared by several published fits on similar structures (Midtgaard *et al.*, 2015). This small deviation is most likely to be the result of the underlying structural dynamics and small structural dispersity of the nanodiscs (Fig. 2), as the model, for simplicity, assumes a static monodisperse structure. Still, the refined model of the empty nanodisc is in line with our expectations in terms of the refined values for the various parameters of the model. In particular, the partial specific molecular volume is refined to a value of 28200 \AA^3 , corresponding to a specific mass density of 1.46 g ml^{-1} , which is in good agreement with the reported value of 1.35 g ml^{-1} for proteins (Mylonas & Svergun, 2007). It is noteworthy, though, that the refined value of the average

area per POPC headgroup was found to be 57.1 \AA^2 . This is somewhat lower than the value of 60 \AA^2 reported for POPC in large bilayer vesicles (Kučerka *et al.*, 2011) and is also lower than the value of 66 \AA^2 determined for MSP1D1–POPC nanodiscs in a previous study (Skar-Gislinge *et al.*, 2010).

The fit of the data acquired from the PR-loaded nanodiscs managed to reproduce the trends of the data consistently (Fig. 5*b*). The POPC area per headgroup was derived as 60 \AA^2 from the refined model, which is in agreement with the value reported for POPC in unperturbed bilayers (Kučerka *et al.*, 2011). However, the value for the specific mass density of MSP1D1 refined from the data on PR-loaded nanodiscs is 1.88 g ml^{-1} and thus differs considerably from the expected value. Similarly, the refined value for the number of lipids in the nanodisc is slightly higher than anticipated. As a consequence, the MSP belt surrounding the loaded nanodisc adopts somewhat unphysical dimensions in the model. However, the molecular volume of MSP1D1 is not well determined, as seen by the large error bars on the value of this parameter. This is of some concern given the good quality of the SAS data. We suspect that it could be the result of a weakly determined concentration in the SEC–SAXS experiment, which is an issue that has been addressed elsewhere (Graewert *et al.*, 2015; Ryan *et al.*, 2018). An accurate estimate of sample concentration is paramount for fitting a complex many-parameter model.

For both the empty nanodiscs and the PR-loaded nanodiscs, the obtained SEC–SANS data showed no indications of aggregates at low q as judged by the flat Guinier regions. Furthermore, the data had a good signal-to-noise ratio up to $q \approx 0.15 \text{ \AA}^{-1}$ (Figs. 5*a* and 5*b*). As seen from the figure, the obtained SEC–SAXS data for empty nanodiscs had a very good signal-to-noise ratio up to the maximum q value, whereas data from PR-loaded nanodiscs had significant error bars from $q \approx 0.25 \text{ \AA}^{-1}$ owing to the lower sample concentration. In the present analysis we found that good (aggregate-free) SANS data up to $q \approx 0.1 \text{ \AA}^{-1}$ were sufficient to have a major impact on the model fit when they were combined with high-quality SAXS data, in particular regarding the parameters describing the lipid component. All in all, the complementary information obtained in the different contrast situations in SANS and SAXS is essential when refining complex models.

6. Conclusions and future perspectives

SEC–SANS was demonstrated to be feasible for studying the structures of empty nanodiscs as well as of membrane-protein-loaded nanodiscs. Data of surprisingly good quality and without signs of aggregation were obtained for all presented samples. The quality of the combined SEC–SANS and SEC–SAXS data obtained was sufficient to refine structural models of empty and PR-loaded nanodiscs.

Complex biological samples are demanding to produce in sufficient amounts and homogeneity for structural measurements, including SAXS and especially SANS. The high flux and a state-of-the-art detector at D22 at the ILL lowers the limit for the required amount of sample, and the new SEC–

SANS setup has made it possible to study truly monodisperse samples. In the presented sets of experiments, the flow rates were limited to 0.3 and 0.2 ml min⁻¹, respectively. With the continued development of the setup, stable flow rates down to 0.01 ml min⁻¹ are now available, which allow better tuning of exposure times, limited at one extent by the amount of sample and at the other by the stability of the sample, *i.e.* the resistance to aggregation over time. This opens the possibility of obtaining better statistics than for the data presented here, and/or of the loading of less material onto the column.

For the present article, the SEC–SANS experiments had to be repeated two or three times with different sample-to-detector distances to cover the desired *q*-range. The D22++ project within the Endurance phase II upgrade (proposal for funding submitted) includes a fixed detector placed at a short sample-to-detector distance, making it possible to cover a large *q*-range in a single SEC run. This upgrade will thus reduce both the time and the amount of sample needed for the experiment. Furthermore, a setup is being implemented with the possibility of running two SEC columns in parallel, enabling the overlap of sample elution from one column with the void volume from the other column, in turn making it possible to continuously measure relevant sample frames instead of wasting time and neutrons on excess buffer frames. Furthermore, this setup also has potential for automating measurements of contrast-variation series by programming the mixing of H₂O-based and D₂O-based buffers.

The anticipated high neutron flux at the upcoming European Spallation Source has potential for much improved SANS instrumentation and for either faster measurements and/or smaller sample volumes than at present sources. This, in combination with the general interest in complex biological structures, drives an increase in the level of ambition with respect to how complex biological samples may be investigated using SANS. As a consequence, dedicated BioSANS instruments or BioSANS instrument activities have become increasingly consolidated in several places in the world during the last decade. The high flux at D22 combined with online SEC is the optimal setup at present for studying complex and potentially unstable biological systems using SANS. Thus, for future studies of extremely precious samples, such as match-out deuterated nanodiscs (Maric *et al.*, 2014), this setup will be preferred.

Acknowledgements

The authors acknowledge the ILL and ESRF for developments in SEC–SAS and allocated beam time. We further thank the D22 technician Mark Jacques and the beamline scientists Marta Brennich and Petra Pernot at BM29 for their support during the beam times. Furthermore, we thank Søren Roi Midtgaard for help during the planning and execution of the first SEC–SANS experiment, as well as Frederik Grønbæk Tidemand for his help with MSP1E3D1 production and his assistance during the second SEC–SAXS beamtime. We thank Professor Steven Sligar for kindly providing the plasmids encoding MSP constructs, Professor Richard Neutze for the

plasmid encoding PR and Professor Mikaela Rapp, Heena Sandhu and Jens Berndtsson (University of Stockholm, Sweden) for the plasmid encoding TmCorA and for helpful advice on protein purification.

Funding information

Funding for this research was provided by the Synergy programme of the Novo Nordisk Foundation, the BRAIN-STRUC programme of the Lundbeck Foundation and BioSynergy KU2016.

References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213–221.
- Als-Nielsen, J. & McMorrow, D. (2011). *Elements of Modern X-ray Physics*, 2nd ed. Chichester: Wiley.
- Appolaire, A., Girard, E., Colombo, M., Durá, M. A., Moulin, M., Härtlein, M., Franzetti, B. & Gabel, F. (2014). *Acta Cryst.* **D70**, 2983–2993.
- Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. (2002). *Nano Lett.* **2**, 853–856.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Bibow, S., Polyhach, Y., Eichmann, C., Chi, C. N., Kowal, J., Albiez, S., McLeod, R. A., Stahlberg, H., Jeschke, G., Güntert, P. & Riek, R. (2017). *Nature Struct. Mol. Biol.* **24**, 187–193.
- David, G. & Pérez, J. (2009). *J. Appl. Cryst.* **42**, 892–900.
- Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. (2004). *J. Am. Chem. Soc.* **126**, 3477–3487.
- Denisov, I. G. & Sligar, S. G. (2016). *Nature Struct. Mol. Biol.* **23**, 481–486.
- Dewhurst, C. (2017). *GRASP v7.15*. <https://www.ill.eu/fr/users-en/scientific-groups/large-scale-structures/grasp/>.
- Fagerberg, L., Jonasson, K., von Heijne, G., Uhlén, M. & Berglund, L. (2010). *Proteomics*, **10**, 1141–1149.
- Förster, S., Apostol, L. & Bras, W. (2010). *J. Appl. Cryst.* **43**, 639–646.
- Gabel, F. (2015). *Methods Enzymol.* **558**, 391–415.
- Gabel, F., Lensink, M. F., Clantin, B., Jacob-Dubuisson, F., Villeret, V. & Ebel, C. (2014). *Biophys. J.* **107**, 185–196.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
- Glatter, O. (1977). *J. Appl. Cryst.* **10**, 415–421.
- Gourdon, P., Alfredsson, A., Pedersen, A., Malmerberg, E., Nyblom, M., Widell, M., Berndtsson, R., Pinhassi, J., Braiman, M., Hansson, Ö., Bonander, N., Karlsson, G. & Neutze, R. (2008). *Protein Expr. Purif.* **58**, 103–113.
- Graewert, M. A., Franke, D., Jeffries, C. M., Blanchet, C. E., Ruskule, D., Kuhle, K., Flieger, A., Schäfer, B., Tartsch, B., Meijers, R. & Svergun, D. I. (2015). *Sci. Rep.* **5**, 10734.
- Graewert, M. A. & Svergun, D. I. (2013). *Curr. Opin. Struct. Biol.* **23**, 748–754.
- Graziano, V., Miller, L. & Yang, L. (2018). *J. Appl. Cryst.* **51**, 157–166.
- Gripion, C., Legrand, L., Risenman, I., Vidal, O., Robert, M. C. & Boué, F. (1997). *J. Cryst. Growth*, **177**, 238–247.
- Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. (2013). *J. Am. Chem. Soc.* **135**, 1919–1925.
- Hansen, S. (2014). *J. Appl. Cryst.* **47**, 1469–1471.
- Hebling, C. M., Morgan, C. R., Stafford, D. W., Jorgenson, J. W., Rand, K. D. & Engen, J. R. (2010). *Anal. Chem.* **82**, 5415–5419.

- Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.
- Jeffries, C. M., Graewert, M. A., Blanchet, C. E., Langley, D. B., Whitten, A. E. & Svergun, D. I. (2016). *Nature Protoc.* **11**, 2122–2153.
- Jordan, A., Jacques, M., Merrick, C., Devos, J., Forsyth, V. T., Porcar, L. & Martel, A. (2016). *J. Appl. Cryst.* **49**, 2015–2020.
- Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, L., Doniach, S. & Plaxco, K. W. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12491–12496.
- Koutsoubas, A. (2017). *Biophys. J.* **113**, 2373–2382.
- Kučerka, N., Nieh, M.-P. & Katsaras, J. (2011). *Biochim. Biophys. Acta*, **1808**, 2761–2771.
- Kynde, S. A. R., Skar-Gislinge, N., Pedersen, M. C., Midtgaard, S. R., Simonsen, J. B., Schweins, R., Mortensen, K. & Arleth, L. (2014). *Acta Cryst. D* **70**, 371–383.
- Larsson, U. (1988). *Eur. J. Biochem.* **174**, 139–144.
- Levenberg, K. (1944). *Q. Appl. Math.* **2**, 164–168.
- Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. (2006). *Bioinformatics*, **22**, 623–625.
- Maric, S., Skar-Gislinge, N., Midtgaard, S., Thygesen, M. B., Schiller, J., Frielinghaus, H., Moulin, M., Haertlein, M., Forsyth, V. T., Pomorski, T. G. & Arleth, L. (2014). *Acta Cryst. D* **70**, 317–328.
- Marquardt, D. W. (1963). *J. Soc. Ind. Appl. Math.* **11**, 431–441.
- Mathew, E., Mirza, A. & Menhart, N. (2004). *J. Synchrotron Rad.* **11**, 314–318.
- Matthies, D., Dalmas, O., Borgnia, M. J., Dominik, P. K., Merk, A., Rao, P., Reddy, B. G., Islam, S., Bartsaghi, A., Perozo, E. & Subramaniam, S. (2016). *Cell*, **164**, 747–756.
- Mehmood, S., Domene, C., Forest, E. & Jault, J.-M. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 10832–10836.
- Midtgaard, S. R., Darwish, T. A., Pedersen, M. C., Huda, P., Larsen, A. H., Jensen, G. V., Kynde, S. A. R., Skar-Gislinge, N., Nielsen, A. J. Z., Olesen, C., Blaise, M., Dorosz, J. J., Thorsen, T. S., Venskutonytė, R., Krintel, C., Møller, J. V., Frielinghaus, H., Gilbert, E. P., Martel, A., Kastrop, J. S., Jensen, P. E., Nissen, P. & Arleth, L. (2018). *FEBS J.* **285**, 357–371.
- Midtgaard, S. R., Pedersen, M. C. & Arleth, L. (2015). *Biophys. J.* **109**, 308–318.
- Morgan, C. R., Hebling, C. M., Rand, K. D., Stafford, D. W., Jorgenson, J. W. & Engen, J. R. (2011). *Mol. Cell. Proteomics*, **10**, M111.010876.
- Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s245–s249.
- Niemann, H. H., Petoukhov, M. V., Härtlein, M., Moulin, M., Gherardi, E., Timmins, P., Heinz, D. W. & Svergun, D. I. (2008). *J. Mol. Biol.* **377**, 489–500.
- Nikolaev, M., Round, E., Gushchin, I., Polovinkin, V., Balandin, T., Kuzmichev, P., Shevchenko, V., Borshchevskiy, V., Kuklin, A., Round, A., Bernhard, F., Willbold, D., Büldt, G. & Gordeliy, V. (2017). *Cryst. Growth Des.* **17**, 945–948.
- Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N. & Dunker, A. K. (2005). *Biochemistry*, **44**, 1989–2000.
- Oldfield, C. J. & Dunker, A. K. (2014). *Annu. Rev. Biochem.* **83**, 553–584.
- Oliver, R. C., Pingali, S. V. & Urban, V. S. (2017). *J. Phys. Chem. Lett.* **8**, 5041–5046.
- Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. (2006). *Nature Rev. Drug Discov.* **5**, 993–996.
- Palombo, I., Daley, D. O. & Rapp, M. (2012). *J. Biol. Chem.* **287**, 27547–27555.
- Pedersen, J. S. (1997). *Adv. Colloid Interface Sci.* **70**, 171–210.
- Pedersen, J. S., Posselt, D. & Mortensen, K. (1990). *J. Appl. Cryst.* **23**, 321–333.
- Pedersen, M. C., Arleth, L. & Mortensen, K. (2013). *J. Appl. Cryst.* **46**, 1894–1898.
- Pedersen, M. C., Hansen, S. L., Markussen, B., Arleth, L. & Mortensen, K. (2014). *J. Appl. Cryst.* **47**, 2000–2010.
- Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., Morawe, C., Mueller-Dieckmann, C., Ohlsson, S., Schmid, W., Surr, J., Theveneau, P., Zerrad, L. & McSweeney, S. (2013). *J. Synchrotron Rad.* **20**, 660–664.
- Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s223–s228.
- Reckel, S., Gottstein, D., Stehle, J., Löhr, F., Verhoefen, M. K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Güntert, P. & Dötsch, V. (2011). *Angew. Chem. Int. Ed.* **50**, 11942–11946.
- Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M. & Sligar, S. G. (2009). *Methods Enzymol.* **464**, 211–231.
- Rubinson, K. A., Stanley, C. & Krueger, S. (2008). *J. Appl. Cryst.* **41**, 456–465.
- Ryan, T. M., Trehwella, J., Murphy, J. M., Keown, J. R., Casey, L., Pearce, F. G., Goldstone, D. C., Chen, K., Luo, Z., Kobe, B., McDevitt, C. A., Watkin, S. A., Hawley, A. M., Mudie, S. T., Samardzic-Boban, V. & Kirby, N. (2018). *J. Appl. Cryst.* **51**, 97–111.
- Schneidman-Duhovny, D., Hammel, M. & Sali, A. (2010). *Nucleic Acids Res.* **38**, W540–W544.
- Skar-Gislinge, N. & Arleth, L. (2011). *Phys. Chem. Chem. Phys.* **13**, 3161–3170.
- Skar-Gislinge, N., Johansen, N. T., Højberg-Nielsen, R. & Arleth, L. (2018). *Langmuir*, **34**, 12569–12582.
- Skar-Gislinge, N., Simonsen, J. B., Mortensen, K., Feidenhans'l, R., Sligar, S. G., Lindberg Møller, B., Bjørnholm, T. & Arleth, L. (2010). *J. Am. Chem. Soc.* **132**, 13713–13722.
- Skou, S., Gillilan, R. E. & Ando, N. (2014). *Nature Protoc.* **9**, 1727–1739.
- Stone, K. M., Voska, J., Kinnebrew, M., Pavlova, A., Junk, M. J. N. & Han, S. (2013). *Biophys. J.* **104**, 472–481.
- Svergun, D. I. & Koch, M. H. J. (2003). *Rep. Prog. Phys.* **66**, 1735–1782.
- Ward, A. B., Sali, A. & Wilson, I. A. (2013). *Science*, **339**, 913–915.
- Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., Trehwella, J. & Langley, D. B. (2007). *J. Mol. Biol.* **368**, 407–420.
- Zaccai, N. R., Sandlin, C. W., Hoopes, J. T., Curtis, J. E., Fleming, P. J., Fleming, K. G. & Krueger, S. (2016). *Methods Enzymol.* **566**, 159–210.