Sample characterization:
Quality control and sample handling
prior to data collection

Marc JAMIN
UMI 3265 UJF-EMBL-CNRS Unit of Virus Host Cell interactions
Grenoble, France

jamin@embl.fr
Take home message

Don’t waste your time with a bad sample
The signal measured in a SAS experiments consists of the averaged contributions from all particles present in the sample.
Small-angle scattering experiments

- SAS data analysis can provide sample quality control
  1. Guinier plot $\rightarrow$ Rg and MM
  2. Concentration dependence
  3. Pair distance distribution function $\rightarrow$ $P(r) vs r$
  4. Kratky plot
  5. Porod-Debye plot
  6. …

... but it is too late
Small-angle scattering experiments

What should you worry about when preparing your sample?

- Purity
- Presence of aggregates
- Association state (monomer, dimer, …)
- Mixture of different oligomers
- For molecular assemblies, presence of excess of some components
Sample purity

- Electrophoresis
- Chromatography
- Mass spectrometry
- N-terminal sequencing
- ...

(c)

Excess of refractive index ($x \times 10^5$)

<table>
<thead>
<tr>
<th>Elution volume (mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>18</td>
</tr>
</tbody>
</table>

- Complex I 85 ± 6 kDa
- Par27 60 ± 5 kDa
- Fha30 32 ± 1 kDa
- Complex II 153 ± 11 kDa
Sample association state

Is my protein a monomer, a dimer, a tetramer ... ?

What is the stoichiometry of my multimolecular complex?

Is my sample monodisperse or polydisperse?
Sample association state

- Measuring molecular mass and molecular dimensions ($R_h$, $R_g$, $V$, ...)

---

*Image of Garfield*
Biophysical methods

A. Hydrodynamic properties
   - Size exclusion chromatography (SEC)
   - Analytical ultracentrifugation (AUC)
   - ...

B. Light scattering
   - Absorbance spectrum
   - Dynamic light scattering
   - Static light scattering
   - ...

C. Combination of techniques
   - SEC-MALLS
Size exclusion chromatography (SEC) is a technique used to separate mixtures based on the size of the particles. Small particles can enter the gel and have more volume to move down. They elute later.

Large particles cannot enter the gel and are excluded. They have less volume to move down and elute sooner.
Size exclusion chromatography - SEC

- The elution profile provides an first evaluation of sample quality
  - Presence or absence of material in the exclusion volume
  - Shape and width of the elution peak
SEC measures the hydrodynamic radius ($R_S$ or $R_h$) of a particle.
Size exclusion chromatography - SEC

Hypothesis:
- Particles have similar conformation than standards
- Particles have similar partial specific volumes
- Particles do not interact with the stationary phase

→ linear relationship between the hydrodynamic radius and the molecular mass
Size exclusion chromatography - SEC

The hydrodynamic radius depends on the shape of the particle
Size exclusion chromatography - SEC

Not all proteins are globular... some are cheating
Size exclusion chromatography - SEC

Rs = 5.2 ± 0.1 nm (VSV P)  MM ~ 240 kDa

MM (Monomer) = 30 kDa
MM (Dimer) = 60 kDa
The shape and width of the elution peak can reveal the presence of unresolved species.
Size exclusion chromatography - SEC

Advantages

• Simple and rapid

• Possibility of separating small oligomers (monomer, dimer, …)

• Purification method

Limitations and problems

• Limited resolution

• Sample dilution

• Loss of large aggregates within the column
Analytical ultracentrifugation - AUC

Thursday, 9 May 2013

Olwyn Byron - Structural information obtained from analytical ultracentrifugation
Analytical ultracentrifugation - AUC

Velocity sedimentation

Equilibrium sedimentation
# Analytical ultracentrifugation - AUC


<table>
<thead>
<tr>
<th>Radius (cm)</th>
<th>Absorbance at 280 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.05</td>
<td>0.0</td>
</tr>
<tr>
<td>6.10</td>
<td>0.4</td>
</tr>
<tr>
<td>6.15</td>
<td>0.8</td>
</tr>
<tr>
<td>6.20</td>
<td>1.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Radius (cm)</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.05</td>
<td>-0.08</td>
</tr>
<tr>
<td>6.10</td>
<td>-0.04</td>
</tr>
<tr>
<td>6.15</td>
<td>0.00</td>
</tr>
<tr>
<td>6.20</td>
<td>0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$S_{20}$ (S)</th>
<th>$c(S_{20})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$S_{20}$ (S)</th>
<th>$c(S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Advantages

- Thorough description of the sample

Limitations and problems

- Time consuming
- Complex analysis
Light scattering

- Light scattering has provided important methods for characterizing macromolecules and macromolecular assemblies.

\[ \mu_{\text{induced}} = \alpha E^* \]

- Light scattering results from the interactions between electrons and the electric field of light.

- The incident light induces an oscillating dipole in the electron cloud. As the dipole changes, energy is radiated or scattered in all directions.

\[ x e E^* \text{induced } = \alpha E^* \]
Light scattering

Light scattering varies with wavelength

\[
\frac{I_s}{I_0} = \frac{16 \pi^4 \alpha^2 \sin^2 \phi}{r^2 \lambda^4}
\]
Rayleigh scattering gives the atmosphere its blue color.
Monochromatic light $I_0, \nu_0$

Elastic Scattering

Scattered light $I_{s,\theta}, \nu_0$

Detector
Light scattering

The absorbance spectrum can already provide a crude evaluation of the sample.
Light scattering

- Light scattering requires fluctuations in polarizability or refractive index.
  - If the wavelength is long as compared to the separation of atoms: perfectly ordered materials show negligible light scattering.
  - Scattering occurs from materials containing fluctuations in the scattering power.

**Perfectly ordered material**

Diagram showing light scattering in a perfectly ordered material.

**Material with fluctuations in density**

Diagram showing light scattering in a material with fluctuations in density.
Imagine if a cuvette, containing particles which are stationary, is illuminated by a laser and a frosted glass screen is used to view the sample cell. A classical speckle pattern would be seen.

The speckle pattern will be stationary both in speckle size and position because the whole system is stationary. The dark spaces are where the phase additions of the scattered light are mutually destructive and cancel each other out. The bright blobs of light in the speckle pattern are where the light scattered from the particles arrives with the same phase and interfere constructively to form a bright patch.

For a system of particles undergoing **Brownian motion**, a speckle pattern is observed where the position of each speckle is seen to be in constant motion. This is because the phase addition from the moving particles is constantly evolving and forming new patterns.
Dynamic Light scattering - DLS

- DLS measures the time-dependence of light scattering.

- DLS measures Brownian motion and relates this to the size of the particles. Brownian motion refers to the random diffusive motion of particles suspended in a liquid or a gas due to the bombardment by the solvent molecules that surround them.

This motion is named after the botanist Robert BROWN who, in 1827, observed the motion of pollen grains in water through his microscope.

Robert Brown
(1773 – 1858)
The hydrodynamic radius of a particle can be calculated from the translational diffusion coefficient by using the Stokes-Einstein equation.

\[ R_h = \frac{k_B T}{6 \pi \eta D} \]

D = translational diffusion coefficient
\( k_B \) = Boltzmann constant
T = temperature (K)
\( \eta \) = viscosity

The velocity of the Brownian motion is defined by a property known as the **translational diffusion coefficient** (\( D \)).

In a one-dimension:
\[ \langle x^2 \rangle = 2Dt \]

In three-dimensions:
\[ \langle r^2 \rangle = 6Dt \]

The **hydrodynamic radius** of a particle can be calculated from the translational diffusion coefficient by using the Stokes-Einstein equation.
The rate at which these intensity fluctuations occur will depend on the size of the particles.

The small particles cause the intensity to fluctuate more rapidly than the large ones.
A correlator is basically a signal comparator. It is designed to measure the degree of similarity between two signals, or one signal with itself at varying time intervals.

If the intensity of a signal is compared with itself at a particular point in time and a time much later, then for a randomly fluctuating signal it is obvious that the intensities are not going to be related in any way, i.e. there will be no correlation between the two signals.

Knowledge of the initial signal intensity will not allow the signal intensity at time $t = \infty$ to be predicted. This will be true of any random process such as diffusion.
However, if the intensity of signal at time $t$ is compared to the intensity a very small time later $(t+\delta t)$, there will be a strong relationship or correlation between the intensities of two signals. The two signals are strongly or well correlated.

The correlation is reducing with time. The period of time $\delta t$ is usually very small, maybe nanoseconds or microseconds and is called the sample time of the correlator.

$t = \infty$ maybe of the order of a millisecond or tens of milliseconds.
Dynamic Light scattering - DLS

- If the signal intensity at \( t \) is compared with itself then there is perfect correlation as the signals are identical.

- Perfect correlation is indicated by unity (1.00) and no correlation is indicated by zero (0.00).

- If the signals at \( t+2\delta t \), \( t+3\delta t \), \( t+4\delta t \) etc. are compared with the signal at \( t \), the correlation of a signal arriving from a random source will decrease with time until at some time, effectively \( t = \infty \), there will be no correlation.

- If the particles are large the signal will be changing slowly and the correlation will persist for a long time. If the particles are small and moving rapidly then correlation will reduce more quickly.

**Autocorrelation function**

\[
F(\tau) = \langle I(t) I(t + \tau) \rangle
\]
Dynamic Light scattering - DLS

- The size distribution obtained is a plot of the relative intensity of light scattered by particles in various size classes and is therefore known as an intensity size distribution.
- The intensity distribution gives the relative contribution of each frequency to the total fluctuation trace.

![Monomodal Intensity Distribution](image)

![Multimodal Intensity Distribution](image)
Dynamic Light scattering - DLS

Advantages

- Simple and rapid
- Non-invasive method - sample recovery
- Possibility to analyze samples containing broad distributions of species of widely different diffusion coefficient
- Detect small amounts of higher mass species (< 0.01 %) – scattering intensity is proportional to the square of protein molecular mass.

Limitations and problems

- Less accurate than SLS and AUC for distinguishing small oligomers
- Complex analysis
The Rayleigh ratio is corrected for geometric factors.
Static Light scattering - SLS

- Rayleigh – Gans – Debye approximation

Zimm's equation:

\[
\frac{K^* C}{R_{0,C}} = \frac{1}{M_w P(\theta)} + 2 A_2 C
\]

- \( R_q \) = Rayleigh ratio at the scattering angle \( q \) and concentration \( C \)
- \( C \) = solute concentration
- \( M_w \) = weight average molar mass of the solute
- \( A_2 \) = second virial coefficient
- \( P(q) \) = form factor – angular dependence of the scattered light

- \( K^* \) = optical constant =

\[
4 \pi^2 n_0^2 \left( \frac{dn}{dC} \right)^2 \frac{1}{\lambda^4 N_A}
\]

- \( N_A \) = Avogadro’s number
- \( n_0 \) = refractive index of the solvent

B.H. Zimm (1948) «The scattering of light and the radial distribution function of high polymer solutions»
J. Chem. Phys. 16, 1093
Static Light scattering - SLS

For diluted solutions of molecules much smaller than the wavelength

- At low concentrations $C < 0.1 \text{ mg/ml}$

$$2 A^2 C M_w < 1$$

thus, the second virial coefficient term ($2 A^2 C$) is negligible

- For particles with $R_g$ value smaller than $1/10$th of the wavelength

$$P(\theta) = 1$$

thus, there is no angular dependence

The intensity of scattered light produced by a macromolecule is proportional to the product of the weight-average molar mass and the concentration of the macromolecule

$$\frac{R_{\theta,C}}{K^*} = M_w C$$
Static Light scattering - SLS

The form factor is important for molecules larger than a tenth of the radiation wavelength.

\[
\frac{KC_B}{R_\theta} = \frac{1}{P(\theta)} \left[ \frac{1}{M} + 2BC_B + \ldots \right]
\]

\[
P(\theta) = 1 - \frac{16\pi^2 R_G^2}{3\lambda^2} \sin^2\left(\frac{\theta}{2}\right) + \ldots
\]

\[
R_G = \left( \frac{\sum m_i r_i^2}{\sum m_i} \right)^{1/2}
\]

For particles larger than 70 nm, light scattering should also measure the radius of gyration.
Static Light scattering - SLS

Advantages
- Simple and rapid
- Sample recovery
- Absolute value of molecular mass

Limitations and problems
- Sensitive to small amounts of aggregates and dust particles
- Cannot be used with mixtures of components

Solution
Combining Size Exclusion Chromatography with on-line detection by static light scattering solves these problems.
- The column acts as a filter
- The column can separate different components present in the sample

Bonus
- Measurement of the dispersity of the sample
Static Light scattering - SLS

Size Exclusion Chromatography combined with static Multi-Angle Laser Light Scattering and refractometry – SEC-MALLS-RI
Scattered intensity is measured at 18 different angles

For each slice of the chromatogram, the intensity of scattered light is measured at different angles and concentration is measured by refractometry (on absorbance)
For each slice of the chromatogram, molar mass is calculated from Debye plot. The intercept on the graph gives the molar mass (MM).
Light scattering is responsible for light reflexion, refraction and diffraction.
Refractive index – differential refractometer

Refractometry – RI
RefRACTOMETRY – RI

\[ n = \frac{dn}{dc} \times l \times C \]
# Refractometry – RI

Refractive index increments of macromolecules

<table>
<thead>
<tr>
<th>Substance</th>
<th>dn/dc (mL/g)</th>
<th>λ (nm)</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins</td>
<td>0.185</td>
<td>690</td>
<td>20 mM Tris/HCl, 150 mM NaCl, pH 7.5</td>
</tr>
<tr>
<td>DNA</td>
<td>0.168</td>
<td>633</td>
<td>100 mM NaCl</td>
</tr>
<tr>
<td>RNA</td>
<td>0.170</td>
<td>690</td>
<td>25 mM Tris/HCl, 1 M NaCl, 1 mM TCEP, 5 mM MgCl2, pH 7.5</td>
</tr>
</tbody>
</table>
Static Light scattering - SLS

- Determination of the oligomeric state
- Monodisperse sample
- VSV P protein is a dimer in solution

\[ \text{MM} = 58.3 \pm 1.5 \text{ kDa} \]

\[ \text{MM (Monomer)} = 29991 \text{ Da} \]

265 a.a.
Phosphoprotein dimer + Nucleoprotein RNA complex

Static Light scattering - SLS

Nucleoprotein-RNA complex + Phosphoprotein dimer → Nucleoprotein-phosphoprotein complex

Stoichiometry 1N : 2P

104 ± 4 kDa

Static Light scattering - SLS

- Ligand-driven protein interaction.

- The plant hormone abscisic acid (ABA) has a central role in coordinating the adaptive response in situations of decreased water availability as well as the regulation of plant growth and development.

- Interaction between Pyr1, the ABA receptor, and the phosphatase Hab1: the molecular mass of 52 ± 4 kDa (red) indicates a 1:1 complex between PYR1 and HAB1, thus revealing that, in the presence of ABA, binding of HAB1 induces the dissociation of the PYR1 dimers.

Pyr1 : MM (Monomer) = 21.5 kDa
DNHAB1 : MM (Dimer) = 37.4 kDa

Santiago et al. (2009) Nature 462, 665-668
Polydisperse sample – sample is composed of discrete entities.

Dimer
130.6 ± 6.0 Da

Monomer
62.4 ± 0.4 Da

Static Light scattering - SLS

Bovine serum albumine
Static Light scattering - SLS

- Polydisperse sample – continuous distribution of molecular masses.
- Mouse prion protein

Number averaged molar mass

\[ M_n = \frac{\sum n_i M_i}{\sum n_i} = \frac{\sum c_i}{\sum c_i/M_i} \]

Weight averaged molar mass

\[ M_w = \frac{\sum n_i M_i^2}{\sum n_i M_i} = \frac{\sum c_i M_i}{\sum c_i} \]

Polydispersity factor

\[ \frac{M_w}{M_n} > 1.01 \]

Vendrely et al. (2005) BBA
Static Light scattering - SLS

- **Advantages**
  - Simple and rapid
  - Absolute value of molecular mass
  - Estimation of the polydispersity in a single chromatographic peak
  - The column act as a filter and remove dust particles and large aggregates

- **Limitations and problems**
  - Loss of large aggregates within the column
  - Sample dilution
Online SEC-SAXS

- Solvent
- HPLC pump
- Injector
- Column
- Réfractometer
- MALLS detector
- UV detector
- WATERS 2487
- X-ray detector
- X-ray beamline
- Waste
Finding optimal solution conditions

Thermofluor

- Thermofluor provides an optical reading of protein thermal melting.
- Thermofluor allows parallel measurements in 386 well-plates.
- Screening of solution conditions.
Finding optimal solution conditions

- High throughput dynamic light scattering using a plate-reader
Conclusions

- **Size exclusion chromatography** → hydrodynamic radius
  - association state, purification,…

- **Analytical ultracentrifugation**

- **DLS** → Translational diffusion coefficient
  - different association states, presence of large aggregates

- **SEC - MALLS – RI** → molecular mass
  - association state, polydispersity