Impurities In Biomolecules
Institute for International Research

Monitoring & Predicting
Biomolecular Aggregation Using
Light Scattering

Kevin Mattison
Malvern Instruments

Ewa Folta-Stogniew
Yale University
Outline

- Light Scattering Technologies
- Batch Light Scattering Applications
  - Kevin Mattison – Malvern Instruments
- Flow Mode Light Scattering Applications
  - Ewa Folta-Stogniew – Yale University
- Closing
- Appendix
Biomolecular Stability

*With regard to aggregation:*

- In the absence of stabilizing “conditions”, when small particles collide, London forces can dominate the interaction, leading to particle aggregation.
- In order to stabilize a formulation against aggregation, particle collisions must be minimized. This can be accomplished using:
  - Electrostatic Effects - wherein the presence of charge leads to a repulsive force between the particles.
  - Steric Effects - wherein the presence of adsorbed or attached additives (known as chaotrophic agents) prohibit particles from getting close enough together for London forces to dominate.
Stabilizing Effects

- In the absence of steric effects, the stability of a system to aggregation is determined by the balance of repulsive and attractive forces which the particles experience as they approach one another. The rule of thumb for electrostatically stable suspension is +/- 30 mV for the zeta potential.

- In the absence of electrostatic effects, steric stability is enhanced upon addition of chaotropic or “structure disrupting” agents, which reduce the likelihood of particles getting close enough for London forces to take over.
Evidence Of Instability?

Insulin formulations at $t = 0$ and 12 months

This insulin formulation can help you recover from hyperglycemic diabetic shock.

This insulin formulation can kill you.
Common Approach – Time Trials

According to sizing results, the sample is completely aggregated within 7 days of preparation.
But Can We Predict?

**Light Scattering:** Low energy photon induces an oscillating dipole in the electron cloud. As the dipole oscillates, energy is radiated in all directions. This radiated energy is called “scattered light.”

*Rayleigh Scattering Profile*
Static Light Scattering (SLS)

Average scattering intensity leads to the particle molecular weight, 2nd virial coefficient, and radius of gyration ($R_g$).

\[
\frac{K C}{R_\theta} = \left( \frac{1}{M} + 2A_2 C \right) P(\theta)
\]

- $K$ = Optical constant
- $M$ = Molecular weight
- $A_2$ = 2nd Virial coefficient
- $C$ = Concentration
- $R_\theta$ = Rayleigh ratio
- $P(\theta)$ = Shape factor
Dynamic Light Scattering (DLS)

Correlation of short time scale ($\mu$s) intensity fluctuations gives the diffusion coefficient, hydrodynamic size, polydispersity, and particle size distribution.

**Stokes-Einstein**

$$R_H = \frac{kT}{6\pi\eta D}$$

$q = \text{Scattering vector}$

$R_H = \text{Radius}$

$T = \text{Temperature}$

$D = \text{Diffusion coefficient}$

$k = \text{Boltzmann constant}$

$\eta = \text{Solvent viscosity}$
Electrophoretic Light Scattering (ELS)

Measured parameter is the frequency shift of the scattered light.

The frequency shift is proportional to the electrophoretic mobility, which is a function of the particle surface potential. Hence ELS gives us information regarding the charge on the particle.

\[
\mu = K\left(\frac{\Delta \nu}{E}\right)
\]
Why Light Scattering?

The scattering intensity:
- varies with the mass and concentration according to the Rayleigh Expression
- is proportional to
  - $M_w$
  - $M_n^2$
  - $R^6$
- is non-invasive
- is ideal for aggregate detection & quantification in low volume, low concentration biological samples.

<table>
<thead>
<tr>
<th>Peak</th>
<th>$D_i$ (nm)</th>
<th>%I</th>
<th>%M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.95</td>
<td>5</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>46.0</td>
<td>94</td>
<td>6</td>
</tr>
</tbody>
</table>
Lysozyme - Comparison Of Radii

Lysozyme

- $M = 14.5\ \text{kDa}$
- $V_p = 0.73\ \text{mL/g}$
- $R_R = 2.25\ \text{nm}$
- $R_H = 1.90\ \text{nm}$
- $R_M = 1.61\ \text{nm}$
- $R_g = 1.47\ \text{nm}$

For

\[
R_g = \begin{align*}
\text{Sphere} & : 0.774 R_H \\
\text{Coil} & : 0.816 R_H \\
\text{Cylinder} & : 1.732 R_H
\end{align*}
\]
What Is $M_W$?

$M_W$ is the mass or weight average molecular weight.

**Number Average**

$$M_N = \frac{\sum_i N_i M_i}{\sum_i N_i}$$

**Weight Average**

$$M_W = \frac{\sum_i m_i M_i}{\sum_i m_i}$$

$N_i =$ the number of particles in each weight class

$M_i =$ the molecular weight of each weight class

$m_i =$ the mass of particles in each weight class

As a consequence, $M_W$ is more heavily weighted by larger particles in the sample.
Light Scattering Applications

Batch Mode Measurements
Monitoring Stability

40% Propofol emulsion

Temperature
- 10 C
- 20 C
- 25 C
- 35 C
- 45 C
- 55 C

Initial
1 Month
3 Months
9 Months

Amplitude
Diameter (nm)
Predicting Stability

**Formulation A**
Mean = -52mV (±0.8)

**Formulation B**
Mean = -24mV (±0.1)

Unstable

Stable
Prediction & Observation

Soy protein formulation

- Stock
- Treated

Z Average Diameter (nm)
- pH
- Zeta Potential (mV)
Zeta Potential - Shelf Life Correlation

Formulations with varying soy fiber content.

![Graph showing the correlation between Zeta Potential (mV) and Shelf Life (Months) for Nutritional Beverages. The graph indicates a negative correlation, with higher Zeta Potential values associated with shorter shelf lives.]
Electrostatic vs. Steric Stability

Polymer adsorption to cationic liposomes in PBS reduces the electrostatic while enhancing the steric stabilization. All are stable.
Salt Effects On Aggregation

Electrostatic shielding enhances BSA aggregation for NaCl concentrations $\geq 500$ mM at the isoionic point. But the aggregation is reversible, suggesting that it is “non-denaturing”.

![Graph showing aggregation of Bovine Serum Albumin, pH 4.8, with different NaCl concentrations](image)
Polydispersity (Pd) From DLS

Pd is representative of the particle size distribution width, with high polydispersity being indicative of oligomerization and/or aggregation.

**Monodisperse**

\[ \% \text{Pd} = 13.1 \]

**Polydisperse**

\[ \% \text{Pd} = 46.6 \]
High Concentration Sizing

**Equilibrium Aggregation**
- Unstable – Short Shelf Life

**Restricted Diffusion**
- Stable – Long Shelf Life

Concentration (W%)
- 50
- 0.5
- 50 (corrected)
Using $T_M$ As A Stability Predictor

In phosphate buffered saline

Melting Points

- Lytic Enzyme: 56 C
- Collagenase: 64 C
- Ribonuclease A: 72 C
Melting Point - Shelf Life Correlation

Samples prepared in phosphate buffered saline at 1.0 mg/mL protein concentration.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Tm (C)</th>
<th>Shelf Life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast Lytic Enzyme</td>
<td>56</td>
<td>&lt; 10 Hrs</td>
</tr>
<tr>
<td>Collagenase</td>
<td>64</td>
<td>50 Hrs</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>72</td>
<td>180 Hrs</td>
</tr>
</tbody>
</table>
Using $A_2$ As A Stability Predictor

$A_2$ is closely correlated with sample solubility

![Diagram showing correlation between $A_2$ and solubility](image-url)
Batch Mode Challenges – Oligomers!

**Static LS**
- **Classical**
  - elevated weight average Molar Mass ($M_w$ weight average)
  - angle dependent intensity

**Dynamic LS**
- **Quasi-elastic**
  - autocorrelation function cannot be fit to single exponential (Cumulant)
  - high polydispersity (%Pd > 15%)

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

**Solutions**
- Sample fractionation followed by batch measurements
- Column separation with simultaneous LS characterization
Light Scattering Applications

Flow Mode Measurements
HPLC system

SEC column

Sample

0.1 μm pre-filtered buffer
0.1 μm “in-line” filter

UV detector

LS detector
DLS+SLS

RI detector

Computer

Waste
Three Detector monitoring

Peak ID - Ova_071305a_01_P_N

- LS #11
- AUX1
- AUX2

Volume (mL)

LS #11, AUX1, AUX2

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

Strip Chart - OVA_b_UV_traces

UV at 280 nm

LS at 90 deg
Intensity of scattered light $\sim M_w^*c$

due to their high Mw aggregates scatter very strongly

*A monomeric protein 43 kDa and aggregates 10 MDa at 2 mg/mL:*
Molar mass distribution for multiple analyses

Ovalbumin 43 kDa

automated template processing of five data sets

Molar Mass vs. Volume

- OVA_e_UV
- OVA_200_a_P_N_UV_templat...
- OVA_b_UV
- OVA_c_UV
- OVA_d_UV
Determination of Weight Fractions
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, (5 runs)

$\text{MM}_w = 108 \pm 17 \text{ kDa}$

$\text{Polydispersity} \; \frac{M_w}{M_n} = 2.3 \pm 0.4$

Ovalbumin (3 runs)

$\text{MM}_w = 141 \pm 3 \text{ kDa}$

$\text{Polydispersity} \; \frac{M_w}{M_n} = 2.92 \pm 0.06$
## 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>Mw = 108 ± 17</td>
<td>Mw = 141 ± 3</td>
<td>Mw = 108 ± 17</td>
<td>Mw = 141 ± 3</td>
</tr>
<tr>
<td></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 x10³</td>
<td>10.9±0.4 x10³</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

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

Zimm Plot

- Peak, Slice : 1,944
- Volume : 7.867 mL
- Fit degree : 1
- Conc. : \((1.915 \pm 0.020)e^{-6}\) g/mL
- Mw : \((2.277 \pm 0.024)e^{7}\) g/mol
- Radius : 46.8±0.2 nm

RMS Radius vs. Volume

- 90° & AUX detector

\( \text{K}^* \text{c} \)
\( \text{R(\theta)} \)
\( \text{M}_w \)
\( \text{<R}^2 > \)
\( \text{sin}^2 \left( \frac{\theta}{2} \right) \)
Inferring conformational information from the relationship between molecular size (Rg) and molecular weight (Molar Mass)

\[ R_g \sim M^\nu \]

\( \log(R_g) \) versus \( \log(MM) \)

Slope = \( \nu \)

<table>
<thead>
<tr>
<th>Form</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(Rg) versus log(MM)

Aggregates of Ovalbumin vs. “amyloid-type” fibers

<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>

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

<table>
<thead>
<tr>
<th>Shape</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

Amyloids
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

Amyloids

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

<table>
<thead>
<tr>
<th>Shape</th>
<th>Value</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>

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

<table>
<thead>
<tr>
<th>Shape</th>
<th>Value</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>

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

Ovalbumin $\rho = 0.91$  -  Coil

Amyloids $\rho = 1.84$  -  Rod

Ova_aggr  $\nu = 0.4$  -  Sphere/Coil

Amyloids $\nu = 0.8$  -  Coil/Rod
Determination of the oligomeric state of modified protein from SEC/LS 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

\[ \text{MW}_p = \frac{k \ast (LS)(UV)}{\varepsilon (RI)^2} \]

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

\[ MW_{p} = 91.39 \times \frac{[(LS) \times (UV)]}{(\varepsilon \times (RI^2))} \]

<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

\[ y = 92.383x - 3.4044 \]
\[ R^2 = 0.9996 \]

Graph showing the relationship between the ratio \((LS)\times(UV)/(A \times (RI^2))\) and MW [kDa]. The red line represents the linear (theoretical MWp) relationship, and the blue diamonds represent the experimental MWp values.
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)
PEG-ylated oligo: 48.3 kDa

8.3 kDa oligo + 40 kDa PEG

PEG (40K)  MM = 41.0 kDa
Polydispersity= 1.001
40K PEG + 8.3 kDa oligo
PEG-oligo  MM = 48.5 kDa
Protein “F”

40.3 kDa

156 kDa

43 kDa

Molar Mass vs. Volume
Protein “F” frictional ratio $R_h/R_s = 1.85$ non-spherical shape

40.3 kDa $R_h = 4.2$ nm

156 kDa $R_h = 4.2$ nm

43 kDa $R_h = 2.9$ nm

Shape Effects

4 fold difference in mass – same size
Protein: 12 kDa; WT and three mutants

Interaction with the column effects

Trimer:
MW = 36 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 \( \nu \), from log(Rg) vs. log(MM) plot
  - via shape factor \( \rho \), 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
# Common Light Scattering Specifications

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size range - DLS</td>
<td>0.6 nm to 6 um Diam</td>
</tr>
<tr>
<td>Size range - Zeta Potential</td>
<td>10 nm to 20 um Diam</td>
</tr>
<tr>
<td>Concentration range</td>
<td>0.1 mg/mL (Lys) to 30w%</td>
</tr>
<tr>
<td>Minimum sample volume</td>
<td>2 uL</td>
</tr>
<tr>
<td>Temperature control</td>
<td>-4 to 130 °C</td>
</tr>
</tbody>
</table>

**Accessories**
- Polarization filters for rotational correlation measurements
- Wavelength filters for fluorescing samples
- Automatic titrators
- Cross-correlation configurations
- Plate readers for high throughput applications
- Multi-angle configurations for full MW & Rg range
- Flow cells for HPLC applications
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
Thank You

Questions?

Kevin Mattison  
*Malvern Instruments*  
kevin.mattison@malvern.com

Ewa Folta-Stogniew  
*Yale University*  
Ewa.Folta-Stogniew@yale.edu