Application of Light Scattering Techniques for Analysis of Oligomerization and Particle Formation

Ewa Folta-Stogniew
Yale University
Outline

• **Light Scattering Technologies**
  – Static and dynamic light scattering
  – Parameters derived from SLS and DLS measurements

• **Batch Light Scattering Applications**
  – Detection of aggregates in DLS and SLS measurement

• **Flow Mode Light Scattering Applications**
  – Molar mass distributions and differences in populations
  – Characterization of morphology of aggregates

• **Determination of an oligomeric state of modified proteins from SEC-LS/UV/RI measurement**

• **Capabilities and limitation of static and dynamic LS measurements**
Light Scattering Experiments

Monochromatic Laser Light

Sample cell

$I_o$ $I$

$\Theta$

detector

Computer
Light Scattering Experiments

- **Static (classical)**
  - time-averaged intensity of scattered light

- **Dynamic (quasielastic)**
  - fluctuation of intensity of scattered light with time

Measurements:
- batch mode
- “in-line” mode combined with a fractionation step,
  - i.e. chromatography, mainly Size Exclusion Chromatography, Flow Field Fractionation
Light Scattering Experiments

• **Static (classical)**
  
  - Time-averaged intensity of scattered light

Parameters derived:

- Molar Mass (weight-average) accuracy ~5%
- \(\langle r_g^2 \rangle^{1/2}\) root mean square radii
  
  for \(\langle r_g^2 \rangle^{1/2} \sim (\lambda/20) \sim 15 \text{ nm}\)
- \(A_2\) second virial coefficient

Rayleigh-Debye-Zimm formalism

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

- \(R(\theta)\) Rayleigh ratio (excess scattered light)
- \(c\) sample concentration (g/ml)
- \(M_w\) weight-average molecular weight (molar mass)
- \(A_2\) second virial coefficient (ml-mol/g2)
- \(P(\theta)\) form factor (angular dependence)

• **Dynamic (quasielastic)**
  
  - Fluctuation of intensity of scattered light with time

Parameters derived:

- \(D_T\) translation diffusion coefficient
- \(R_h\) hydrodynamic radius (Stokes radius)

Uncertainty of ~10% for monodisperse sample

Stokes-Einstein

\[
D_T = \frac{kT}{6\pi\eta R_h}
\]

- \(D_T\) translational diffusion coefficient
- \(k\) Boltzmann constant
- \(T\) temperature
- \(R_h\) radius
- \(\eta\) solvent viscosity
Why Light Scattering?

• Scattering Intensity, $R(\Theta) \sim Mw^c$
  
  because of their big Mw, aggregates scatter strongly even when present at low concentrations; easily detectable

• Angular variation of the scattered light is related to the size of the molecule
  
  the light scattering signal from aggregates will show angular dependence, while LS signal produces by lower order oligomers like dimers, trimers, tetramers, et c. will not

• LS measurements are non-invasive and non-destructive
  
  • small sample volumes
  • great dynamic range for sizing: hydrodynamic radii ~ 2nm to 500 nm
  • great dynamic range for Mw determination: < 1kDa to >10 MDa
  • wide range of concentrations (non-ideality can be addressed through the determination of second virial coefficient)
  • perfectly suited for determination of oligomeric state of modified proteins without prior knowledge of extend of modification (glycosylated, modified by polyethylene glycol, or membrane proteins present as complexes with lipids and detergents
Determination of hydrodynamic radius, $R_h$, from a Dynamic LS experiment

Ovalbumin; monomer: 43 kDa; $R_h=3.0$ nm

$R_h = 8\pm7$ nm from Cumulant Fit (Polydispersity 93%)

Regularization Fit:

<table>
<thead>
<tr>
<th>Peak</th>
<th>Rh (nm)</th>
<th>Polydispersity (%)</th>
<th>MW (R) kDa</th>
<th>% Intensity</th>
<th>% Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.1</td>
<td>12.8</td>
<td>46</td>
<td>54</td>
<td>99.9</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>17.8</td>
<td>&gt;1MDa</td>
<td>23</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>13.4</td>
<td>&gt;1MDa</td>
<td>23</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>

Intensity Autocorrelation

% Intensity

Time (µs)

Radius (nm)
Results from a batch mode Dynamic LS experiment:

Ovalbumin 43 kDa; Rh=3.0 nm

Rh = 3.2±0.6 nm from Cumulant Fit (Polydispersity 19%)
Dissociation of aggregates upon dilution; time course

Protein H 23 kDa; Rh=2.3 nm

Rh=94 nm

Rh=2.3 nm; Pd=42%
Determination of Molar Mass and second virial coefficient from a batch static LS experiment

BSA 66 kDa

Mw = 62 kDa

A₂ = (5.226 ± 0.316)e⁻⁴ mol mL/g²

**Zimm plot analysis of static light scattering data**

\[
\frac{K \cdot c}{R(\theta)} = \frac{1}{M_w P(\theta)} + 2A_2 c
\]

and Rh from DLS

Batch Mode Static MALLS experiment

Monomer 14 kDa

Angular dependence of scattered light clearly indicates presence of aggregates

<table>
<thead>
<tr>
<th>Sample</th>
<th>Weight Average MM, Mw ± SD* [kDa]</th>
<th>RMS [nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15 ± 1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>126 ± 8</td>
<td>56 ± 10</td>
</tr>
</tbody>
</table>
Feature detected in a batch mode LS measurements for sample containing aggregates

- **Static (classical)**
  - Aggregates present:
    - elevated weight average Molar Mass ($M_w$ weight average)
    - angular dependence in scattered light

- **Dynamic (quasielastic)**
  - Aggregates present:
    - autocorrelation function cannot be described by single exponential (cumulant fit)
    - polydispersity from cumulant fit >15%

**Missing information:** how much and what size?

**Solutions**

- Sample fractionation followed by batch measurements
- **Column separation with simultaneous LS characterization**
Three Detector monitoring

Peak ID - Ova_071305a_01_P_N

- UV at 280 nm
- RI
- LS at 90°
Ovalbumin 43 kDa

- 88% monomer
- 8% dimer
- 1.5% trimer
- 3% aggregates < 1MDa
- 0.4% 1-100 MDa

**UV at 280 nm**

**LS at 90 deg**

Detector: AUX1

Detector: 11

Strip Chart - OVA_b_UV_traces
Intensity of scattered light $\sim M_w^c$

due to their high $M_w$ aggregates scatter very strongly

A monomeric protein 43 kDa and aggregates 10 MDa at 2 mg/mL:
Ovalbumin 43 kDa

Aggregates
angular dependence of scattered light

Lower order oligomers
no angular dependence of scattered light
Molar mass distribution for multiple analyses

Ovalbumin 43 kDa automated template processing of five data sets
Determination of Weight Fractions

Cumulative Molar Mass

Molar Mass (g/mol)

Cumulative Weight Fraction $W$

- OVA_e_UV
- OVA_200_a_P_N_UV_template...
- OVA_b_UV
- OVA_c_UV
- OVA_d_UV
Differences in population based on molar mass distribution

Ovalbumin (5 runs)
Mw = 108 ± 17 kDa
Polydispersity Mw/Mn 2.3 ± 0.4

Ovalbumin (3 runs)
Mw = 141 ± 3 kDa
Polydispersity Mw/Mn 2.92 ± 0.06
Differences in population based on molar mass distribution

Ovalbumin 43 kDa

- **Ovalbumin (5 runs)**
  - $\text{MMw} = 108 \pm 17 \text{ kDa}$
  - Polydispersity $\text{Mw/Mn} = 2.3 \pm 0.4$

- **Ovalbumin (3 runs)**
  - $\text{MMw} = 141 \pm 3 \text{ kDa}$
  - Polydispersity $\text{Mw/Mn} = 2.92 \pm 0.06$
Differences in population based on molar mass distribution

Ovalbumin 43 kDa

<table>
<thead>
<tr>
<th>Oligomeric state</th>
<th>Average Mw ± SD [kDa] (5 analyses)</th>
<th>Average Mw ± SD [kDa] (3 analyses)</th>
<th>Fraction of Mass [% of total] (5 analyses)</th>
<th>Fraction of Mass [% of total] (3 analyses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono (20-50 kDa)</td>
<td>43.0 ± 0.1</td>
<td>42.80 ± 0.02</td>
<td>88.1 ± 0.1</td>
<td>85.23 ± 0.06</td>
</tr>
<tr>
<td>Di (50-96 kDa)</td>
<td>82.7 ± 0.4</td>
<td>84.1 ± 0.2</td>
<td>7.68 ± 0.04</td>
<td>9.4 ± 0.0</td>
</tr>
<tr>
<td>Tri (96-130 kDa)</td>
<td>114 ± 4</td>
<td>121.8 ± 0.7</td>
<td>1.54 ± 0.05</td>
<td>1.9 ± 0.0</td>
</tr>
<tr>
<td>Agg. (0.13 –1 MDa)</td>
<td>270 ±10</td>
<td>284 ± 2</td>
<td>2.18 ± 0.08</td>
<td>2.87 ± 0.06</td>
</tr>
<tr>
<td>Agg. (1 –100 MDa)</td>
<td>10±1 x10^3</td>
<td>10.9±0.4 x10^3</td>
<td>0.4 ± 0.0</td>
<td>0.6 ± 0.0</td>
</tr>
</tbody>
</table>
Morphology of aggregates from angular dependence of LS signal;

size determination- Rg

Determination of radius of gyration, 
\( R_g \) (root mean square radius, 
R.M.S.,) from angular dependence 
of scattered light

Zimm Plot

\[
\frac{K*c}{R(\theta)} = \frac{1}{M_v} (1 + \frac{16 \pi^2}{3 \lambda^2}) <R^2 \sin^2(\frac{\theta}{2})>
\]

Radius: 46.8±0.2 nm

- Peak, Slice : 1, 944
- Volume : 7.867 mL
- Fit degree : 1
- Conc. : (1.915 ± 0.020)e-6 g/mL
- Mw : (2.277 ± 0.024)e+7 g/mol

- RMS Radius vs. Volume

- Ova_122105a_Rg-Rh
Inferring conformational information from the relationship between molecular size ($R_g$) and molecular weight (Molar Mass)

$$R_g \sim M^\nu$$

log($R_g$) versus log(MM)

Slope = $\nu$

<table>
<thead>
<tr>
<th>For</th>
<th>$\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>0.33</td>
</tr>
<tr>
<td>Coil</td>
<td>0.5</td>
</tr>
<tr>
<td>Rod</td>
<td>1</td>
</tr>
</tbody>
</table>

Shape analysis: $\log(R_g)$ versus $\log(MM)$

Aggregates of Ovalbumin vs. “amyloid-type” fibers

For

<table>
<thead>
<tr>
<th>Shape</th>
<th>$\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>0.33</td>
</tr>
<tr>
<td>Coil</td>
<td>0.5</td>
</tr>
<tr>
<td>Rod</td>
<td>1</td>
</tr>
</tbody>
</table>

Ova_aggr $\nu = 0.4$ Sphere/Coil

Amyloids $\nu = 0.8$ Coil/Rod
Shape analysis: shape factor $\rho = \frac{R_g}{R_h}$

Aggregates of Ovalbumin vs. amyloid fibers

Shape factor: $\rho = \frac{R_g}{R_h}$

Combination of MALS ($R_g$) and DLS ($R_h$)

For $\rho = \frac{R_g}{R_h}$
- Sphere: 0.774
- Coil: 0.816
- Rod: 1.732

Ovalbumin

Amyloids
Shape analysis: \textbf{shape factor} $\rho = \frac{R_g}{R_h}$

Aggregates of \textit{Ovalbumin} vs. \textit{amyloid fibers}

Shape factor: \quad $\rho = \frac{R_g}{R_h}$

\textbf{Combination of MALS ($R_g$) and DLS ($R_h$)}

\begin{itemize}
  \item \textbf{Ovalbumin}
  \item \textbf{Amyloids}
\end{itemize}

Shape analysis: shape factor $\rho = \frac{R_g}{R_h}$

Aggregates of Ovalbumin vs. amyloid fibers

Shape factor: $\rho = \frac{R_g}{R_h}$

Combination of MALS ($R_g$) and DLS ($R_h$)

Ovalbumin

$R_g/R_h = 0.91$  Coil

Amyloids

$R_g/R_h = 1.84$  Rod
Shape analysis: shape factor $\rho = \frac{R_g}{R_h}$

Aggregates of Ovalbumin vs. amyloid fibers

Shape factor: $\rho = \frac{R_g}{R_h}$

Combination of MALLS ($R_g$) and DLS ($R_h$)

For $\rho = \frac{R_g}{R_h}$

<table>
<thead>
<tr>
<th></th>
<th>$\rho = \frac{R_g}{R_h}$</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>

Ovalbumin

$R_g/R_h = 0.91$  
Ova_aggr $\nu = 0.4$  
Sphere/Coil

Amyloids

$R_g/R_h = 1.84$  
Amyloids $\nu = 0.8$  
Coil/Rod
Various uses of Light Scattering for assessing protein aggregates

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Detects Aggregates</th>
<th>Information about population (distribution)</th>
<th>Challenge in use</th>
<th>Sample dilution</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLS</td>
<td>Yes</td>
<td>No</td>
<td>Low</td>
<td>No</td>
<td>Fast</td>
</tr>
<tr>
<td>Micro-batch MALS</td>
<td>Yes</td>
<td>No</td>
<td>High</td>
<td>No</td>
<td>Medium</td>
</tr>
<tr>
<td>SEC/MALLS/DLS</td>
<td>Yes</td>
<td>Yes</td>
<td>Medium</td>
<td>Yes</td>
<td>Medium</td>
</tr>
</tbody>
</table>
Determination of the oligomeric state of modified proteins from SEC-LS/UV/RI analysis

1. Glycosylated proteins
2. Proteins conjugated with polyethylene glycol
3. Membrane protein present as a complex with lipids and detergents

Input:
• Polypeptide sequence
• Chemical nature of the modifier

Results:
• Oligomeric state of the polypeptide
• Extend of modification (grams of modifier /gram of polypeptide)

“three detector method”
Three Detector Method

\[ MW_p = \frac{k \cdot (LS)(UV)}{\varepsilon(RI)^2} \]

- \( MW_p \): Molecular Weight (polypeptide)
- \( \varepsilon \): extinction coefficient
- \( LS \): light scattering intensity
- \( UV \): absorbance (\( \varepsilon \))
- \( RI \): refractive index change
- \( k \): calibration constant

\[ MW_p = \frac{k \cdot (LS)(UV)}{\varepsilon(RI)^2} \]

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ova</td>
<td>43</td>
</tr>
<tr>
<td>BSA(1)</td>
<td>66</td>
</tr>
<tr>
<td>BSA(2)</td>
<td>132</td>
</tr>
<tr>
<td>Ald</td>
<td>156</td>
</tr>
<tr>
<td>Apo-Fer</td>
<td>475</td>
</tr>
</tbody>
</table>

Three-detector calibration
10-17-01

\[ y = 92.383x - 3.4044 \]
\[ R^2 = 0.9996 \]
PEG-ylated protein: 75 kDa

36 kDa polypeptide + 39 kDa PEG

Polypeptide: 146 kDa
(tetramer: 144 kDa)

Full protein: 291 kDa
(tetramer: 300 kDa)
PEG-ylated protein: 75 kDa

36 kDa polypeptide + 39 kDa PEG

Polypeptide: 146 kDa
(tetramer: 144 kDa)

Full protein: 291 kDa
(tetramer: 300 kDa)
Capabilities

Static LS
• fast and accurate determination of molar masses (weight average)
  – glycosylated protein, conjugated with PEG, protein-lipids-detergent complexes,
    protein-nucleic acid complexes
• accuracy of ± 5% in Molar Mass determination
• easy to implement, fully automated (data collection and data analysis)
• highly reproducible (no operator bias)
• SEC/MALS excellent in detecting and quantifying population with various oligomeric state
  in protein

Dynamic LS
• very fast detection of aggregates
• great dynamic range
• well suited to study kinetics of aggregation
• DLS detector available in a plate reader format for high volume analyses

Combined data about MM, Rg and Rh - shape information (multiangle static and dynamic LS)
• via frictional ratio Rh/Rs
• via shape factor ν, from log(Rg) vs. log(MM) plot
• via shape factor ρ, from Rg/Rh ratio
Limitations

**Static LS**
- measures weight average molar mass – needs fractionation to resolve different oligomeric states
- possible losses of sample during filtration and fractionation
- limitation on solvent choices (related to a fractionation step)
- SEC/SLS/DLS dilution during experiment

**Dynamic LS**
- measures hydrodynamic radius, which is affected by shape
- cannot discriminate between shape effects and changes in oligomeric states, *i.e.* non-spherical shape mimics oligomerization
- needs fractionation to resolve low number oligomers when present in mixture
Ken Williams
Director of W.M. Keck Biotechnology Resource Laboratory at Yale University School of Medicine

NIH

Users of SEC/LS Service

http://info.med.yale.edu/wmkeck/biophysics
Ewa.Folta-Stogniew@yale.edu