

Size-exclusion chromatography small-angle X-ray scattering of water soluble proteins on a laboratory instrument

Saskia Bucciarelli,^a Søren Roi Midtgaard,^b Martin Nors Pedersen,^b Søren Skou,^c Lise Arleth^b and Bente Vestergaard^{a*}

Received 11 May 2018

Accepted 13 October 2018

Edited by D. I. Svergun, European Molecular Biology Laboratory, Hamburg, Germany

Keywords: size-exclusion chromatography; small-angle X-ray scattering; SEC-SAXS; molecular weight; polydisperse proteins; in-house small-angle X-ray scattering.

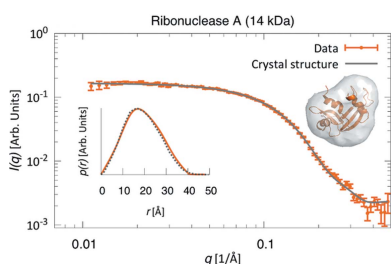
Supporting information: this article has supporting information at journals.iucr.org/j

^aDepartment of Drug Design and Pharmacology, University of Copenhagen, Denmark, ^bStructural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark, and ^cXenocs Nordic, Hørsholm, Denmark. *Correspondence e-mail: bente.vestergaard@sund.ku.dk

Coupling of size-exclusion chromatography with biological solution small-angle X-ray scattering (SEC-SAXS) on dedicated synchrotron beamlines enables structural analysis of challenging samples such as labile proteins and low-affinity complexes. For this reason, the approach has gained increased popularity during the past decade. Transportation of perishable samples to synchrotrons might, however, compromise the experiments, and the limited availability of synchrotron beamtime renders iterative sample optimization tedious and lengthy. Here, the successful setup of laboratory-based SEC-SAXS is described in a proof-of-concept study. It is demonstrated that sufficient quality data can be obtained on a laboratory instrument with small sample consumption, comparable to typical synchrotron SEC-SAXS demands. UV/vis measurements directly on the SAXS exposure cell ensure accurate concentration determination, crucial for direct molecular weight determination from the scattering data. The absence of radiation damage implies that the sample can be fractionated and subjected to complementary analysis available at the home institution after SEC-SAXS. Laboratory-based SEC-SAXS opens the field for analysis of biological samples at the home institution, thus increasing productivity of biostructural research. It may further ensure that synchrotron beamtime is used primarily for the most suitable and optimized samples.

1. Introduction

Solution-based small-angle X-ray scattering (SAXS) has over recent decades gained popularity in structural biology, owing to its potential to investigate the structure, dynamics and interactions of biomolecules directly in solution (Petoukhov & Svergun, 2013; Bizien *et al.*, 2016; Vestergaard, 2016). Such structural insights usually require monodisperse samples devoid of aggregates and impurities, as each individual component of a solution contributes to the total scattering pattern (Jacques & Trehella, 2010; Rambo & Tainer, 2013; Jeffries *et al.*, 2016; Rambo, 2017). This requirement can be challenging to meet for biological samples from a sample preparation point of view, not least because a vast number of biological processes, including pathological cases, involve structurally heterogeneous and aggregation-prone multi-domain proteins (Han *et al.*, 2007) and heterogeneous protein complexes in dynamic equilibria (Ali & Imperiali, 2005; Berger *et al.*, 2011; Knowles *et al.*, 2014; Marsh & Teichmann, 2015; Vestergaard, 2016). Structural investigation of such systems is highly relevant, but often obstructed by the structural dispersity and unstable or transient nature of the samples. Size-exclusion chromatography (SEC) coupled with



OPEN ACCESS

SAXS (SEC-SAXS) has emerged as a valuable tool to mitigate these difficulties (David & Pérez, 2009; Pérez & Nishino, 2012; Perez & Vachette, 2017) and has been successfully integrated at most synchrotron SAXS beamlines with a focus on biological SAXS (BioSAXS) (Mathew *et al.*, 2004; David & Pérez, 2009; Watanabe & Inoko, 2009; Gunn *et al.*, 2011; Graewert *et al.*, 2015; Brennich *et al.*, 2017; Ryan *et al.*, 2018). The technique has since resulted in prominent publications (*e.g.* Berthaud *et al.*, 2012; Meisburger *et al.*, 2016; Pflüger *et al.*, 2018) and is employed by an increasing number of synchrotron users. The availability of synchrotron beamtime is, though, limited, despite great efforts from beamline administrators to adapt to user needs, and is strictly confined to the scheduled time. Samples should thus be optimized as much as possible ahead of synchrotron experiments. However, the final optimization often needs to be done during the SEC-SAXS experiment and is an iterative process. With beamtime typically allocated only a few times per year, such optimization often becomes tedious and lengthy. The workflow of BioSAXS laboratories would therefore be significantly expedited by the possibility to optimize samples at the home institution. Additionally, the high X-ray flux of synchrotrons may induce radiation damage in the samples. While attenuating the beam resolves this issue, it also leads to longer measurement times and, correspondingly, consumption of precious beamtime.

In some cases, it can be difficult to obtain complete separation of the proteins eluting from the size-exclusion column (so-called baseline separation), leaving the eluting protein solutions polydisperse. In general, separation of the individual protein peaks improves with lower flow rates (Cheng & Hollis, 1987; Ricker & Sandoval, 1996; Hong *et al.*, 2012). This, however, again leads to increased spending of beamtime. Significant efforts are therefore being made to develop data processing tools to deal with overlapping peaks. In cases where there is sufficient separation of the eluting species, it is possible to identify data regions corresponding to monodisperse samples, which may be further analyzed (Malaby *et al.*, 2015; Panjkovich & Svergun, 2018). On the other hand, in cases where there is significant overlap of the eluting species, this approach will result in the unwanted effect of discarding a major portion of the collected data (corresponding to the data from a mixture of species); it is then necessary to decompose the data before further data evaluation (Brookes *et al.*, 2013, 2016), which often poses analytical challenges (Brookes *et al.*, 2016; Herranz-Trillo *et al.*, 2017). It is thus highly desirable to optimize sample conditions and flow rates to achieve better baseline separation in order to fully exploit the potential of SEC-SAXS.

Hence, it would be of great value to the field to be able to perform SEC-SAXS on an in-house instrument which is close to the home laboratory and where experiments are not affected by radiation damage or time limitations to the same extent as on synchrotron BioSAXS beamlines. Over the past decade, advances in the development of laboratory SAXS instruments optimized for solution measurements have enabled investigation of biological samples using static (not coupled to in-line chromatography) SAXS at the home

laboratory (Mortuza *et al.*, 2014; Sibillano *et al.*, 2014; Dupont *et al.*, 2015; Bruetzel *et al.*, 2016; Malmos *et al.*, 2016; Mortensen *et al.*, 2017). To our knowledge, however, there are to date no reports of SEC-SAXS having been implemented on any laboratory instruments. This is probably because of the comparatively low X-ray flux provided by classical generators, inducing the necessity for longer exposures on more concentrated samples, which is not compatible with SEC-SAXS. Although it has been suggested that laboratory-based SEC-SAXS should indeed be feasible with modern detectors (Wright *et al.*, 2013), it has recently been stated by Ryan *et al.* (2018) that ‘in-line SEC-SAXS [...] is generally beyond the capability of current laboratory SAXS instruments for most proteins’.

Nevertheless, having access to a state-of-the-art laboratory BioSAXS instrument (a Xenocs BioXolver L, equipped with a powerful MetalJet X-ray source and a single-photon-counting detector), we have devised a proof-of-concept study to assess the feasibility of laboratory-based SEC-SAXS. Using an array of proteins covering a wide range of molecular weights, we demonstrate that laboratory-based SEC-SAXS yields data of sufficient statistical quality within the time used to perform a standard SEC run, while consuming no more protein than is routinely used for synchrotron SEC-SAXS measurements (typically a few milligrams). We demonstrate that laboratory-based SEC-SAXS can produce publication quality data and highlight additional advantages related to the use of a laboratory setup.

2. Methods

2.1. Sample preparation

All proteins were from commercial sources. Ribonuclease A (RNase A), carbonic anhydrase (CAH), ovalbumin (OVA) and conalbumin (CA) were purchased from GE Healthcare. Bovine serum albumin (BSA) and horse apoferritin (HAF) were obtained from Sigma–Aldrich. Human insulin (HI) was obtained from Novo Nordisk A/S as a zinc-free powder and formulated according to the protocol described by Nygaard *et al.* (2012), except that the initial pH lowering described by Nygaard *et al.* was avoided. Hence, after the protein had been dissolved in water, Zn(OAc)₂, phenol, NaCl and sodium phosphate buffer were added and the concentration adjusted to reach a concentration of 600 µM HI in 7 mM sodium phosphate (pH 7.4), 60 mM phenol, 200 µM Zn(OAc)₂ and 23 mM NaCl. The pH of the sample was then checked and gently adjusted to pH 7.4 using small amounts of HCl or NaOH. Static SAXS measurements were performed on BSA in 50 mM HEPES buffer at pH 7.5 prepared from lyophilized protein without additional purification. Protein concentrations were determined using UV/vis spectroscopy at 280 nm on a NanoDrop 1000 using the protein extinction coefficients in Table S1 in the supporting information (SI). In the case of the static BSA measurement, the sample concentration was additionally determined using the in-line UV/vis capabilities of the SAXS instrument.

2.2. Size-exclusion chromatography

SEC was performed using an ÄKTA Purifier 100 high-performance liquid chromatography (HPLC) system from GE Healthcare coupled with a Superdex 200 Increase 10/300 GL column with a bed volume of 24 ml. In all instances, 0.5 ml of sample per measurement were loaded on the column. The flow rate was set to 0.5 ml min⁻¹ and decreased to 0.1 ml min⁻¹ when the protein eluted from the column, in order to ensure long enough exposure times and accordingly better counting statistics in the obtained data. The SEC runs were performed with the column kept at room temperature.

2.3. Small-angle X-ray scattering, laboratory

SAXS experiments were performed on a BioXolver L, a commercial laboratory instrument from Xenocs (<http://xenocs.com/en/solutions/bioxolver/>), equipped with a 250 W liquid gallium alloy X-ray source (MetalJet) with wavelength $\lambda = 1.34 \text{ \AA}$, and a BioCUBE, a temperature-controlled flow-through cell which allows UV/vis measurements directly on the SAXS exposure volume. The flux at the sample position was of the order of $3 \times 10^8 \text{ photons s}^{-1}$ with a beam size of approximately $1 \times 1 \text{ mm}$. SAXS data were collected at 298 K in 30 s exposures with variable sample–detector distance, d , altering the probed scattering angle θ and thus the range of the scattering vector \mathbf{q} [$|\mathbf{q}| = q = (4\pi/\lambda)\sin(\theta/2)$]. Two different medium-resolution settings were used for the experiments, depending on the desired q range: $d = 654 \text{ mm}$, $q = (0.011\text{--}0.50) \text{ \AA}^{-1}$, suitable for most proteins, and $d = 1507 \text{ mm}$, $q = (0.0075\text{--}0.22) \text{ \AA}^{-1}$, necessary for larger macromolecules. The sample concentration during the SEC-SAXS experiment was monitored using UV absorption at 280 nm, performed directly on the SAXS exposure volume. For comparison, a static SAXS measurement was performed on a 5 mg ml^{-1} BSA sample at $d = 654 \text{ mm}$, with a sample volume of 5 \mu l and an exposure time of 60 s.

2.4. Small-angle X-ray scattering, synchrotron

Static synchrotron SAXS data were collected on 4 mg ml^{-1} BSA at the EMBL beamline P12 at Petra III in Hamburg, Germany (Blanchet *et al.*, 2015), at 293 K, covering a q range from 0.0023 to 0.51 \AA^{-1} . The exposure time was 1 s, and the sample volume of 25 \mu l was flowed through the beam during exposure in order to reduce radiation damage. A synchrotron SEC-SAXS experiment was performed at 283 K at the BioSAXS beamline BM29 at ESRF, Grenoble (Pernot *et al.*, 2013), using the same column as for the laboratory experiments and an 8 mg ml^{-1} BSA sample (in PBS buffer), a loading volume of 500 \mu l , a beam attenuation of 45% and an exposure time of 1 s per frame. In order to reduce radiation damage, 1 mM dithiothreitol (DTT) was used, rather than glycerol. The advantage of DTT over glycerol is that it barely affects the scattering contrast (see SI for a detailed calculation), although in some cases it might affect the structure of the protein (Jeffries *et al.*, 2015).

2.5. Data analysis

The two-dimensional images were radially averaged using the software *RAW* (Nielsen *et al.*, 2009; Hopkins *et al.*, 2017). The resulting one-dimensional curves were integrated in the region between $q = 0.05$ and 0.1 \AA^{-1} and the integrated intensity was plotted as a function of time. The resulting curve was used to identify the frames corresponding to the eluting protein. The numbers of frames used for the data analysis are shown in Table S1 and selected frames in Figs. S1 and S2 in the SI. Background subtraction was performed by selecting and averaging frames corresponding to the background before and after the protein elution (see Figs. S1 and S2 in the SI). No further correction of the background subtraction was necessary. Measurements for absolute scale calibration were not performed for the SEC-SAXS measurements. For each individual frame over the monomer peak, the radius of gyration, R_g , and the scattering intensity in the forward direction, $I(0)$, were determined by the Guinier approximation through *AUTORG* (Petoukhov *et al.*, 2007). Theoretical scattering patterns were calculated from the crystal structures to the maximum experimentally measured q (listed in Table S1 in the SI) using *CRY SOL* (Svergun *et al.*, 1995) with background corrections enabled. The corresponding pair-distance distribution functions, $p(r)$, radii of gyration, $R_{g,\text{cryst}}$, and longest extensions, $D_{\text{max,cryst}}$ were extracted as described by Midtgaard *et al.* (2018). For the experimental data, pair-distance distribution functions, $p(r)$, longest extensions, $D_{\text{max,exp}}$, radii of gyration, $R_{g,\text{exp}}$, numbers of Shannon channels, N_s , and so-called numbers of good parameters, N_p , were determined using *BAYESapp* (Vestergaard & Hansen, 2006; Hansen, 2012). The experimental molecular weights, $M_{w,\text{exp}}$, were estimated from the Porod volume using *SAXSMoW* (Fischer *et al.*, 2010). For the static BSA sample measured on the BioXolver L, the molecular weight was also determined from scattering data on absolute scale in the software *RAW* after normalization of the recorded two-dimensional images by transmitted intensity of the direct beam, measured directly on the beamstop-free detector. A pure water sample was used as a secondary standard (Orthaber *et al.*, 2000). The useful q range was estimated using the Shannon-channel-based approach implemented in the program *Shanum* (Konarev & Svergun, 2015). For the generation of *ab initio* models, *DAMMIF* (Franke & Svergun, 2009) was used (see Table S1 for details), run ten times, aligned and filtered. The presented models are the end results from *damfilt* (Volkov & Svergun, 2003) of the ten aligned models, visualized by *PyMol* (<https://pymol.org/>) with their respective known crystal structures.

3. Results and discussion

3.1. The SEC-SAXS setup

The instrumental setup, schematically outlined in Fig. 1, consisted of the aforementioned HPLC unit (ÄKTA Purifier 100 and SEC column) mounted on a mobile table placed in close proximity to the SAXS instrument (BioXolver L) and connected to the flow-through cell (BioCUBE) which allows

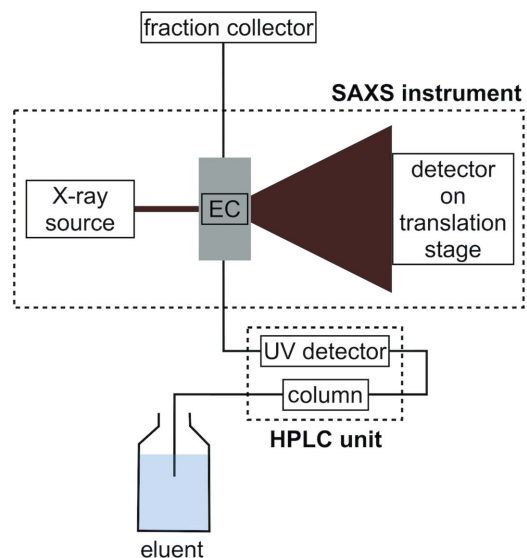


Figure 1
Schematic of our laboratory-based SEC-SAXS setup with the HPLC unit, composed of an ÄKTA chromatography system, a SEC column and a UV detector, the SAXS instrument (Xenocs BioXolver L) with an exposure cell (EC) for UV/vis and SAXS measurements, the X-ray source and the detector on a translation stage, and the fraction collector for sample collection after SEC-SAXS.

UV/vis measurements directly on the SAXS exposure volume (see Fig. S3 in the SI for photographs of the actual setup). The tubing volume from the column to the point of X-ray exposure was 415 μl . A fraction collector was placed immediately after the SAXS exposure cell to collect samples for further analysis. Synchrotron setups are typically placed in a safety interlock system which can only be accessed upon completion of the experiment, after closing the shutter in front of the X-ray beam. With a laboratory-based setup, however, access to the instrument during the course of the measurements offers the option to immediately collect the individual samples for further complementary analysis as they elute from the X-ray experiment. This is particularly important for labile samples. The X-ray exposure cell is temperature controlled, and additional temperature control on buffers and samples is feasible, but not implemented in this study.

3.2. Sample demands and data quality

A 60 s static SAXS measurement (without the SEC setup) on a 5 mg ml^{-1} BSA sample on the BioXolver L (Table 1) contains 3.9 good parameters, N_p . For reference, an equivalent static synchrotron data set is found to have 7.9 good parameters, corresponding to a clearly more information-rich data set. This is to be expected from the larger q range, due to the larger detector, and five orders of magnitude higher X-ray flux available at the beamline used. Nonetheless, the result from the laboratory instrument shows that data with good enough statistical quality can be obtained with short enough exposure times to enable SEC-SAXS.

To evaluate the sensitivity of the system, a series of measurements were performed using BSA, a 66 kDa protein

Table 1
Data quality and structural parameters.

Data quality and structural parameters from the experimental SEC-SAXS data of BSA. Also shown are the values obtained from static measurements on our laboratory instrument and a synchrotron BioSAXS beamline (P12, EMBL-Hamburg), as well as from a SEC-SAXS experiment on the beamline BM29, ESRF, Grenoble, for comparison. N_s : number of Shannon channels; N_p : number of good parameters; q_{\min} and q_{\max} : smallest and largest measured q ; q_{eff} : largest effectively useful q (from *Shanum*); $R_{g,\text{exp}}$: experimentally determined radius of gyration (from *BAYESapp*); $M_{w,\text{exp}}$: experimentally determined molecular weight (from *SAXSMoW*); c_{peak} : protein concentration at the maximum of the elution peak. Numbers in parentheses are uncertainties on the least significant digit.

| Sample | N_s | N_p | q_{\min} (\AA^{-1}) | q_{\max} (\AA^{-1}) | q_{eff} (\AA^{-1}) | $R_{g,\text{exp}}$ (\AA) | $M_{w,\text{exp}}$ (kDa) | c_{peak} (mg ml^{-1}) |
|--|-------|---------|-------------------------------------|-------------------------------------|---|--|-----------------------------|--|
| 7.6 mg ml^{-1} | 5.2 | 4.3 (2) | 0.011 | 0.50 | 0.50 | 27.5 (1) | 65 | 3.8 |
| 3.7 mg ml^{-1} | 5.3 | 3.4 (3) | 0.011 | 0.50 | 0.46 | 28.3 (1) | 67 | 1.8 |
| 2.0 mg ml^{-1} | 4.5 | 2.7 (3) | 0.011 | 0.50 | 0.46 | 27.1 (1) | 64 | 0.9 |
| 1.0 mg ml^{-1} | 4.8 | 2.4 (3) | 0.011 | 0.50 | 0.40 | 27.5 (2) | 61 | 0.4 |
| 7.6 mg ml^{-1} , low q | 4.8 | 2.4 (5) | 0.0075 | 0.22 | 0.19 | 27.8 (2) | 61 | 3.8 |
| 4.6 mg ml^{-1} , static | 6.0 | 3.9 (2) | 0.011 | 0.45 | 0.33 | 28.4 (1) | 66 | n/a |
| 4.3 mg ml^{-1} , static, P12 | 9.4 | 7.9 (2) | 0.0023 | 0.51 | 0.50 | 28.0 (1) | 66 | n/a |
| 8.1 mg ml^{-1} , BM29 | 11.1 | 8.6 (4) | 0.0054 | 0.48 | 0.48 | 27.1 (1) | 63 | n/a |

typically used as a calibration standard for SAXS experiments. The measurements were performed using the medium- q setting and covered a range of protein stock concentrations (8, 4, 2 and 1 mg ml^{-1}) with loaded volumes of 0.5 ml. The radii of gyration, R_g , and the forward scattering intensities, $I(0)$, of the individual frames across the monomer peak of the SEC-SAXS data sets (Fig. S4 in the SI), obtained from a Guinier analysis of each frame, demonstrate the monodispersity of the samples, allowing us to average multiple exposures to improve data quality for further analysis. The results are shown in Figs. 2(a)–2(d), with the corresponding Guinier plots shown in Figs. S5A–S5D in the SI. To evaluate the data quality from the instrumental low- q setting, resulting in a reduced X-ray flux, the experiment was repeated in this setting for the most concentrated sample (Figs. 2e, S4E and S5E). For comparison, we also show a SEC-SAXS data set of 8 mg ml^{-1} BSA obtained with a loading volume of 0.5 ml at the synchrotron BioSAXS beamline BM29, ESRF, Grenoble (Figs. 2f, S4F and S5F). In addition, we collected static SAXS data (*i.e.* without the SEC setup) from a sample at 5 mg ml^{-1} on our laboratory setup and 4 mg ml^{-1} at the beamline P12 (data shown in Fig. S6 in the SI and the corresponding Guinier plots in Figs. S5G–S5H). All data correspond well to the theoretically calculated curves based on the crystal structure of monomeric BSA (Fig. 2), evaluating parameters both in reciprocal [$I(q)$] and in direct space [$p(r)$]. The experimentally determined molecular weights $M_{w,\text{exp}}$ from *SAXSMoW* yield values within less than 10% of the theoretical value of 66 kDa (Table 1).

In an effort to quantify the data quality, we determined the number of Shannon channels, N_s , and good parameters, N_p , in the SEC-SAXS data and reference static SAXS measurements, as well as the usable q range. These values are reported in Table 1, together with the experimentally determined

structural parameters. As expected, the number of Shannon channels is roughly constant for data collected with the same detector setting (the variation between the values is caused by changes of the experimental value of $D_{\max,\text{exp}}$). The number of good parameters, however, is a better measure of the information content. Here, we observe a clear correlation between the noise level of the data and the number of good parameters available (Vestergaard & Hansen, 2006; Hansen, 2012; Pedersen *et al.*, 2014). In the data set from 8 mg ml⁻¹ BSA, 4.3 good parameters are found available in the data, gradually

decreasing to 2.4 for the 1 mg ml⁻¹ BSA data, clearly demonstrating the lower data quality associated with the lower protein concentration. Estimating the data quality by the largest meaningful q value from *Shanum*, it is also evident from Table 1 that the high- q region of the data is useful for all measurements, except for the lowest concentration. The SEC-SAXS data set from a synchrotron beamline from 8 mg ml⁻¹ BSA (Fig. 2*f*) contains 8.6 good parameters and a maximum useful q of 0.48 Å⁻¹, again associated with the higher X-ray flux of the synchrotron.

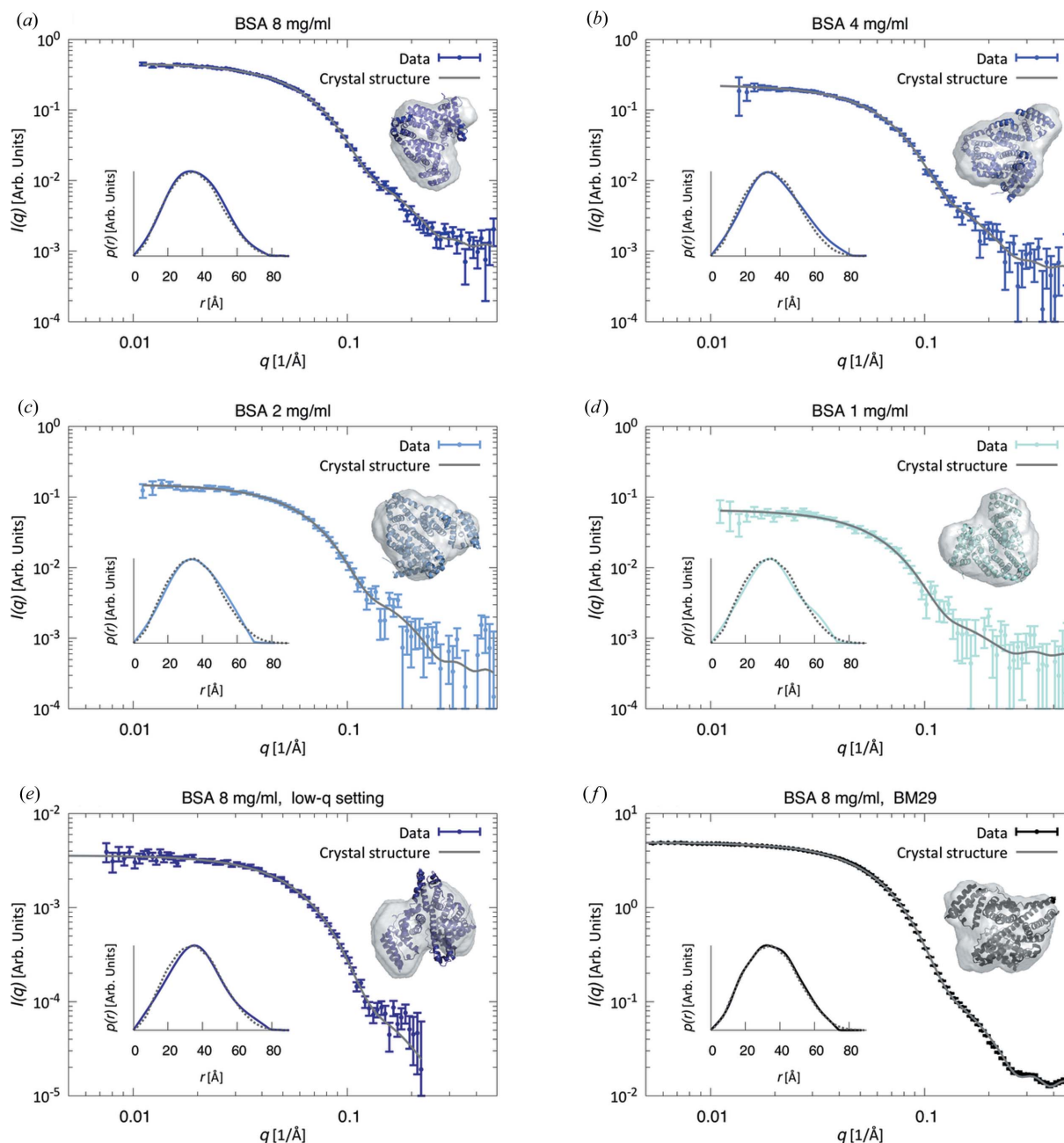


Figure 2

(a)–(e) BSA monomer data from SEC-SAXS (data points) at different concentrations and detector settings [(a)–(d) $d = 654$ mm, $q = (0.011\text{--}0.50)$ Å⁻¹; (e) $d = 1507$ mm, $q = (0.0075\text{--}0.22)$ Å⁻¹], together with the theoretical scattering curves calculated from the known crystal structure of monomeric BSA (gray lines). Insets show *ab initio* models based on the experimental data together with the crystal structure of BSA, viewed from different angles, as well as the corresponding pair-distance distribution functions, $p(r)$ [solid lines: experimental data; dotted lines: calculated $p(r)$]. (f) SEC-SAXS BSA monomer data set from a synchrotron BioSAXS beamline (BM29, ESRF, Grenoble).

Overall, the analysis presented here demonstrates that our setup produces data of good statistical quality over a range of protein concentrations and sample–detector distances. Ultimately, the minimum loaded amount of material guaranteeing sufficient data quality depends on the protein size, as well as the monodispersity of the sample, since higher polydispersity implies a lower total amount of protein in the main elution peak.

3.3. Molecular weight range

Having established the sensitivity of our setup, we demonstrate the general applicability of laboratory SEC-SAXS by

presenting data from six additional proteins (Figs. 3, S7 and S8 and Table 2), spanning a range of molecular weights from 14 to 476 kDa, and applying two different sample–detector distances for the largest protein (apoferritin) (see Table S1 in the SI for details).

The scattering curves of all proteins are consistent with those calculated from their respective crystal structures and the data contain between 2.7 and 6.9 good parameters, N_p , depending on the size of the protein (larger proteins have larger N_p) and the concentration of the sample. All measured radii of gyration (listed in Table 2) are in agreement with values obtained from the crystal structures, and the molecular

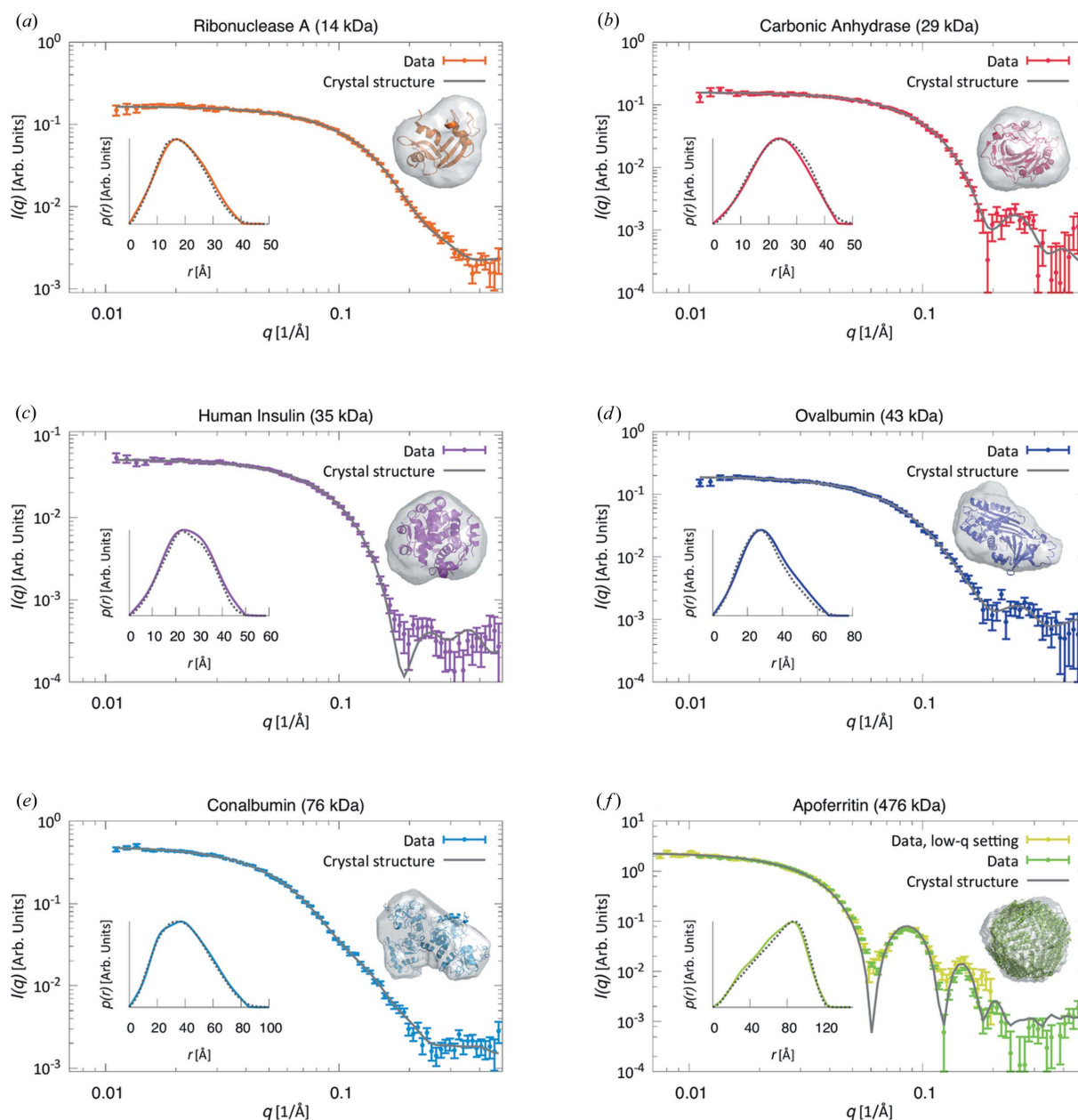


Figure 3 (a)–(e) SEC-SAXS data sets (data points) of various proteins, together with the *ab initio* models obtained from the experimental data and the known crystal structures for comparison. Also shown are the theoretical scattering curves (gray lines) calculated from the latter. Insets: pair-distance distribution functions $p(r)$ from the experimental data (solid lines) and from the crystal structures (dotted lines). (f) The low- q data were scaled to overlap with the data from the high- q setting. The *CRYSOLOG* fit, $p(r)$ and *ab initio* model were calculated on the basis of merged data from both settings.

Table 2

Structural and other parameters of various proteins.

Structural parameters of various proteins calculated on the basis of the scattering profiles from the SEC-SAXS experiments [$R_{g,exp}$ (from *BAYESapp*) and $M_{w,exp}$ (from *SAXSMoW*)] and from the known crystal structures ($R_{g,cryst}$), as well as the known molecular weight of the proteins ($M_{w,theo}$). Also shown are the number of good parameters, N_p , and the largest effectively useful q (from *Shanum*), q_{eff} , of the experimental scattering data, together with the smallest and largest measured q , q_{min} and q_{max} , and the concentration at the maximum of the elution peak, c_{peak} , determined using the UV/vis detection on the SAXS exposure cell. Note that for HI no peak concentration is available because of the presence of phenol, which strongly absorbs at 280 nm, in the running buffer. Numbers in parentheses are uncertainties on the least significant digit.

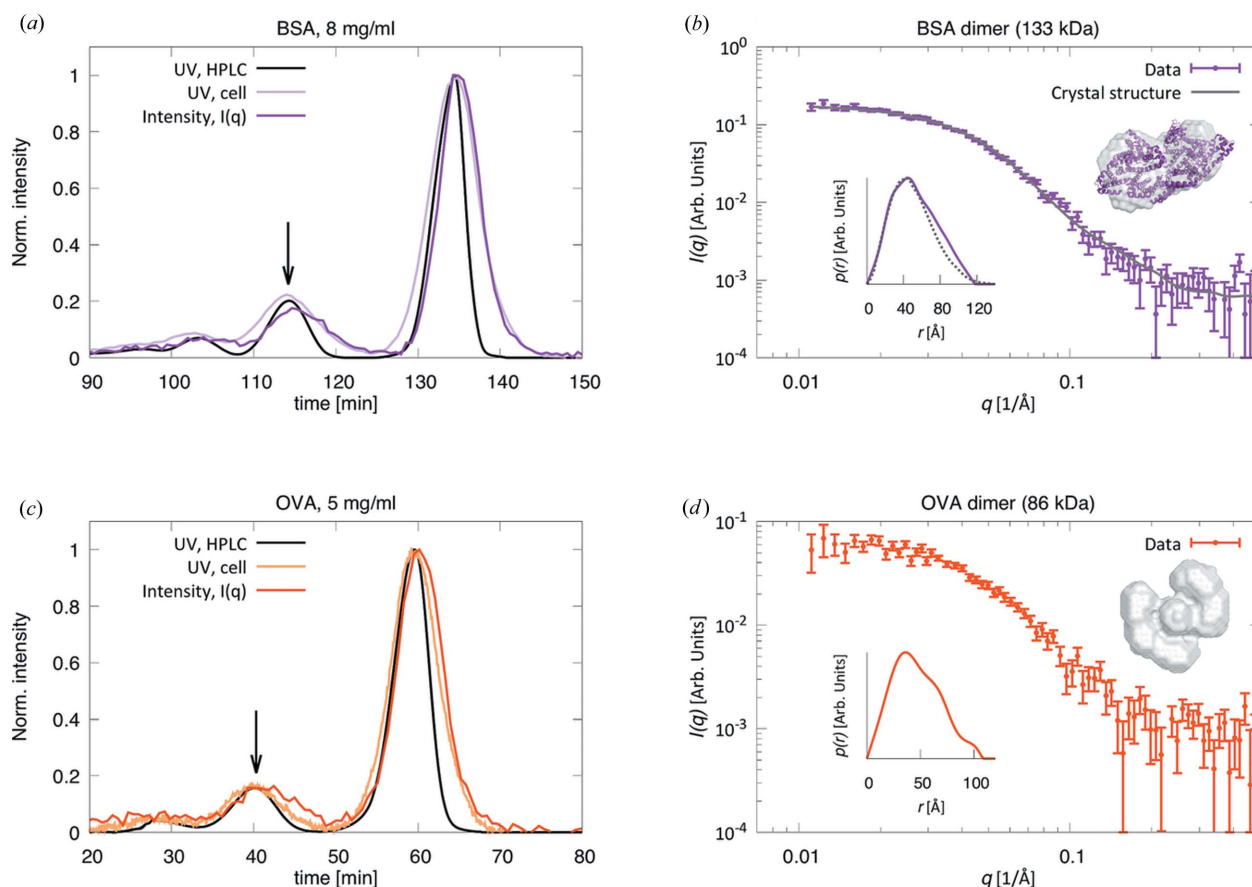
| Protein | $R_{g,exp}$ (Å) | $R_{g,cryst}$ (Å) | $M_{w,exp}$ (kDa) | $M_{w,theo}$ (kDa) | N_p | q_{min} (Å ⁻¹) | q_{max} (Å ⁻¹) | q_{eff} (Å ⁻¹) | c_{peak} (mg ml ⁻¹) |
|-----------------|--------------------|----------------------|----------------------|-----------------------|---------|---------------------------------|---------------------------------|---------------------------------|--------------------------------------|
| RNase A | 14.7 (1) | 14.4 | 10 | 14 | 3.7 (6) | 0.011 | 0.50 | 0.46 | 4.2 |
| CAH | 17.9 (1) | 18.3 | 31 | 29 | 2.7 (1) | 0.011 | 0.50 | 0.46 | 2.5 |
| HI | 19.1 (1) | 18.8 | 29 | 35 | 3.7 (1) | 0.011 | 0.50 | 0.49 | n/a |
| OVA | 23.7 (1) | 22.8 | 40 | 43 | 3.5 (5) | 0.011 | 0.50 | 0.43 | 2.0 |
| CA | 30.4 (1) | 30.3 | 76 | 76 | 5.0 (5) | 0.011 | 0.50 | 0.50 | 2.6 |
| HAF | 52.0 (1) | 53.0 | 434 | 476 | 6.9 (2) | 0.011 | 0.50 | 0.49 | 3.7 |
| HAF, low q | 52.3 (1) | 53.0 | 485 | 476 | 4.4 (5) | 0.0075 | 0.22 | 0.19 | 3.7 |
| HAF, merged | 52.2 (1) | 53.0 | 437 | 476 | 6.2 (2) | 0.0075 | 0.50 | 0.47 | 3.7 |

weights estimated from the experimental data using *SAXSMoW* are in good agreement with the known molecular weights (Tables 1 and 2). The size of the X-ray beam leads to a small instrumental smearing on laboratory sources. Such resolution effects, present in the experimental data, were not included in the *CRY SOL* fits presented here, leading to the sharper features of the theoretical scattering curves in Figs. 3(c) and 3(f) compared to the experimental data.

Fig. 3(f) shows good agreement between data obtained with different sample–detector distances and demonstrates the potential to cover a broader q range by merging data sets from different settings.

3.4. Direct UV/vis measurements on the SAXS exposure cell

Using low flow rates, we obtain baseline separation of monomer and dimer elution peaks in all cases (see Fig. 4 for two examples, BSA and OVA), except for HAF, the largest of the proteins investigated here. The concentrations of the protein dimers are significantly lower than those of the corresponding monomers, evident from the UV traces shown in Figs. 4(a) and 4(c).


Figure 4

Data for (a), (b) BSA and (c), (d) OVA. (a), (c) UV traces (at 280 nm) from the HPLC unit (black lines) and on the SAXS exposure cell (light purple and light orange), together with the integrated SAXS intensity $I(q)$ over time (dark purple and dark orange). Arrows indicate the dimer elution peaks. (b), (d) SEC-SAXS data sets (data points) corresponding to the protein dimers, marked by the arrows in panels (a) and (c), and theoretical scattering curve calculated from the known crystallographic dimer structure of BSA (gray line), together with the corresponding pair-distance distribution functions, $p(r)$, from the experimental data (solid lines) and from the crystal structure (dotted line). *Ab initio* models obtained from the experimental data and the crystal structure (for BSA) are shown as insets.

Table 3

Structural and data validation parameters of BSA and OVA dimers.

$R_{g,exp}$: radius of gyration (from *BAYESapp*); N_p : number of good parameters; q_{min} and q_{max} : smallest and largest measured q ; q_{eff} : largest effectively useful q (from *Shanum*). $M_{w,exp}$: molecular weight from *SAXSMoW*. For the OVA dimer, the software was not able to calculate a molecular weight from the experimental data. Numbers in parentheses are uncertainties on the least significant digit.

| Protein | $R_{g,exp}$ (Å) | $M_{w,exp}$ (kDa) | N_p | q_{min} (Å ⁻¹) | q_{max} (Å ⁻¹) | q_{eff} (Å ⁻¹) |
|-----------|--------------------|----------------------|---------|---------------------------------|---------------------------------|---------------------------------|
| BSA dimer | 40.8 (2) | 134 | 3.5 (4) | 0.011 | 0.50 | 0.43 |
| OVA dimer | 36.4 (2) | n/a | 2.8 (3) | 0.011 | 0.50 | 0.43 |

Broadening of the eluting protein peak always happens in a SEC-SAXS experiment owing to the large difference in diameter between standard HPLC tubing and the SAXS capillary and to Taylor dispersion in the tubing (Taylor, 1953). Correct estimation of this broadening is crucial for accurate concentration determination of the sample, necessary for direct molecular weight determination from the scattering data (Mylonas & Svergun, 2007). The ability of our setup to perform UV/vis measurements directly on the SAXS exposure cell, in conjunction with SAXS data acquisition, enables accurate determination of the chromatogram, and thus the protein concentration, at the point of exposure (Figs. 4a and 4c). In fact, the UV trace from the SAXS exposure cell presents a wider peak than that from the HPLC unit and overlaps with the integrated SAXS intensity. Establishment of the exact SEC-SAXS chromatogram allows a full correspondence between the fractionated samples after SAXS exposure and the corresponding measured SAXS data. Further characterization by additional biophysical or biochemical techniques can thus be linked directly to potential structural differences in the eluting proteins.

As measurements for absolute scale calibration were not performed during this study, we instead demonstrate the accuracy of the direct molecular weight determination from scattering data on absolute scale by means of a static BSA measurement performed on the same instrument (Fig. S6 in the SI). The in-line UV/vis and the NanoDrop spectrophotometer indeed yield the same value for the protein concentration (4.5 mg ml⁻¹), and the experimentally determined molecular weight from data on absolute scale (64 kDa) differs by less than 5% from the known molecular weight of BSA (66 kDa).

We demonstrate that the relatively high X-ray flux and low background-to-noise level of our laboratory setup yield data of sufficient statistical quality to even enable structural investigation of minor solution components, as illustrated by the BSA and OVA dimers (Table 3). While the peak broadening associated with increased flow rates might not be an impediment if SEC is merely employed to separate monomeric protein from higher oligomer species, more complex samples may require optimized peak resolution and thus low flow rates. Here, laboratory-based SEC-SAXS optimally complements synchrotron setups.

4. General discussion

Synchrotron SEC-SAXS is in many aspects superior to laboratory-based experiments, given the high flux, allowing for a higher signal-to-noise ratio, the more focused beam and hence minimal smearing effects, and the wider q range through larger detectors. However, laboratory-based SEC-SAXS has some intrinsic advantages that should be mentioned here. Synchrotron SEC-SAXS requires transportation of samples to the beamline, with inherent challenges related to the sensitivity of biological samples to temperature differences, mechanical stress and time. Particularly short-lived perishable samples might even necessitate preparation on site, thus requiring dedicated facilities at the synchrotron (Boivin *et al.*, 2016). In conjunction with the time pressure associated with synchrotron experiments, necessary extensive sample handling on site increases the risk of compromising the experiments. In the current study, we have used standard proteins for a proof-of-concept study, but it follows that vulnerable samples would greatly benefit from the ability to perform SEC-SAXS experiments on a laboratory-based instrument, immediately following careful and optimized preparation. In addition, easy access and available space around laboratory instruments, compared to synchrotron setups, in general allow for more complicated setups on the sample side and give additional flexibility when optimizing experiments. In particular, it is possible to collect the fractionated samples directly after the SAXS measurements and subject them to further biochemical or biophysical analysis. Given the accurate UV/vis assessment of the elution profile directly on the SAXS exposure cell, the results can be directly correlated with potential structural differences in each fraction, detected *via* the SAXS analysis.

Even for samples where the improved conditions on synchrotron beamlines are needed, laboratory-based SEC-SAXS evidently can serve as a valuable tool for initial investigations and optimization of samples prior to synchrotron SEC-SAXS.

Finally, small-angle neutron scattering (SANS) is gradually gaining in popularity in the biostructural research field (Chaudhuri, 2015; Gabel, 2015) and significant development is taking place, now also enabling implementation of in-line SEC (SEC-SANS) (Jordan *et al.*, 2016). SANS is complementary to SAXS, yielding similar information but with the added benefit of being able to distinguish between different components of a solution through contrast variation (Skar-Gislinge *et al.*, 2010; Hennig *et al.*, 2013; Kynde *et al.*, 2014). It is therefore frequently used in conjunction with SAXS at large-scale facilities. Bar a few exceptions, this typically requires multiple experiments at different facilities. This not only renders such experiments logistically challenging, but may also lead to data acquired on different samples, as many biological samples need to be freshly prepared immediately prior to the experiment. SANS beamlines starting to offer SEC-SANS to investigate complex and/or unstable biological samples would thus greatly benefit from having a laboratory-based SAXS instrument with an integrated SEC-SAXS setup directly on site. This would enable users to perform both experiments in

parallel, on the same sample, thus avoiding ambiguity related to sample variations.

5. Conclusion

We demonstrate that it is possible to perform SEC-SAXS on a laboratory-based instrument, provided it is optimized for high X-ray flux and signal-to-noise ratio. Recovery of the sample for further investigation using complementary techniques is possible immediately following SEC-SAXS, and thus users benefit from the presence of additional biophysics and biochemistry instruments at the home institution. UV/vis measurements on the SAXS exposure cell furthermore allow in-line concentration determination during the SEC-SAXS experiment and accurate correlation between SAXS data and the fractionated samples, obtained *via* correct estimation of the peak broadening. This enhances the opportunity for assignment of structure–function relationships. The laboratory-based setup presented here can be used not only to optimally prepare for synchrotron SEC-SAXS by optimizing samples at the home institution but also as an alternative to synchrotron experiments for a vast range of samples, as the obtained data are demonstrated here to be of useful statistical quality, leaving the precious synchrotron beamtime for samples requiring the higher flux or broader q range. Available measurement time not being as much a concern on a laboratory-based instrument as it is on synchrotrons, the former can also serve as a tool to investigate complex samples with multiple components which require slow flow rates for good baseline resolution.

Acknowledgements

We thank PhD student Lasse S. Dreyer for preparation of the human insulin samples. We gratefully acknowledge beamtime at the beamline P12, EMBL, Hamburg. The synchrotron SEC-SAXS experiment was performed on beamline BM29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Mark Tully at the ESRF for providing assistance in using beamline BM29, and to PhD students Nicolai T. Johansen, Frederik G. Tidemand, Noah Kassem and Marie Lycksell and to Dr Martin C. Pedersen for performance of this experiment.

Funding information

We acknowledge funding from the Novo Nordisk Foundation through the SYNERGY project (grant No. NNF15OC0016670), the Lundbeck foundation through the BRAINSTRUC project (grant No. 2015-2666) and funding for beamtime travels from the DANSCATT program.

References

Ali, M. & Imperiali, B. (2005). *Bioorg. Med. Chem.* **13**, 5013–5020.
 Berger, I., Blanco, A., Boelens, R., Cavarelli, J., Coll, M., Folkers, G., Nie, Y., Pogenberg, V., Schultz, P., Wilmanns, M., Moras, D. & Poterszman, A. (2011). *J. Struct. Biol.* **175**, 135–146.

Berthaud, A., Manzi, J., Pérez, J. & Mangenot, S. (2012). *J. Am. Chem. Soc.* **134**, 10080–10088.
 Bizien, T., Durand, D., Roblina, P., Thureau, A., Vachette, P. & Pérez, J. (2016). *Protein Pept. Lett.* **23**, 217–231.
 Blanchet, C. E., Spilotros, A., Schwemmer, F., Graewert, M. A., Kikhney, A., Jeffries, C. M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., Fiedler, S., Roessle, M. & Svergun, D. I. (2015). *J. Appl. Cryst.* **48**, 431–443.
 Boivin, S., Kozak, S., Rasmussen, G., Nemtanu, I., Vieira, V. & Meijers, R. (2016). *Methods*, **95**, 70–77.
 Brennich, M. E., Round, A. R. & Hutin, S. (2017). *J. Vis. Exp.* **119**, 54861.
 Brookes, E., Pérez, J., Cardinali, B., Profumo, A., Vachette, P. & Rocco, M. (2013). *J. Appl. Cryst.* **46**, 1823–1833.
 Brookes, E., Vachette, P., Rocco, M. & Pérez, J. (2016). *J. Appl. Cryst.* **49**, 1827–1841.
 Bruetzel, L. K., Fischer, S., Salditt, A., Sedlak, S. M., Nickel, B. & Lipfert, J. (2016). *Rev. Sci. Instrum.* **87**, 025103.
 Chaudhuri, B. (2015). *Protein Sci.* **24**, 267–276.
 Cheng, W. & Hollis, D. (1987). *J. Chromatogr. A*, **408**, 9–19.
 David, G. & Pérez, J. (2009). *J. Appl. Cryst.* **42**, 892–900.
 Dupont, D. M., Thuesen, C. K., Bøtkjaer, K. A., Behrens, M. A., Dam, K., Sørensen, H. P., Pedersen, J. S., Ploug, M., Jensen, J. K. & Andreasen, P. A. (2015). *PLoS One*, **10**, e0119207.
 Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. & Craievich, A. F. (2010). *J. Appl. Cryst.* **43**, 101–109.
 Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
 Gabel, F. (2015). *Structures of Large RNA Molecules and Their Complexes*, Methods in Enzymology, Vol. 558, pp. 391–415. Academic Press.
 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. (2015). *Sci. Rep.* **5**, 10734.
 Gunn, N. J., Gorman, M. A., Dobson, R. C. J., Parker, M. W. & Mulhern, T. D. (2011). *Acta Cryst.* **F67**, 336–339.
 Han, J., Batey, S., Nickson, A. A., Teichmann, S. A. & Clarke, J. (2007). *Nat. Rev. Mol. Cell Biol.* **8**, 319–330.
 Hansen, S. (2012). *J. Appl. Cryst.* **45**, 566–567.
 Hennig, J., Wang, I., Sonntag, M., Gabel, F. & Sattler, M. (2013). *J. Biomol. NMR*, **56**, 17–30.
 Herranz-Trillo, F., Groenning, M., van Maarschalkerweerd, A., Tauler, R., Vestergaard, B. & Bernado, P. (2017). *Structure*, **25**, 1–11.
 Hong, P., Koza, S. & Bouvier, E. S. (2012). *J. Liq. Chromatogr. Relat. Technol.* **35**, 2923–2950.
 Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.
 Jacques, D. A. & Trehwella, J. (2010). *Protein Sci.* **19**, 642–657.
 Jeffries, C. M., Graewert, M. A., Blanchet, C. E., Langley, D., Whitten, A. & Svergun, D. I. (2016). *Nat. Protoc.* **11**, 2122–2153.
 Jeffries, C. M., Graewert, M. A., Svergun, D. I. & Blanchet, C. E. (2015). *J. Synchrotron Rad.* **22**, 273–279.
 Jordan, A., Jacques, M., Merrick, C., Devos, J., Forsyth, V. T., Porcar, L. & Martel, A. (2016). *J. Appl. Cryst.* **49**, 2015–2020.
 Knowles, T. P., Vendruscolo, M. & Dobson, C. M. (2014). *Nat. Rev. Mol. Cell Biol.* **15**, 384–396.
 Konarev, P. V. & Svergun, D. I. (2015). *IUCrJ*, **2**, 352–360.
 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.* **D70**, 371–383.
 Malaby, A. W., Chakravarthy, S., Irving, T. C., Kathuria, S. V., Bilsel, O. & Lambright, D. G. (2015). *J. Appl. Cryst.* **48**, 1102–1113.
 Malmos, K. G., Bjerring, M., Jessen, C. M., Nielsen, E. H. T., Poulsen, E. T., Christiansen, G., Vosegaard, T., Skrydstrup, T., Enghild, J. J., Pedersen, J. S. & Otzen, D. E. (2016). *J. Biol. Chem.* **291**, 16849–16862.
 Marsh, J. A. & Teichmann, S. A. (2015). *Annu. Rev. Biochem.* **84**, 551–575.

- Mathew, E., Mirza, A. & Menhart, N. (2004). *J. Synchrotron Rad.* **11**, 314–318.
- Meisburger, S. P., Taylor, A. B., Khan, C. A., Zhang, S., Fitzpatrick, P. F. & Ando, N. (2016). *J. Am. Chem. Soc.* **138**, 6506–6516.
- Midtgaard, S. *et al.* (2018). *FEBS J.* **285**, 357–371.
- Mortensen, H. G., Madsen, J. K., Andersen, K. K., Vosegaard, T., Deen, G. R., Otzen, D. E. & Pedersen, J. S. (2017). *Biophys. J.* **113**, 2621–2633.
- Mortuza, G. B., Cavazza, T., Garcia-Mayoral, M. F., Hermida, D., Peset, I., Pedrero, J. G., Merino, N., Blanco, F. J., Lyngsø, J., Bruix, M., Pedersen, J. S., Vernos, I. & Montoya, G. (2014). *Nat. Commun.* **5**, 5072.
- Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s245–s249.
- Nielsen, S. S., Toft, K. N., Snakenborg, D., Jeppesen, M. G., Jacobsen, J. K., Vestergaard, B., Kutter, J. P. & Arleth, L. (2009). *J. Appl. Cryst.* **42**, 959–964.
- Nygaard, J., Munch, H. K., Thulstrup, P. W., Christensen, N. J., Hoeg-Jensen, T., Jensen, K. J. & Arleth, L. (2012). *Langmuir*, **28**, 12159–12170.
- Orthaber, D., Bergmann, A. & Glatter, O. (2000). *J. Appl. Cryst.* **33**, 218–225.
- Panjikovich, A. & Svergun, D. I. (2018). *Bioinformatics*, **34**, 1944–1946.
- Pedersen, M. C., Hansen, S. L., Markussen, B., Arleth, L. & Mortensen, K. (2014). *J. Appl. Cryst.* **47**, 2000–2010.
- Pérez, J. & Nishino, Y. (2012). *Curr. Opin. Struct. Biol.* **22**, 670–678.
- Perez, J. & Vachette, P. (2017). *Biological Small Angle Scattering: Techniques, Strategies and Tips*, Advances in Experimental Medicine and Biology, Vol. 1009, pp. 183–199. Singapore; Springer.
- Pernot, P. *et al.* (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.
- Petoukhov, M. & Svergun, D. (2013). *Int. J. Biochem. Cell Biol.* **45**, 429–437.
- Pflüger, T., Hernández, C. F., Lewe, P., Frank, F., Mertens, H., Svergun, D., Baumstark, M. W., Lunin, V. Y., Jetten, M. S. & Andrade, S. L. (2018). *Nat. Commun.* **9**, 164.
- Rambo, R. P. (2017). *Biological Small Angle Scattering: Techniques, Strategies and Tips*, Advances in Experimental Medicine and Biology, Vol. 1009, pp. 31–45. Singapore: Springer.
- Rambo, R. P. & Tainer, J. A. (2013). *Annu. Rev. Biophys.* **42**, 415–441.
- Ricker, R. & Sandoval, L. (1996). *J. Chromatogr. A*, **743**, 43–50.
- Ryan, T. M. *et al.* (2018). *J. Appl. Cryst.* **51**, 97–111.
- Sibillano, T., De Caro, L., Altamura, D., Siliqi, D., Ramella, M., Boccafosci, F., Ciasca, G., Campi, G., Tirinato, L., Di Fabrizio, E. & Giannini, C. (2014). *Sci. Rep.* **4**, 6985.
- 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.
- Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
- Taylor, G. (1953). *Proc. R. Soc. London Ser. A*, **219**, 186–203.
- Vestergaard, B. (2016). *Arch. Biochem. Biophys.* **602**, 69–79.
- Vestergaard, B. & Hansen, S. (2006). *J. Appl. Cryst.* **39**, 797–804.
- Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
- Watanabe, Y. & Inoko, Y. (2009). *J. Chromatogr. A*, **1216**, 7461–7465.
- Wright, G. S. A., Lee, H. C., Schulze-Briese, C., Grossmann, J. G., Strange, R. W. & Hasnain, S. S. (2013). *J. Synchrotron Rad.* **20**, 383–385.