Sample Preparation and Characterization

Cy Jeffries
EMBL Hamburg
Small-angle scattering (SAS)

What are the most robust parameters and related structural information that can be extracted from SAXS and SANS data from biomacromolecules in solution?

i) Radius of gyration ($R_g$) maximum particle dimension ($D_{max}$), volume ($V$).

ii) Molecular mass estimates ($MM$).

iii) Probable frequency of distances ($r$) within single particles ($p(r)$ vs $r$).

iv) Scaling parameters – compact, flexible, flat, rod, hollow.

v) Interparticle interactions: Attractive or repulsive.

vi) Size distributions and volume fractions.
Typically we want to develop models that describe the SAS data.

SAXS and SANS: Model interpretation starts with obtaining quality scattering data.

**SAXS models:** cardiac myosin binding protein C.

**SANS models:** Ribosome.

Jeffries et al. (2011) *J. Mol. Biol.* 414(5):735-748


...but you need to bootstrap information.

However…

What remains crucial to the interpretation of a scattering profile is the quality of the sample that is placed into an X-ray or neutron beam.

High Quality samples:

Good instrument: Quality data: Quality models:

…the model is as only as good as the data, which is as only as good as the sample.


The issue...

“Our protein looks just like BSA, that is exactly what we believed!”

In 2012 guidelines for publishing SAS data were, themselves, published. Several wwwPDB SAS taskforce meetings in the intervening years have developed standards for reporting SAS data, data formats, models, conventions, etc.
New Tables for reporting.

Four main themes.

1) **Sample details.**
2) SAS data acquisition and reduction.
3) Data presentation and validation.
4) Structural modelling – software tools and fit evaluations.

Table 1
Summary of guidelines for sample details.

- **Source of samples**, including sample-purification protocol, a measure of the final purity and how it was determined.
- Composition of the sample, including protein or nucleic acid sequences as measured, or FASTA IDs with the relevant ranges specified, plus fusion tags, ligands, cofactors, glycosylation or other modifications and the predicted molecular mass.
- Solvent/buffer pH and composition, including additives such as free-radical scavengers used to minimize the effects of radiation damage during SAXS data acquisition, and a statement of how the SAS-measured solvent blank was obtained (e.g., last-step dialyse, concentrator or column flowthrough).
- Sample concentration(s) and method(s) of determination, including extinction coefficients and wavelengths when UV absorbance measurements are used. In the case of combined SEC–SAXS experiments, a description (or reference) to the system, column size/type/resin, injection sample concentration and volume and flow rate.
- In the case of SANS contrast-variation experiments, the deuteration level of each biomolecular component (e.g., from mass spectrometry) and of the solvent (e.g., from densitometry or transmissions).
- Any SAS-independent assessments of monodispersity over a range of conditions (e.g., analytical ultracentrifugation, dynamic light scattering and/or aggregate-free gel filtration and/or multi-angle laser light scattering) that complement the SAS-based assessments.

*Plus data deposition into the Small Angle Scattering Biological Databank (SASBDB)*

www.sasbdb.org
Sample preparation is particularly important.

Preparation monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments

Cy M Jeffries¹, Melissa A Graewert¹, Clément E Blanchet¹, David B Langley²,³, Andrew E Whitten⁴ & Dmitri I Svergun¹

¹European Molecular Biology Laboratory (EMBL) Hamburg Outstation, DESY, Hamburg, Germany. ²Victor Chang Cardiac Research Institute, Darlinghurst, New South Wales, Australia. ³Garvan Institute, Darlinghurst, New South Wales, Australia. ⁴Australian Centre for Neutron Science, ANSTO, Lucas Heights, New South Wales, Australia. Correspondence should be addressed to C.M.J. (cy.jeffries@embl-hamburg.de).
Sample quality: SAXS and SANS:

Sample quality often ‘less forgiving’ compared to X-ray crystallography, NMR, EM, AFM, etc.

Why?

There is nowhere to hide…

For SAXS, every electron in a sample – or any electron between the sample and a detector – has the potential to scatter X-rays.

For SANS, every atomic nucleus in a sample – or any atomic nucleus between the sample and a detector – has the potential to scatter neutrons, either coherently or incoherently (depending on the isotope).
A reminder what X-rays and neutrons ‘feel’

X-rays are scattered by electrons.
Electrons = chemistry = potential X-ray damage

Neutrons are primarily scattered by atomic nuclei.

...basically empty space

...deep penetration of materials

...1.25 kilometers to nearest electron.

...unlike X-rays, low-to-no radiation damage!
The consequences.

1) Understand your sample.

...and

2) Understand the background scattering contributions!

**SAXS and SANS:**

- **Subtractive techniques.**
- Subtract all background scattering contributions to ‘reveal’ scattering contributions by the macromolecules in solution.

The relationship that guides sample preparation:

Is the **SUM** of the scattering from every particle $i$ within a sample... 

...and the form factor

Weighted by the difference in scattering length density of particles against the solvent and the volume of the particles **SQUARED**...

$I(q) = \sum_{i}^{n} \left[(\Delta \rho_i V_i)^2 P_i(q)\right] S(q)$

Scattering intensity...

So what?

Because small-angle scattering represents the **summed-weighted** contribution from all particles in a sample, the capacity to extract accurate shape information from a scattering data is reliant on:

1) **Sample purity**, i.e., a sample is sufficiently free of contaminants and is sufficiently dilute to limit between-particle interactions. If these conditions are met then:

   \[ I(q) = N(\Delta \rho V)^2 P(q) \]

   where \( N \) is the number density of homogeneous particles.

2) The accurate subtraction of **background scattering** contributions is critical, for example, from the supporting solvent, capillary, instrument, etc. This can be achieved by ensuring the experimental conditions used to measure the sample and the background are identical that includes the measurement of a exactly-matched solvent blank.
What can be controlled in the wet lab:

\[ \sum_{i}^{n} V_i^2 \]  \quad \text{Sample purity:}

\[ N \rightarrow \text{Sample concentration:} \quad \rightarrow C \]

\[ \Delta \rho^2 \]  \quad \text{Contrast}

\[ S(q) \]  \quad \text{Interparticle interactions: Change } N \text{ or the sample environment (e.g., change ionic strength).}
Polyacrylamide gel electrophoresis is a first step to assess the quality of a sample.

\[ \sum_{i}^{n} V_i^2 \]

**SDS-PAGE**

**Native-PAGE**

*Tip*: also run PAGE using reducing and non-reducing loading buffer!

Mokbel et al. (2013) *Brain*. 136; 494–507
Sample purity and contaminants.

**A. The ideal outcome when purifying a sample.** After background corrections have been made, the scattering from each individual protein within a population of pure monodisperse 14 kDa protein sum to produce a total scattering profile (red). Therefore, both the scattering data and the $P(r) \text{ vs } r$ represents the scattering of a single protein in solution.

**PAGE results 2: Sample with trace low MW contaminants.**

\[
\sum_{i}^{n} V_{i}^{2}
\]

Sample purity and contaminants.

**B. A less-ideal situation.** If low-MW contaminants are present, the total scattering (red) will be comprised of the sum of the scattering from each different species in proportion to their volume squared and concentration. Here, a low molecular weight (MW) contaminant (~ 5 kDa, 2% of the sample, blue) is present in the 14 kDa protein sample (grey). However, the total contribution to the scattering made by the low MW contaminant is small and does not significantly affect \( I(q) \) vs \( q \) or \( P(r) \) vs \( r \). (Approximate rule of thumb: \( 2 \times 5^2 \ll 98 \times 14^2 \).)

Sample purity and contaminants.

**C. Something to avoid.** High MW contaminants have disastrous consequences on $I(q)$ vs $q$ (red). The scattering contributions made by trace $\sim 100$ kDa protein (blue) doubles $I(0)$ even though the target 14 kDa protein (grey) is 98% pure. The effect on $P(r)$ vs $r$ is significant as it is the sum-weighted contribution made by the 14 kDa protein plus the 100 kDa contaminant. (Approximate rule of thumb: $2 \times 100^2 \sim 98 \times 14^2$).
Sample purity and contaminants.

**D. Flexibility.** A 100 kDa protein is both pure and monomeric. However, the protein is flexible. Therefore, the total $I(q)$ and subsequent $P(r)$ is the weighted-summed contribution of the volume fraction of each population.

A note on what SAXS people mean when they say: “Your sample is aggregated.”
Sample characterisation: SDS-PAGE, plus Size Exclusion Chromatography (SEC).

- The SDS-PAGE results might indicate that a protein sample is reasonably pure and consists of monomers.

- This result can be misleading if not backed up by further sample characterisation. The SEC trace indicates that the sample is comprised of a heterogeneous population of particles that include self-associated aggregates, dimers and monomers.
Sample characterisation: The advantage of DLS as a stand-alone tool.

- The disadvantage of SEC is that it can take time to screen solvent conditions that stabilise a component for SAXS investigations.
- Dynamic light scattering (DLS) acts both as an additional quality assurance tool to probe the polydispersity and hydrodynamic radius of a sample as well as a quick method to screen diverse sample environments.
Assess sample handling procedures with DLS

Protein solution: Different freeze-thaw protocols.

Incorporate shape-factor information. \( R_g(\text{SAXS})/R_h(\text{DLS}) \) ratio.

### A

20 mM NaCl

Low NaCl buffer: fast and slow thawing.

### B

170 mM NaCl

High NaCl buffer: fast and slow thawing. **Slow thaw aggregates!**

For \( \rho = R_g/R_h \)

<table>
<thead>
<tr>
<th>Shape</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>0.774</td>
</tr>
<tr>
<td>Coil</td>
<td>0.816</td>
</tr>
<tr>
<td>Rod</td>
<td>1.732</td>
</tr>
</tbody>
</table>

\( \text{pH} \), Ionic strength, Expensive additives
Do you have a quality sample? Molecular weight (MW) from $I(0)$.

The molecular mass and volume are extremely important parameters to experimentally obtain from the SAS data as they speak to sample quality.

$$I(0) \approx N (\Delta \rho V)^2$$

Where $N$ is the number-density, or a fancy way of saying concentration. From this concentration-dependent relationship the MW can be determined using two methods:

**Absolute scaling**

$$MW_{sample} = \frac{I(0)_{sample} N_A}{c_{sample} (\Delta \rho V_{sample})^2}$$

Absolute scaling - requires partial specific volume, contrast and concentration.

**Scaling relative to a known standard**

$$MW_{protein} = \frac{I(0)_{protein} \cdot c_{standard} \cdot MW_{standard}}{c_{protein} \cdot I(0)_{standard}}$$

An assumption that a target has a similar scattering length density and partial specific volume as the secondary standard.
If the molecular weight standard *does not* have the same scattering length density, partial specific volume, or is in a solution with a different contrast relative to the sample, you have to correct the relationship:

\[
\text{MW}_{\text{sample}} = \frac{I(0)_{\text{sample}}}{I(0)_{\text{standard}}} \times \frac{c_{\text{standard}}\Delta\rho_{\text{standard}}^2 v_{\text{standard}}^2}{c_{\text{sample}}\Delta\rho_{\text{sample}}^2 v_{\text{sample}}^2} \times \text{MW}_{\text{standard}}
\]

...need to take into account the contrast and partial specific volume

The issue with concentration-dependent MW methods.

1) You need to measure the concentration of the sample.

**Spectrophotometry**

Yes you can! – recommend 4 μl.
Protein Abs 280 nm:
Extinction coefficient from amino acid sequence
Protparam: web.expasy.org/protparam/

Polynucleotides Abs 260 nm:
Careful dilution series!

**Refractive index**

dRI – recommend 4 μl.

Anton Paar Abbemat 550
Refractive index.

- Temperature dependent.
- Wavelength dependent.

RI is an extremely useful tool for assessing protein concentration as the RI increment for proteins (~0.185 ml.g\(^{-1}\)) is — unlike A280 nm extinction coefficients — relatively stable. In other words, the refractive index increment of proteins is relatively independent from the amino acid or nucleotide sequence (but refer to the paper by Zhao et al., above).

The RI increment can also be adjusted for polynucleotides (DNA: ~0.17 ml.g\(^{-1}\) and RNA: 0.17–0.19 ml.g\(^{-1}\)).

Consequently, RI may be more useful for determining the concentration of, for example, proteins with low \(A_{280\,nm}\) extinction coefficients or protein/DNA complexes.

*Great for IDP sample concentration determination.*
Sample Concentration: How much material?

\[ C \]

Depends on the size of a macromolecule (due to \( V^2 \)), but generally \( 1-10 \text{ mg.ml}^{-1} \).

Larger proteins, lower concentration; smaller proteins, higher concentration.

For SAXS: 10-30 \( \mu \text{l} \) sample  
(Microfluidics = nl – \( \mu \text{l} \))

For SANS: 200-400 \( \mu \text{l} \) sample

*For SEC-SAS: Refer to up-and-coming lecture.*
Sample Concentration: concentration series.

Monitor $I(0)/C$ and $R_g$ through different concentrations.

Repulsive interparticle interference?
- Decrease sample concentration.
- Increase sample ionic strength (e.g., NaCl).

Alternative sample concentration methods

**Amino acid analysis:**
- Based on digesting the entire protein into component amino acids.
- Extremely accurate, but requires specialised equipment and personnel.

**Dye-based assays:**
- For example Bradford reagent.
- Generally less accurate unless standardised (case by case basis).
- Reducing agents can interfere with readings (e.g., DTT)
The issue with concentration-dependent methods II

The SAS commission recommends placing $l(s)$, i.e., the scattering intensities, on an absolute scale, cm$^{-1}$.

a) You need to know the contrast of the system, i.e., the difference in scattering length density between a macromolecule and the solvent:

$$\Delta \rho = \bar{\rho}_m - \bar{\rho}_s,$$

where $\bar{\rho}_m$ and $\bar{\rho}_s$ are the mean scattering length densities of the particle and the solvent, respectively.

For this to work properly, the sample has to be exactly-matched with the solvent in terms of mean scattering length density.

b) You need to know the partial specific volume, $psv$, of the macromolecule:

$$\nu = \text{cm}^3\cdot\text{g}^{-1}$$
How to calculate psv.

Experimentally – good luck, you will need about a gigaton of material.

OR

For amino-acid only proteins, use the ATSAS tool `seqstat`

a) Have the sequence ready as a .txt file, one letter amino acid code, no headers or footers.

b) Open a terminal, using the command line type:

   `>seqstat nameofsequence.txt`

The last column lists the psv based on sequence (tip: to get an idea what the other columns are type seqstat --h).

OR

Use the PSV and volume calculator from NucProt for proteins and RNA:  
`http://geometry.molmovdb.org/NucProt/`

OR

Use MULCh
MULCh: *Contrast* module

Modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies.

http://smb-research.smb.usyd.edu.au/NCVWeb/

*Contrast* calculates X-ray and neutron scattering contrasts the sample (and psv). All you need is the buffer composition and macromolecular sequence plus any bound cofactors/metal.

If a crystal structure is available, *CRYSOL* and *CRYSON* can be used (ATSAS).

Need help? Refer to Box 2/Figure11 of *Protocols*.

A note on background solvent 1:

\[ \Delta \rho^2 \]

Once again, SAS is a **subtractive** technique. The subtraction of solvent scattering contributions from the sample are necessary to obtain the scattering from a population of macromolecules.

Different buffers have different mean scattering length densities = scattering intensities. Therefore sample and solvent need to be matched.

![Different Buffers](image1.png)

![Matching Sample with Buffer](image2.png)
Is exact solvent matching really necessary?

- Incorrect background subtraction will affect structural parameters and modelling.

\[ \Delta \rho^2 \]

SAS can be used to identify and model structural flexibility.

40% of human proteins have regions of structural flexibility and 25% are predicted to be disordered from ‘beginning to end.’

A note on background solvent and Kratky plots.

\( \Delta \rho^2 \)

- Incorrect background subtraction will affect the interpretation of Kratky plots when attempting to identify molecular flexibility from a SAXS profile.

Incorrect background subtraction will affect the interpretation of Kratky plots when attempting to identify molecular flexibility from a SAXS profile.

Mismatched solvent subtraction (ca 7\% less scattering compared to correct solvent.)

Correctly matched solvent subtraction
Obtaining equivalent sample and matched solvent blanks.

**How?**
- Dialysis.
- Size exclusion chromatography.
- (Sometimes) spin concentrators.

$\Delta \rho^2$

**Do not:**
Use buffers/solvents that are ‘close enough’.

Replicating the scattering length density and absorption properties of a solvent is very difficult to reproduce.
Special considerations for SAXS samples.

- For SAXS, avoid really high concentrations of ‘electron-rich’ components in the supporting solvent, sucrose, glycerol, etc.
- For example, 5% v/v glycerol will reduce the background-corrected scattering intensities by approximately 20%.

- Another example: sucrose
Special considerations for SANS

Scattering lengths, biological elements.

\[ b_{(coherent)} \text{ values } (10^{-12} \text{ cm}) \]

-0.3741  \( ^1\text{H} \)

0.6671  \( ^2\text{H} \)

0.6651  \( ^{12}\text{C} \)

0.9370  \( ^{14}\text{N} \)

0.5803  \( ^{16}\text{O} \)

0.2804  \( ^{32}\text{S} \)

0.5130  \( ^{31}\text{P} \)

The \( b \) for \( ^1\text{H} \) is negative:
Attractive interaction potential.

\[ B_{(incoherent)} (10^{-12} \text{ cm}) \]

The spin state of \( ^{12}\text{C}, \, ^{16}\text{O} \) and \( ^{32}\text{S} \) disallow incoherent scattering events from the \( \frac{1}{2} \) spin state of neutrons.

As it happens, the incoherent scattering length of \( ^1\text{H} \) is enormous – one of the longest incoherent scattering lengths!

Lysozyme in 100% \( ^1\text{H}_2\text{O} \)

Lysozyme in 100% \( ^2\text{H}_2\text{O} \)

Incoherent neutron scattering in \( ^1\text{H}_2\text{O} \) is quite evident!
So now if we now take a solution.

We can sum the \textit{coherent} scattering lengths for each isotope of each atom per unit volume to obtain the average scattering length density, $\rho$, of the solution.

$$\sum \frac{b_i}{V}$$

Substituting $^1\text{H}$ with $^2\text{H}$, e.g., replacing regular light water ($^1\text{H}_2\text{O}$) with heavy water ($^2\text{H}_2\text{O}$, or D$_2$O): $\rho$ will change.
If we now take a macromolecule and put it into two different ρ solutions...

Low contrast = weaker coherent scattering intensities.

High contrast = stronger coherent scattering intensities.

We obtain two samples each with a different contrast.

\[ \Delta \rho = \rho_{\text{macromolecule}} - \rho_{\text{solvent}} \]

Contrast = the difference in the average scattering length density of the macromolecule and the average scattering length density of the solvent.

\[ I(q) = \sum_i^n [(\Delta \rho_i V_i)^2 P_i(q)] S(q) \]

\[ I(q) \propto \Delta \rho^2 \]
If we now take a complex with two *different* regions of scattering length density and put it into two different \( \rho \) solutions...

We can access the coherent scattering contributions from the *individual* components of the complex (while bound together), depending on the \(^1\text{H}/^2\text{H}\) ratios in the solvent (i.e., the \% v/v \(^2\text{H}_2\text{O}\)).
Different classes of macromolecules have different average scattering lengths.

The reason? Proteins, polynucleotides, and lipids naturally have different $^1$H per unit volume.

If hetero-macromolecular complexes, e.g., protein bound to DNA, are placed into the appropriate % v/v $^1$H$_2$O/$^2$H$_2$O solvents, it is possible to extract the scattering contributions for the whole complex (e.g., 100% v/v $^1$H$_2$O buffers) and from the individual components at the respective match points (for example 43% and 65% v/v $^2$H$_2$O).

Match point: $\Delta \rho = 0$

This type of experiment is called *contrast matching*.

Contrast variation.

Contrast variation means to collect SANS data from samples and buffers across several % v/v $^2$H$_2$O concentrations in the solvent.

0%, 20% 40%, 80% 90% 100%

A series of linear equations can be used to *extrapolate* the component scattering functions.

For a single component (One scattering length density)

$$I(q_{\text{total}}) \propto \Delta \rho^2$$

For a complex (Two scattering length densities)

$$I(q_{\text{total}}) \propto \Delta \rho_1^2 I(q)_1 + \Delta \rho_1 \Delta \rho_2 I(q)_{12} + \Delta \rho_2^2 I(q)_2$$

Component 1 contribution

‘between’ component contributions (cross-term)

Component 2 contribution
Protein-protein complexes require biodeuteration.

Almost all proteins have the same average neutron scattering length density because the $^1$H per unit volume from one protein to the next is roughly equivalent.

Therefore SANS will not provide any more information than SAXS regarding the shape of a protein-protein complex.

There is simple solution! Split the match points of the components by changing the scattering length density of one of the components!

How?

Biodeuteration – alter the $^1$H per unit volume of one of the proteins by substituting non-exchangeable $^1$H with non-exchangeable $^2$H.

The deuteration of recombinant proteins in *Escherichia coli* is predictable.

Basic minimal media with a selected ratio of $^2$H$_2$O.

Non-linear relationship. For example for 60% non-exchangeable $^2$H in the target protein, we require 80% v/v $^2$H$_2$O media.

Must use an *E. coli* B strain (e.g., BL21) – K12 strains (DH5α) do not grow.

Growth is VERY slow and requires cell adaption to the $^2$H$_2$O. This can take several days to a week.

**Box 3 | Deuteration of recombinant proteins using a laboratory-based $^2$H labeling protocol**

● TIMING 5–10 d + additional time if screening of bacterial growth conditions is necessary.


Aspect to consider: SANS sample preparation.

*Is your complex a complex in $^2\text{H}_2\text{O}$?*

The hydrogen bonding strength of $^2\text{H}_2\text{O}$ is stronger than that of $^1\text{H}_2\text{O}$.

This can significantly alter the solubility of macromolecules in solution, both the individual components and any resulting complex.

$$p\text{D} = p\text{H} \text{ (measured on pH meter)} + 0.4$$

$p^2\text{H}$ and $p\text{H}$ are also different!

For contrast variation studies you must $p\text{D}$ adjust the 100 % $^2\text{H}_2\text{O}$ buffer, taking into account the correction factor above, then $p\text{H}$ adjust the 100 % $^1\text{H}_2\text{O}$ buffer. Then mix the solutions in the desired % v/v $^2\text{H}_2\text{O}$. Once mixed you *cannot* $p\text{H}/p\text{D}$ adjust the solutions!
Aspect to consider: SANS sample preparation.

Compared to SAXS, you have to prepare more material. Neutron beams are typically 7–12 mm in diameter. The sample cells hold between 200–400 μl of sample. For a 5-point contrast series this means preparing 1–2 ml of sample (5–7 mg.ml\(^{-1}\))!

Exposure times are LONG.

The sample must be time-stable!

Neutron sources are orders of magnitude less intense than X-rays. Exposure times are typically between 15 min–1.5 hrs, depending on the detector position and sample concentration. To cover a \(q\)-range, usually more than one detector position is required. *Remember as a component is matched out, the scattering intensities will decrease!*

**For 5 samples and 5 buffers, MT cell and blocked beam = 6 – 20 hrs!**
**SANS: Incoherent scattering from $^1$H.**

- The introduction of contaminating $^1$H will change $\Delta \rho$ for a SANS experiment.
- The introduction of contaminating $^1$H will introduce additional incoherent scattering intensities.

**Atmospheric water:**
- Changes the contrast.
- Introduces $^1$H incoherent scattering.

---

**Keep water vapor out!**

Perform dialysis in sealed snap-lock bags.

Keep pipettes, tips, cassettes, etc dry.

Dedicate equipment to handle either $^1$H$_2$O or $^2$H$_2$O solutions (do not mix).
However: low percentage glycerol (less than 5% v/v), sucrose and other polyols as well as DTT, TCEP, ascorbic acid (1-5 mM), can reduce the effects of X-ray induced aggregation!
Radiation damage detected

**Instrument modifications**

- **Advantage**
  - No need to modify the sample.
- **Disadvantage**
  - Can result in a reduction in data quality.

  - **Decrease exposure time**
    - Desired outcome
      - Decreased signal-to-noise per data frame compensates by an increase in the number of frames unaffected by damage.

  - **Reduce flux: beam attenuation**

  **Sample modifications**

- **Advantage**
  - Maintain data quality.
- **Disadvantage**
  - Can alter the sample via chemical or physical effects (e.g., solubility/change in association state).
  - Very careful sample and solvent handling is necessary to prevent background mismatch.

  - **Increase sample flow**
    - Disadvantage
      - Potentially higher sample consumption.

  - **Add small molecules to sample and background solvent**
    - Alert
      - Requires highly accurate pipettes or a micro-balance for small volume additions of concentrated stock solutions, or
      - Requires additional dialysis time/SEC to obtain matched solvent blanks.

**Additives**

- 1–2 mM DTT
  - Alert
    - Proteins with oxidized disulphide bonds may undergo reduction.
    - Oxidized DTT has different 280 nm absorption properties

  **Tris or HEPES buffers + additive concentration screening.**
  - Alternative = TCEP, sucrose, ethylene glycol

- 1–2 mM Ascorbic acid
  - Alert
    - Adjusting the pH of ascorbate solutions before addition is advised to prevent pH shock.

- 3–5% (vol/vol) Glycerol
  - Alert
    - X-ray contrast is decreased.
    - Concentrated glycerol stock solutions require pH adjustment. Viscous stocks are difficult to add to small sample and solvent volumes.
Combining sample characterization with SAS?

- SEC-SAXS!

In-line SEC trace and RI/RALS MW validation.

MW: from RI/RALS and SAXS Porod volume

<table>
<thead>
<tr>
<th>Protein Mixture</th>
<th>Porod Volume (nm³)</th>
<th>MW - SAXS (kDa)</th>
<th>MW (Malvern) (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monomer</td>
<td>110</td>
<td>65</td>
<td>67</td>
</tr>
<tr>
<td>Dimer</td>
<td>207</td>
<td>122</td>
<td>134</td>
</tr>
</tbody>
</table>

Real-space distance distribution and ab initio modelling.

Up-and-coming lecture: Inline purification of equilibrium mixtures.
Acknowledgements

• Melissa Graewert

• Clement Blanchet

• Daniel Franke

• BioSAXS Group at EMBL-HH