Protein sizing by light scattering, molecular weight and polydispersity

Ulf Nobbmann  <Ulf.Nobbmann@malvern.co.uk>
Malvern Instruments
Malvern, Worcestershire WR14 1XZ
Outline

- Why light scattering?
  - Crystallization…
- Theory
  - Static light scattering (SLS) $\rightarrow$ molecular weight
  - Dynamic light scattering (DLS) $\rightarrow$ polydispersity
  - Electrophoretic light scattering (ELS) $\rightarrow$ zeta potential
- Application examples
  - Molecular weight
  - Sizing
  - Polydispersity
- Malvern Instruments & the Zetasizer Nano
Why Light Scattering?

- The scattering intensity is a function of the molecular weight and concentration.
- Non-invasive technique, giving information on the size, mass, and charge of a protein sample.
- Light scattering is extremely sensitive to the presence of small amounts of aggregates.
- The velocity of a particle under an applied electric field is proportional to the charge.
The direct link to Diamond

Crystallization!

- Need crystal for structure determination
- Search for optimum conditions for growth
- Light scattering can give indication of likelihood of success

Rule of thumb:
Low polydispersity (Pd ≤ 20%) for best chance of crystal growth
Light - Matter Interactions

As light is sent through material there are several potential interactions:

- Transmission
- Absorption
- Fluorescence

- Scattering!
Light - Matter Interactions: Scattering

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

The scattering signal may be analysed by several methods:

- Average signal strength: **static**, ‘classic’
- Fluctuations of signal: **dynamic**, quasi-elastic
- Shift of the signal: **electrophoretic**
Static Light Scattering

Molecular Weight Measurements
Static Light Scattering (SLS)

Average scattering intensity is a function of the (particle) molecular weight and the 2\textsuperscript{nd} virial coefficient.

\[ \frac{KC}{R_\theta} = \left( \frac{1}{M} + 2A_2C \right) \frac{1}{P(\theta)} \]

- \( K \): Optical constant
- \( C \): Concentration
- \( M \): Molecular weight
- \( R_\theta \): Rayleigh ratio
- \( A_2 \): 2\textsuperscript{nd} Virial coefficient
- \( P(\theta) \): Shape (or form) factor
Static Light Scattering (SLS)

\[
\frac{KC}{R_\theta} = \left( \frac{1}{M} + 2A_2C \right) \frac{1}{P_\theta}
\]

\[
K = \frac{2\pi^2}{\lambda_o^4 N_A} \left( n_o \frac{dn}{dc} \right)^2
\]

\[
\lambda_o = \text{laser wavelength}
\]

\[
N_A = \text{Avogadro's number}
\]

\[
n_o = \text{Solvent RI}
\]

\[
dn/dc = \text{differential RI increment}
\]

\[
P_\theta = 1 + \frac{16\pi^2 n_o^2 R_g^2}{3\lambda_o^2} \sin^2 \left( \frac{\theta}{2} \right)
\]

\[
R_g = \text{Radius of gyration}
\]

\[
\theta = \text{Measurement angle}
\]

\[
R_\theta = \frac{I_A n_o^2}{I_T n_T^2} R_T
\]

\[
I_A = \text{Intensity of analyte (sample I – solvent I)}
\]

\[
n_o = \text{Solvent RI}
\]

\[
I_T = \text{Intensity of standard (toluene)}
\]

\[
n_T = \text{Standard (toluene) RI}
\]

\[
R_T = \text{Rayleigh ratio of standard (toluene)}
\]
Static light scattering (SLS)

- The intensity of scattered light that a macromolecule produces is proportional to the product of the weight-average molecular weight and the concentration of the macromolecule \( I \propto (M_W)(C) \)

- For molecules which show no angular dependence in their scattering intensity, accurate molecular weight determinations can be made at a single angle (Rayleigh scatterers, isotropic scattering)

- This is called a Debye plot and allows for the determination of
  - Absolute Molecular Weight
  - 2nd Virial Coefficient \( (A_2) \)
Debye plots: What do the measurements involve?

- Preparation of a number of concentrations of the unknown molecule (protein) in a suitable buffer

- Typical concentrations: 1, 2, 3 and 5 mg/mL
Static Light Scattering (SLS)

\[
\frac{KC}{R_\theta} = \left( \frac{1}{M} + 2A_2C \right) \frac{1}{P_\theta}
\]

For Rayleigh scatterers, \( P(\theta) = 1 \) and the equation is simplified to

\[
\frac{KC}{R_\theta} = \left( \frac{1}{M} + 2A_2C \right) \quad (y = b + mx)
\]

Therefore a plot of \( KC/R_\theta \) versus \( C \) should give a straight line whose intercept at zero concentration will be \( 1/M \) and whose gradient will be \( A_2 \)
Molecular Weight Example (Lysozyme in PBS)

\[ \frac{dn}{dc} = 0.185 \text{ (mL/g)} \]

- \( I_{\text{tol}} = 192630 \text{ (counts/sec)} \)
- \( I_{\text{sol}} = 21870 \text{ (counts/sec)} \)

<table>
<thead>
<tr>
<th>Lysozyme Concentration (mg/mL)</th>
<th>Measured Intensity (counts/sec)</th>
<th>Intensity of Analyte (counts/sec)</th>
<th>KC/R(_0) (1/Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.006</td>
<td>87,830</td>
<td>65,960</td>
<td>6.1994 x 10(^{-5})</td>
</tr>
<tr>
<td>3.018</td>
<td>222,900</td>
<td>201,030</td>
<td>6.4765 x 10(^{-5})</td>
</tr>
<tr>
<td>5.029</td>
<td>366,770</td>
<td>344,900</td>
<td>6.6682 x 10(^{-5})</td>
</tr>
<tr>
<td>10.059</td>
<td>742,570</td>
<td>720,700</td>
<td>6.7743 x 10(^{-5})</td>
</tr>
</tbody>
</table>
Molecular Weight Example
(Lysozyme in PBS)

\[ \frac{1}{1/\text{Intercept}} = 14.6 \text{KDa} \]

\[ \text{Slope} = -3.23 \times 10^{-4} \]
2nd virial coefficient

- A thermodynamic property describing the interaction strength between the molecule and the solvent
- For samples where $A_2 > 0$, the molecules tend to stay in solution (protein molecules prefer contact with buffer)
- When $A_2 = 0$, the molecule-solvent interaction strength is equivalent to the molecule-molecule interaction strength – the solvent is described as being a theta solvent (protein doesn’t mind buffer)
- When $A_2 < 0$, the molecule will tend to fall out of solution or aggregate (protein doesn’t like buffer)
Dynamic Light Scattering

Molecular Size Measurements
Dynamic Light Scattering

Dynamic light scattering is a technique for measuring the size of molecules and nanoparticles.

DLS measures the time dependent fluctuations in the scattering intensity to determine the translational diffusion coefficient ($D_T$), and subsequently the hydrodynamic radius ($R_H$).
Dynamic Light Scattering

- Dynamic light scattering is a technique for measuring the size of molecules and nanoparticles.
- DLS measures the time dependent fluctuations in the scattering intensity to determine the translational diffusion coefficient (D), and subsequently the hydrodynamic size.
- The rate of intensity fluctuation is dependent upon the size of the particle/molecule.
Dynamic Light Scattering (DLS)

Fluctuations are a result of Brownian motion and can be correlated with the particle diffusion coefficient and size.

\[ g(\tau) = \frac{\langle I(t)I(t+\tau) \rangle}{\langle I(t) \rangle^2} = A + \sum B_\ell e^{-2q^2D\tau} \]

**Stokes-Einstein**

\[ R_H = \frac{kT}{6\pi\eta D} \]

**Symbols**
- \( q = \) Scattering vector
- \( R_H = \) Radius
- \( D = \) Diffusion coefficient
- \( k = \) Boltzmann constant
- \( T = \) Temperature
- \( \eta = \) Solvent viscosity
Brownian Motion

- Random movement of particles due to the bombardment by the solvent molecules that surround them
Brownian Motion

Temperature must be
- accurately known (viscosity)
- stable (otherwise convection present)

The larger the particle the more slowly the Brownian motion will be

The higher the temperature the more rapid the Brownian motion will be

‘Velocity’ of the Brownian motion is defined by the translational diffusion coefficient ($D_T$)
Physical Constraints

The non-randomness of the intensity trace is a consequence of the physical confinement of the particles to be in locations very near to their initial locations across very short time intervals.
Dynamic Light Scattering (DLS)

Fluctuations are a result of Brownian motion and can be correlated with the particle diffusion coefficient and size.

\[ g(\tau) = \frac{\langle I(t)I(t+\tau) \rangle}{\langle I(t) \rangle^2} = A + \sum B e^{[-2q^2D\tau]} \]

**Stokes-Einstein**

\[ R_H = \frac{kT}{6\pi\eta D} \]

$q =$ Scattering vector  
$D =$ Diffusion coefficient  
$R_H =$ Radius  
$k =$ Boltzmann constant  
$T =$ Temperature  
$\eta =$ Solvent viscosity
Stokes-Einstein Equation

\[ R_H = \frac{k_B T}{6 \pi \eta D_T} \]

where
- \( R_H \) = hydrodynamic diameter
- \( k_B \) = Boltzmann’s constant
- \( T \) = absolute temperature
- \( \eta \) = viscosity
- \( D_T \) = diffusion coefficient
Comparative Protein $R_H$ Values

Lysozyme  
$M_W=14.5$ kDa  
$R_H=1.9$ nm

Insulin - pH 7  
$M_W=34.2$ kDa  
$R_H=2.7$ nm

Immunoglobulin G  
$M_W=160$ kDa  
$R_H=7.1$ nm

Thyroglobulin  
$M_W=650$ kDa  
$R_H=10.1$ nm
Correlogram Interpretation

- Small Molecules
- Large Particles
Correlogram Interpretation
Correlogram Interpretation

The time at which the correlation of the signal starts to decay gives information about the mean diameter.
The time at which the correlation of the signal starts to decay gives information about the mean diameter.

The angle of decay gives information about the polydispersity of the distribution.
The time at which the correlation of the signal starts to decay gives information about the mean diameter. The angle of decay gives information about the polydispersity of the distribution. The baseline gives information about the presence of large particles/aggregates.
Distributions By DLS

Comparison of Z average (Cumulants) size to multimodal distribution results.

Z average: 12.4 nm

(Single species assumption)
Intensity, Volume And Number Distributions

Mixture containing equal numbers of 5 and 50nm spherical particles

**NUMBER**

**VOLUME**

\[ V = \frac{4}{3} \pi r^3 \]

**INTENSITY**

\[ I = r^6 \]
Benefits Of Sizing By DLS

- Non-invasive
- High sensitivity (< 0.1 mg/mL for typical proteins)
- Low volume (12 µL)
- Scattering intensity is proportional to the square of the protein molecular weight, making the technique ideal for identifying the presence of trace amounts of aggregate.
Electrophoretic Light Scattering

Molecular Charge Measurements
Electrophoretic Light Scattering (ELS)

Measured parameter is the frequency shift ($\Delta \nu$) of the light scattered from a moving particle.

$\nu = \nu_0 + \Delta \nu$

$\mu = K \left( \frac{\Delta \nu}{E} \right)$

$\mu$ is the electrophoretic mobility, $E$ is the electric field strength, and $K$ is a constant.
Measuring Electrophoretic Mobility

- Classical capillary electrophoresis (light microscope, stopwatch)

The particles move with a characteristic velocity which is dependent on:

- Field strength
- Dielectric constant of medium
- Viscosity of the medium
- Zeta potential
Electrophoresis

- Electrophoresis is the movement of a charged particle relative to the liquid it is suspended in under the influence of an applied electric field.
- The electrophoretic mobility of a colloidal dispersion can be used to determine the zeta potential.
- Zeta potential is the charge a particle acquires in a particular medium.
- Zeta potential measurements can be used to predict dispersion stability.
- Influenced by: pH, salts, concentration, additives,…
Electroosmosis is the movement of liquid relative to a stationary charged surface under the influence of an applied field.
Measuring Electrophoretic Mobility

- Laser Doppler electrophoresis (LDE)
  - Phase analysis light scattering (PALS)
  - Mixed mode measurements (M3)
- A laser beam is passed through the sample in the capillary cell undergoing electrophoresis
- Scattered light from moving particles is frequency shifted
- These small frequency shifts are measured
- The frequency shift $\Delta f$ is equal to:

$$\Delta f = 2v \sin(\theta/2)/\lambda$$

$v =$ the particle velocity
$\lambda =$ laser wavelength
$\theta =$ scattering angle

...measure phase instead
Mixed Mode Measurement (M3)

- Mixed mode measurement (M3) is a patented method that allows measurement at any point in a capillary cell.
- It eliminates electroosmosis by reversing the applied field at a high frequency.
- Malvern have combined M3 with PALS to improve the measurement sensitivity and accuracy (M3-PALS).
Phase Analysis Light Scattering

Phase Difference Demonstration
Phase Plot From General Purpose
Phase Plot From General Purpose

FFR

PALS only to obtain mean $(E_p)$
Phase Plot From General Purpose

FFR

SFR

PALS only to obtain mean ($E_p$)

PALS to obtain mean ($E_o + E_p$) + FT to obtain distribution (width)
Zetapotential Distribution Plot
General Purpose
Light Scattering Return

- Hydrodynamic Radius
- Distribution & Polydispersity
- Solution Composition
- Molecular Weight
- 2nd Virial Coefficient
- Conformation
- Shape Estimates
- Zeta Potential
- pI & Charge Estimates
- Formulation Stability
Application Example

Molecular Weight Measurements
Absolute Protein Molecular Weight

![Graph comparing different proteins' molecular weight vs. concentration](image)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Conditions</th>
<th>Measured</th>
<th>Known</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>0.1 M NaCl, pH 7</td>
<td>14.9</td>
<td>14.3</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>0.1 M NaHAc, pH 4.2, 4% NaCl</td>
<td>14.6</td>
<td>14.3</td>
</tr>
<tr>
<td>Interferon γ</td>
<td>5 mM Histidine, pH 5</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Interferon γ</td>
<td>5 mM Succinate, pH 5</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>0.1 M NaCl, pH 7</td>
<td>48</td>
<td>43</td>
</tr>
<tr>
<td>BSA</td>
<td>0.1 M NaCl, pH 7</td>
<td>64</td>
<td>67</td>
</tr>
</tbody>
</table>


www.malvern.co.uk © Malvern Instruments Ltd. - A spectris Company
**A₂ Crystallization Window**

Crystallization Window: 

\[-0.8 > A₂ > -8 \times 10^{-4} \text{ mol mL / g}^2\]


**Rayleigh Equation**

\[
\frac{KC}{R_0} = \frac{1}{M} + 2A₂C
\]
Application Example

Molecular Size Measurements
## Common Proteins

Comparison of known to estimated molecular weight for common proteins.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Dia (nm)</th>
<th>Known MW (kDa)</th>
<th>Estimated MW (kDa)</th>
<th>%Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>3.8</td>
<td>14.7</td>
<td>15.1</td>
<td>-2.7</td>
</tr>
<tr>
<td>Chymotrypsinogen</td>
<td>4.8</td>
<td>25</td>
<td>26.1</td>
<td>-4.4</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>5.2</td>
<td>29</td>
<td>31.5</td>
<td>-8.6</td>
</tr>
<tr>
<td>Human Insulin (pH 7)</td>
<td>5.4</td>
<td>34.2</td>
<td>34.4</td>
<td>-0.6</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>6</td>
<td>43</td>
<td>44</td>
<td>-2.3</td>
</tr>
<tr>
<td>Hexokinase sub-unit</td>
<td>6.6</td>
<td>51</td>
<td>55</td>
<td>-7.8</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>7</td>
<td>65</td>
<td>63</td>
<td>3.1</td>
</tr>
<tr>
<td>Bovine Serum Albumin</td>
<td>7.1</td>
<td>67</td>
<td>65.3</td>
<td>2.5</td>
</tr>
<tr>
<td>Horse Alcohol Dehydrogenase</td>
<td>7.4</td>
<td>80</td>
<td>71.9</td>
<td>10.1</td>
</tr>
<tr>
<td>Amyloglucosidase</td>
<td>7.8</td>
<td>99</td>
<td>81.3</td>
<td>17.9</td>
</tr>
<tr>
<td>Hexokinase</td>
<td>8.6</td>
<td>102</td>
<td>102.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>Yeast Alcohol Dehydrogenase</td>
<td>9.8</td>
<td>150</td>
<td>138.7</td>
<td>7.5</td>
</tr>
<tr>
<td>Apoferritin</td>
<td>16.4</td>
<td>443</td>
<td>462.7</td>
<td>-4.4</td>
</tr>
<tr>
<td>Thyroglobulin</td>
<td>20.2</td>
<td>669</td>
<td>753.5</td>
<td>-12.6</td>
</tr>
</tbody>
</table>
DLS results indicate an insulin structure that is dimeric at pH 2 and hexameric at pH 7, consistent with crystallographic data.
Application Example

Polydispersity Measurements
Polydispersity (Pd) From DLS

Pd is representative of the particle size distribution width.

**Monodisperse**

- %Pd = 13.1

**Polydisperse**

- %Pd = 46.6

60 second measurement
Crystal Screening Using DLS

A$_2$ ZONE

I

Too soluble

II

High quality crystals

II + Impurity

Low quality or no crystals

III

Amorphous precipitate

Low Concentration
Light Scattering Conditions

High Concentration
Crystal Growth Conditions
Monoclonal Antibody Fragment
Size, zeta potential and molecular weight

Mean diameter = 5.1nm

Mean zeta potential = -7.6mV

\[ M_W = 20.7 \text{KDa} \]
\[ A_2 = -0.0049 \text{ ml mol/g}^2 \]
Malvern Instruments

& the Zetasizer Nano
Zetasizer Nano
Optics of the Zetasizer Nano
## Zetasizer Technical Specifications

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sizing range</td>
<td>0.6 nm to 6 μm Diam</td>
</tr>
<tr>
<td>Concentration range</td>
<td>0.1 mg/mL Lys to 30w%</td>
</tr>
<tr>
<td>Min sizing sample volume</td>
<td>12 μL</td>
</tr>
<tr>
<td>Min zeta sample volume</td>
<td>0.75 mL</td>
</tr>
<tr>
<td>Temperature control</td>
<td>2 to 90 °C</td>
</tr>
<tr>
<td>Conductivity range</td>
<td>0 to 200 mS/cm</td>
</tr>
<tr>
<td>Laser</td>
<td>3 mW 633 nm HeNe</td>
</tr>
<tr>
<td>Detector</td>
<td>APD</td>
</tr>
</tbody>
</table>

- Crystal screening
- Protein & polymer characterization
- CMC measurements
- Drug delivery systems
- Formulation stability
- Biological assemblies
- Virus & vaccine characterization
- Macromolecular critical points
More information

- Application notes
- Multimedia presentations
- Brochures
- Detailed specifications

www.malvern.co.uk/proteins