# 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

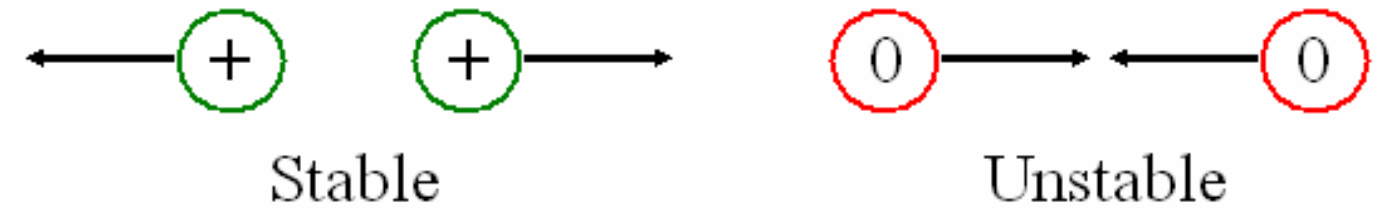
# Biomolecular Stability

*With regard to aggregation:*

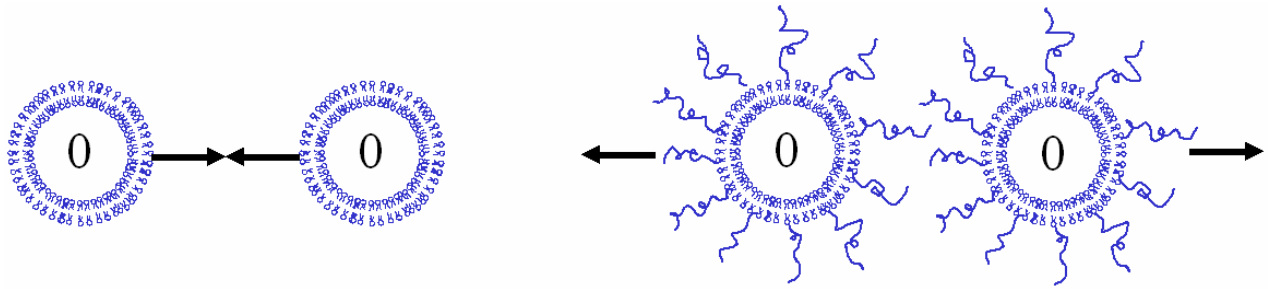
- In the absence of stabilizing “conditions”, when small particles collide, London forces can dominate the interaction, leading to particle aggregation.
- In order to stabilize a formulation against aggregation, particle collisions must be minimized.  This can be accomplished using:
- Electrostatic Effects - 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 chaotropic 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.

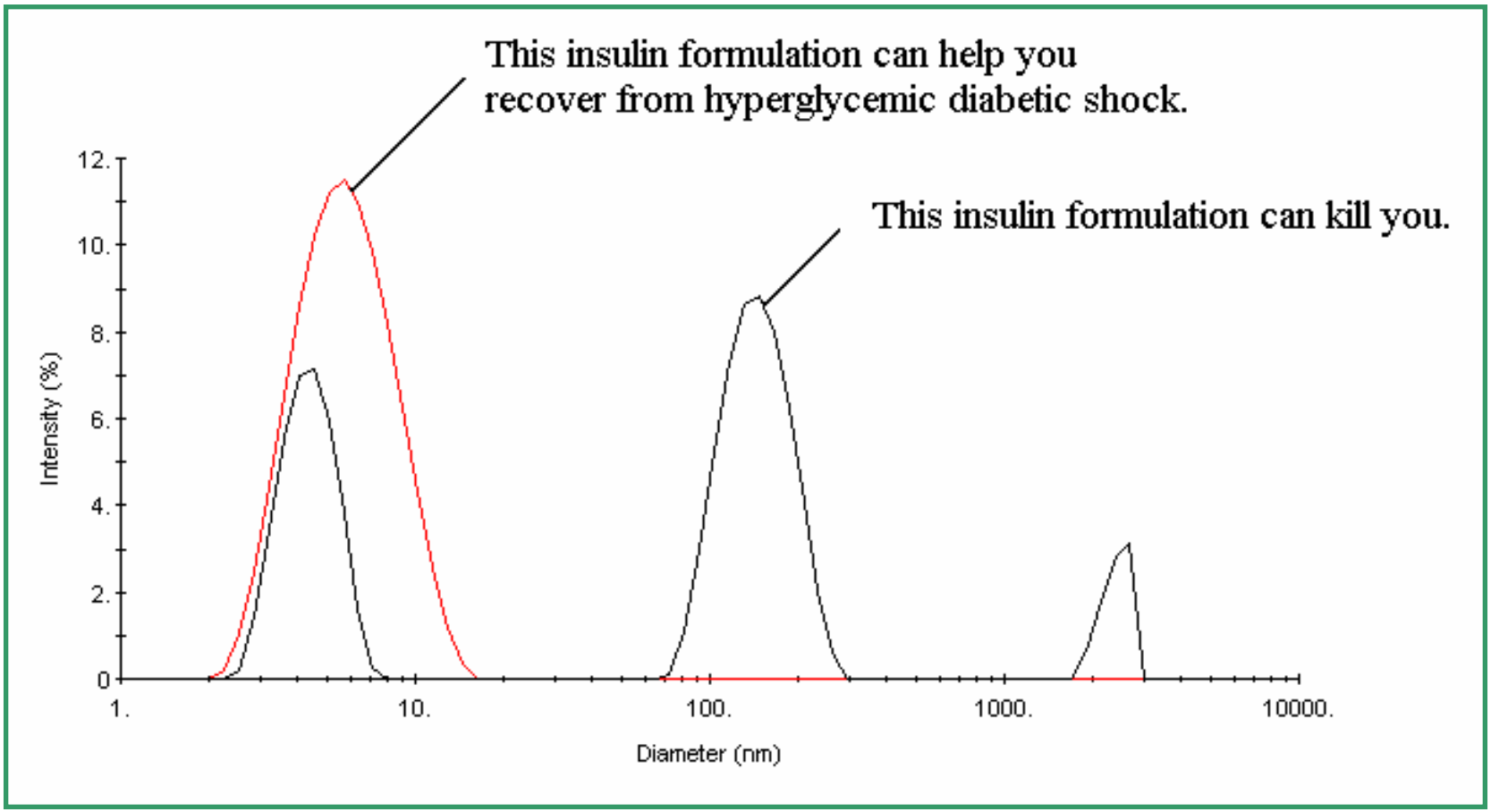



- In the absence of electrostatic effects, steric stability is enhanced upon addition of chaotropic or “structure disrupting” agents, which reduce the likelihood of particles getting close enough for London forces to take over.



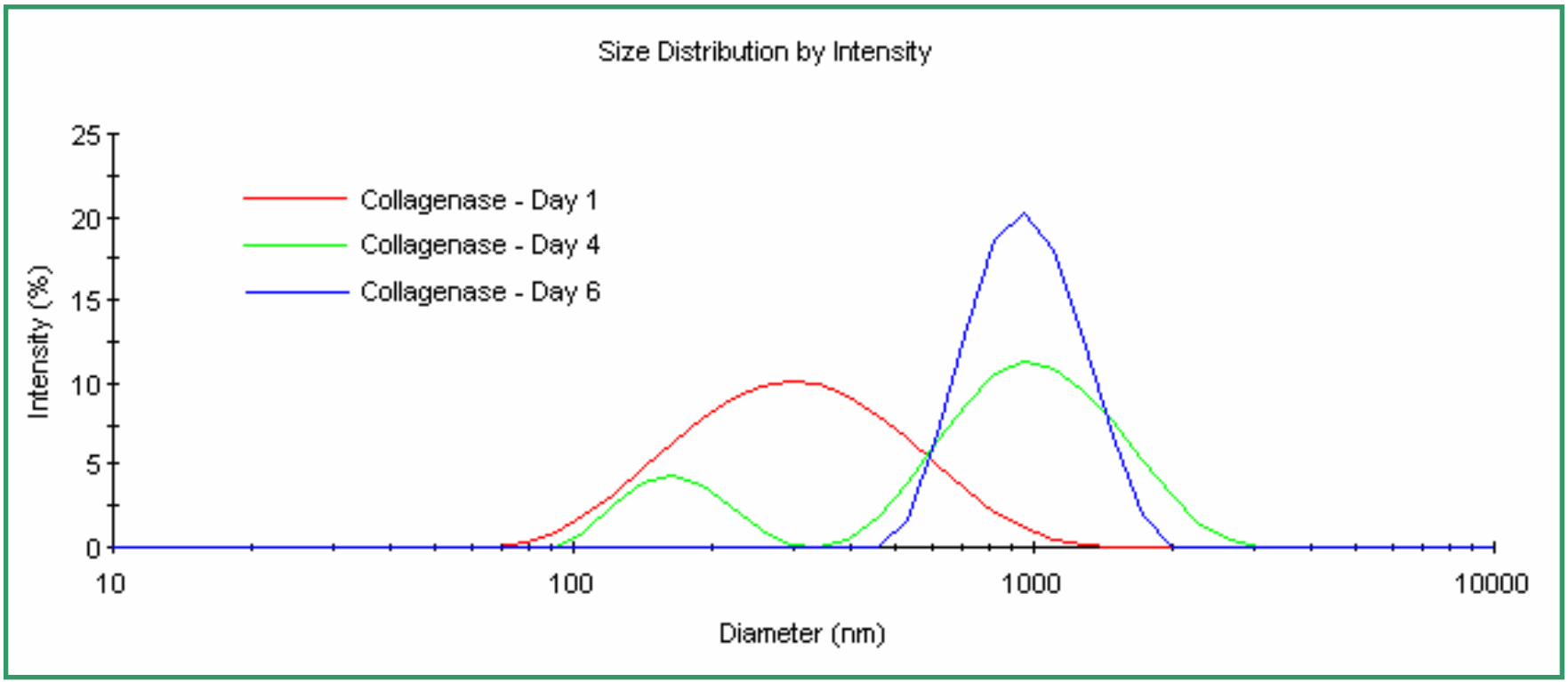
# Evidence Of Instability?

Insulin formulations at t = 0 and 12 months



# Common Approach – Time Trials

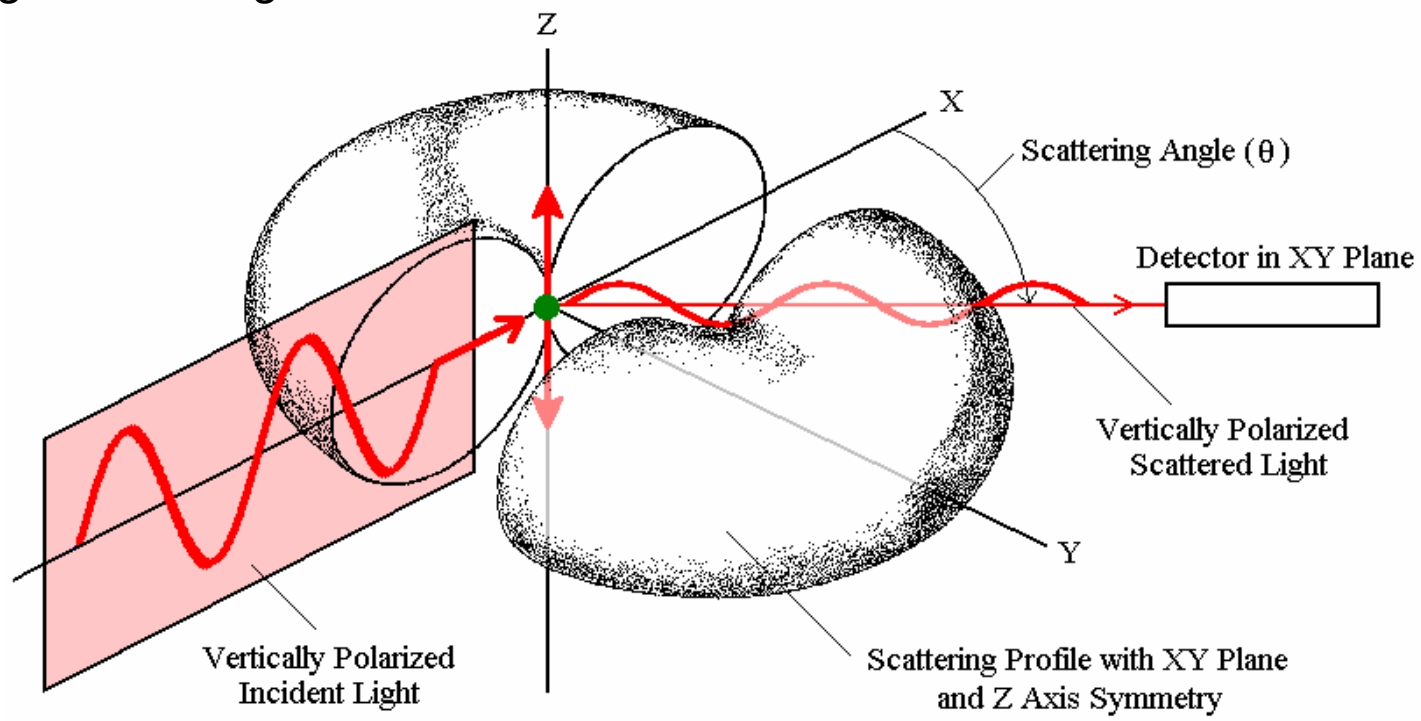
According to sizing results, the sample is completely aggregated within 7 days of preparation.



# But Can We Predict?

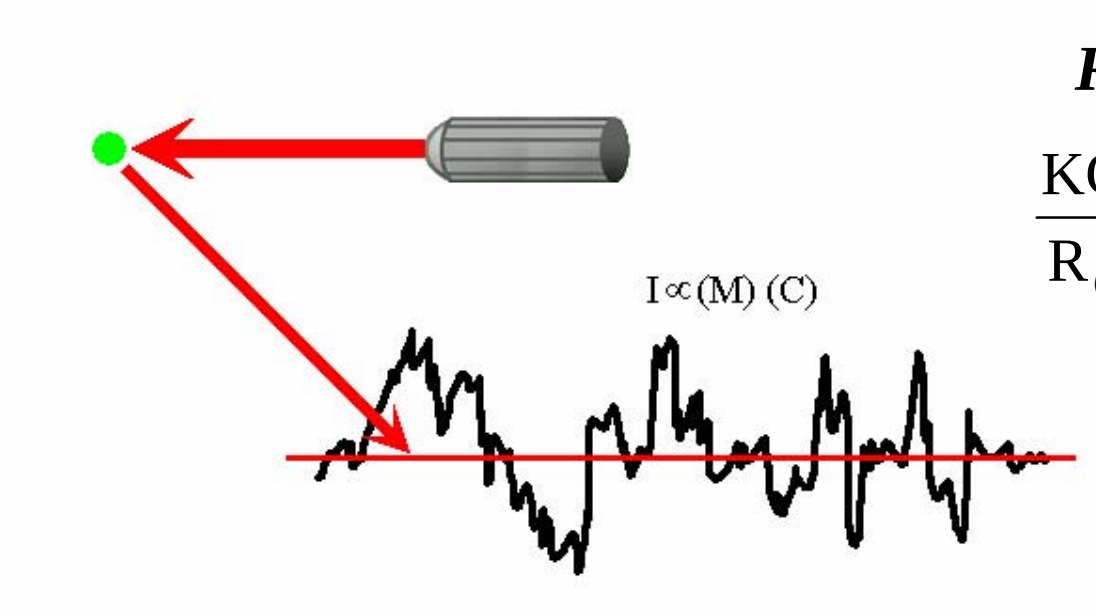
**Light Scattering:** Low energy photon induces an oscillating dipole in the electron cloud. As the dipole oscillates, energy is radiated in all directions. This radiated energy is called “scattered light.”

*Rayleigh Scattering Profile*



# Static Light Scattering (SLS)

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




***Rayleigh Equation***

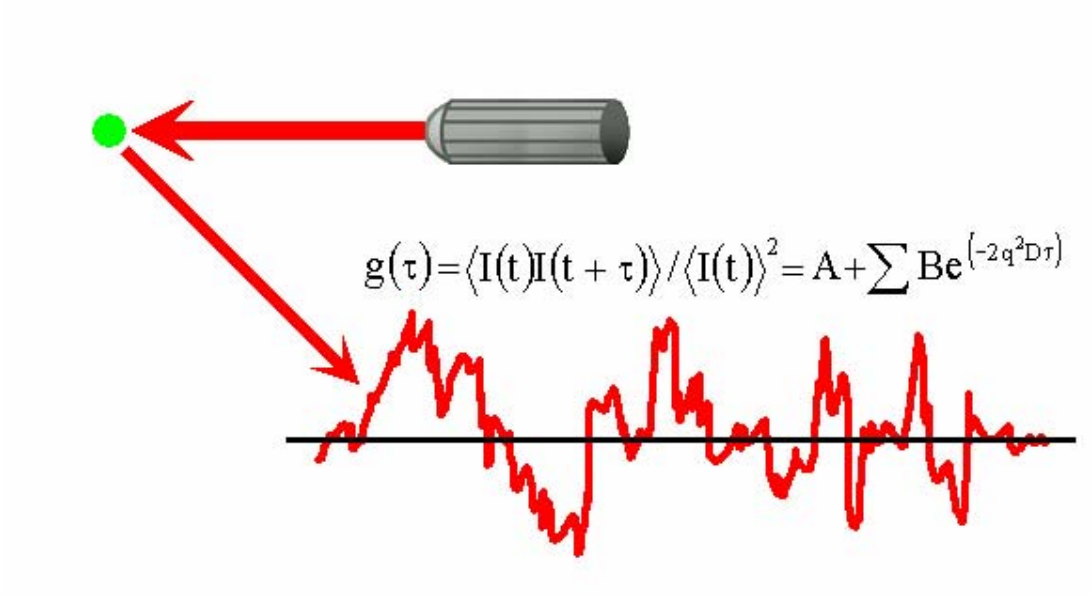
$$\frac{KC}{R_\theta} = \left(\frac{1}{M} + 2A_2C\right)P(\theta)$$

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 (μs) intensity fluctuations gives the diffusion coefficient, hydrodynamic size, polydispersity, and particle size distribution.




***Stokes-Einstein***

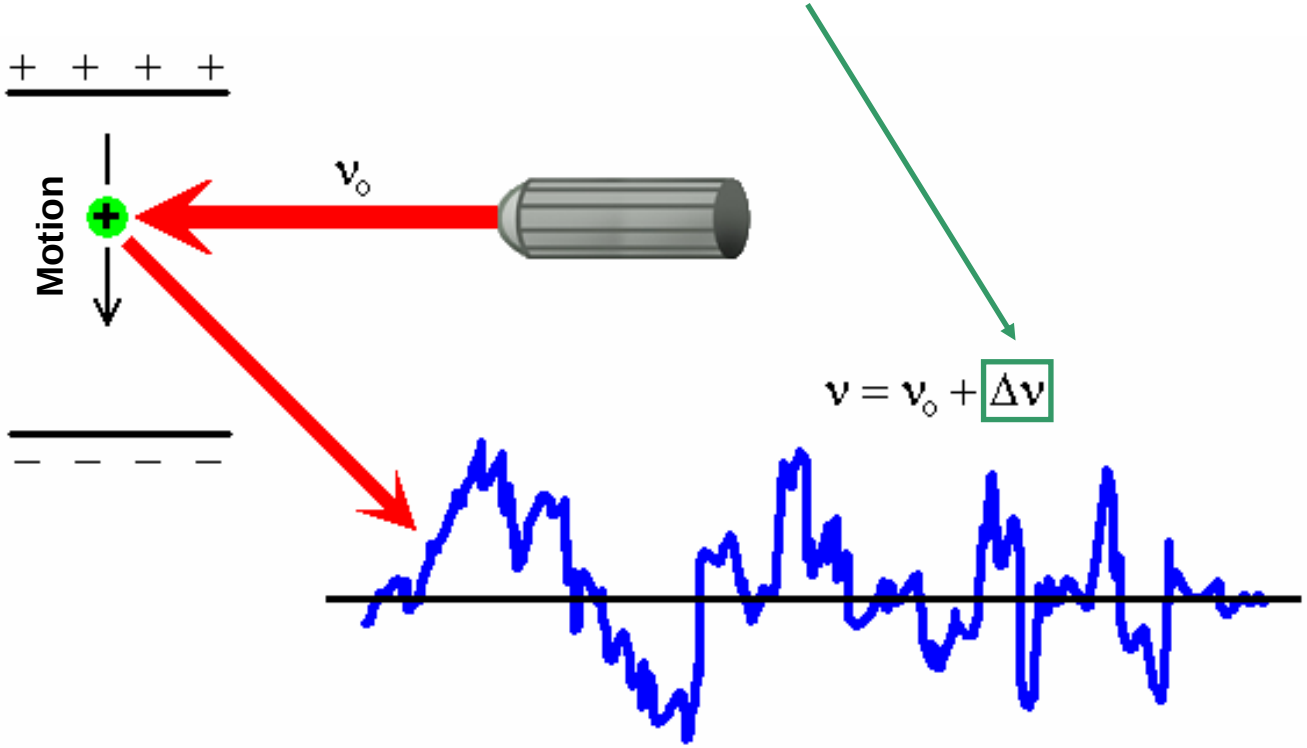
$$R_H = \frac{kT}{6\pi\eta D}$$

q = Scattering vector
$R_H$ = Radius
T = Temperature

D = Diffusion coefficient
k = Boltzmann constant
η = Solvent viscosity

# Electrophoretic Light Scattering (ELS)

Measured parameter is the frequency shift of the scattered light.



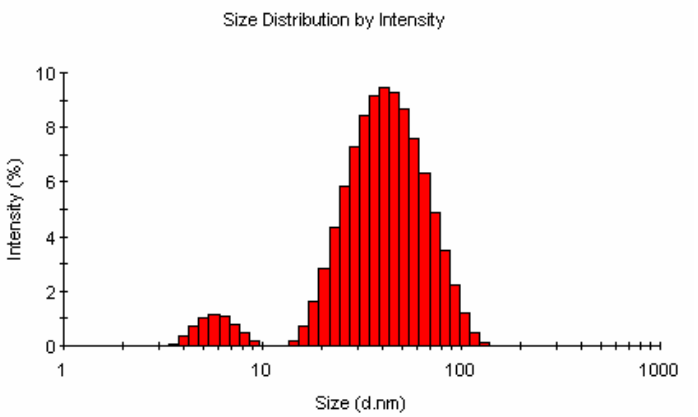

The frequency shift is proportional to the electrophoretic mobility, which is a function of the particle surface potential. Hence ELS gives us information regarding the charge on the particle.

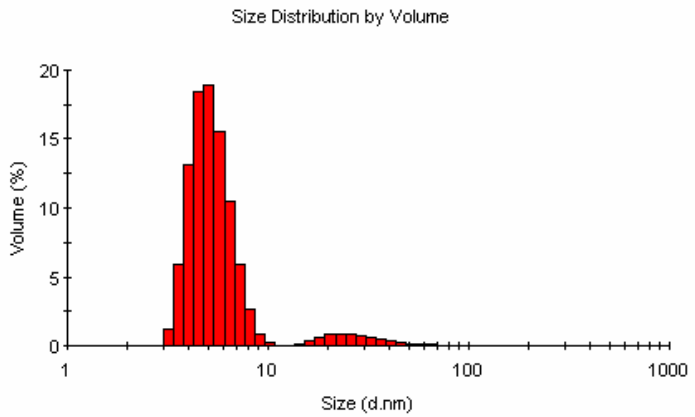
$$\mu = K\left(\frac{\Delta\upsilon}{E}\right)$$

# Why Light Scattering?

The scattering intensity:

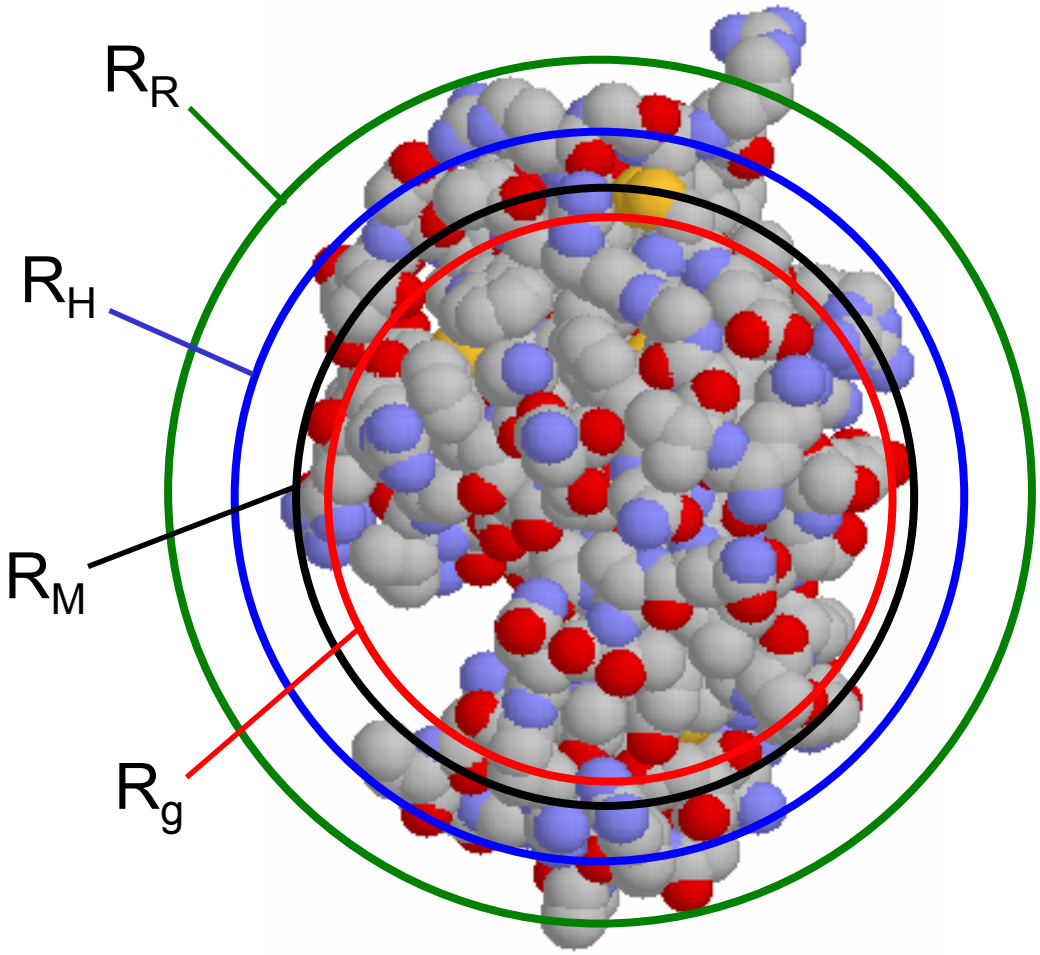
- varies with the mass and concentration according to the Rayleigh Expression
- is proportional to
  - $M_w$
  - $M_n^2$
  - $R^6$
- is non-invasive
- is ideal for aggregate detection & quantification in low volume, low concentration biological samples.




| Peak | $D_I$ (nm) | %I | %M |
|---|---|---|---|
| 1 | 5.95 | 5 | 93 |
| 2 | 46.0 | 94 | 6 |

# Lysozyme - Comparison Of Radii




## Lysozyme

M = 14.5 kDa

$V_p$ = 0.73 mL/g

$R_R$ = 2.25 nm

$R_H$ = 1.90 nm

$R_M$ = 1.61 nm

$R_g$ = 1.47 nm

| For | $R_g$ |
|---|---|
| Sphere | = 0.774 $R_H$ |
| Coil | = 0.816 $R_H$ |
| Cylinder | = 1.732 $R_H$ |

# 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

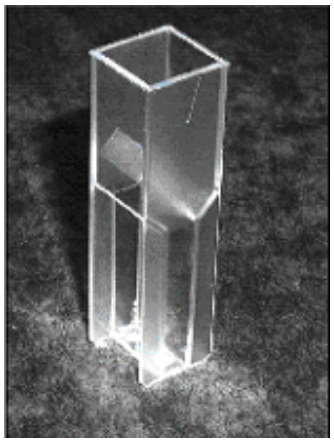
$M_i$ = the molecular weight of each weight class

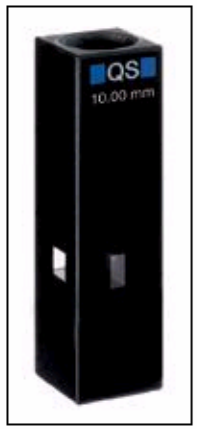
$m_i$ = the mass of particles in each weight class

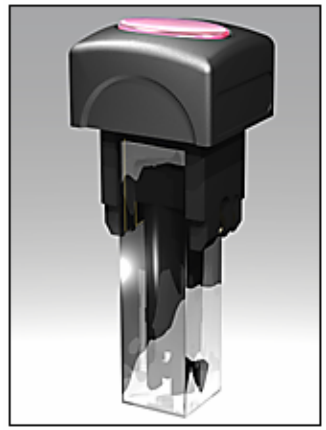
*As a consequence, $M_W$ is more heavily weighted by larger particles in the sample.*

# Light Scattering Applications

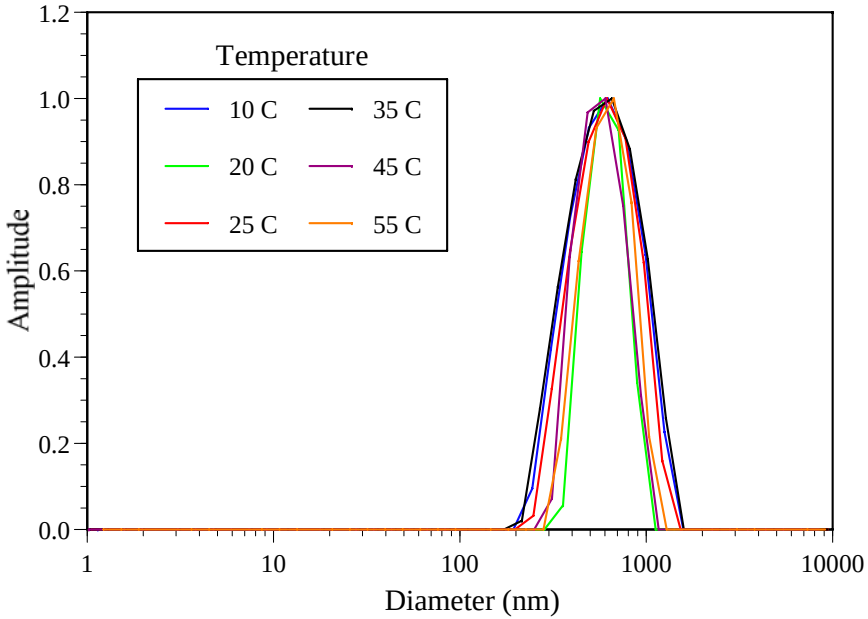
## Batch Mode Measurements

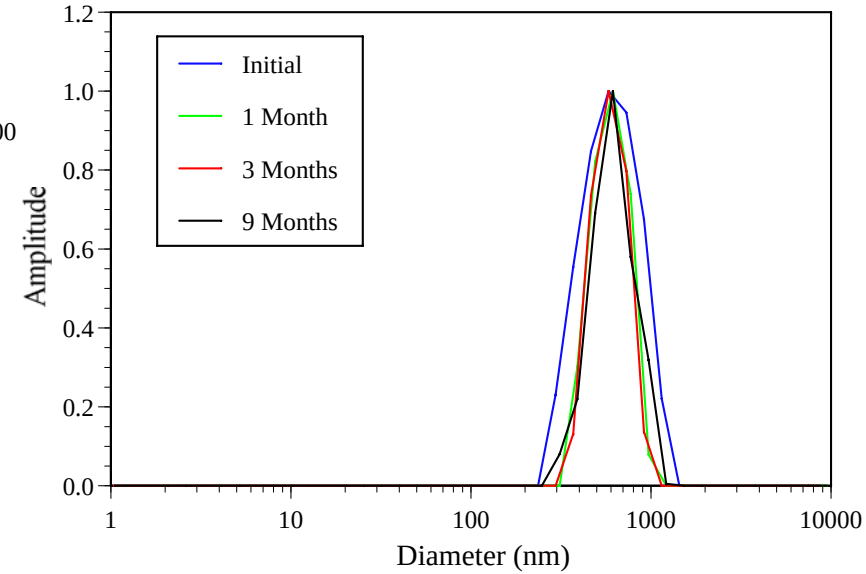


# Monitoring Stability

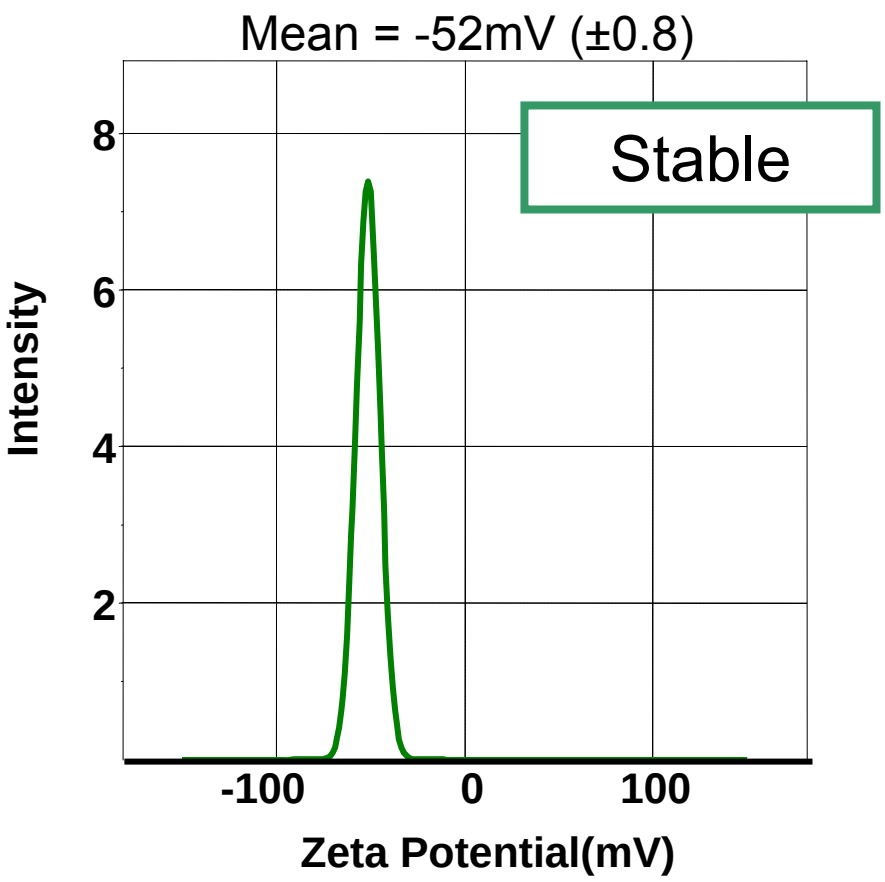



40% Propofol emulsion

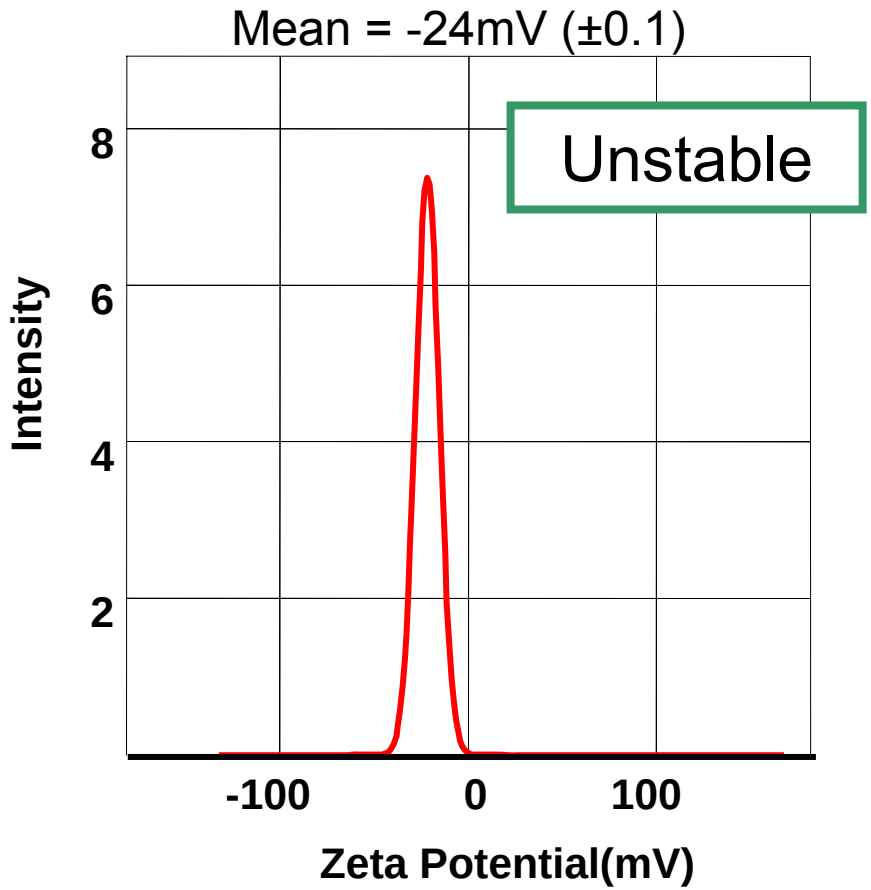


# Predicting Stability

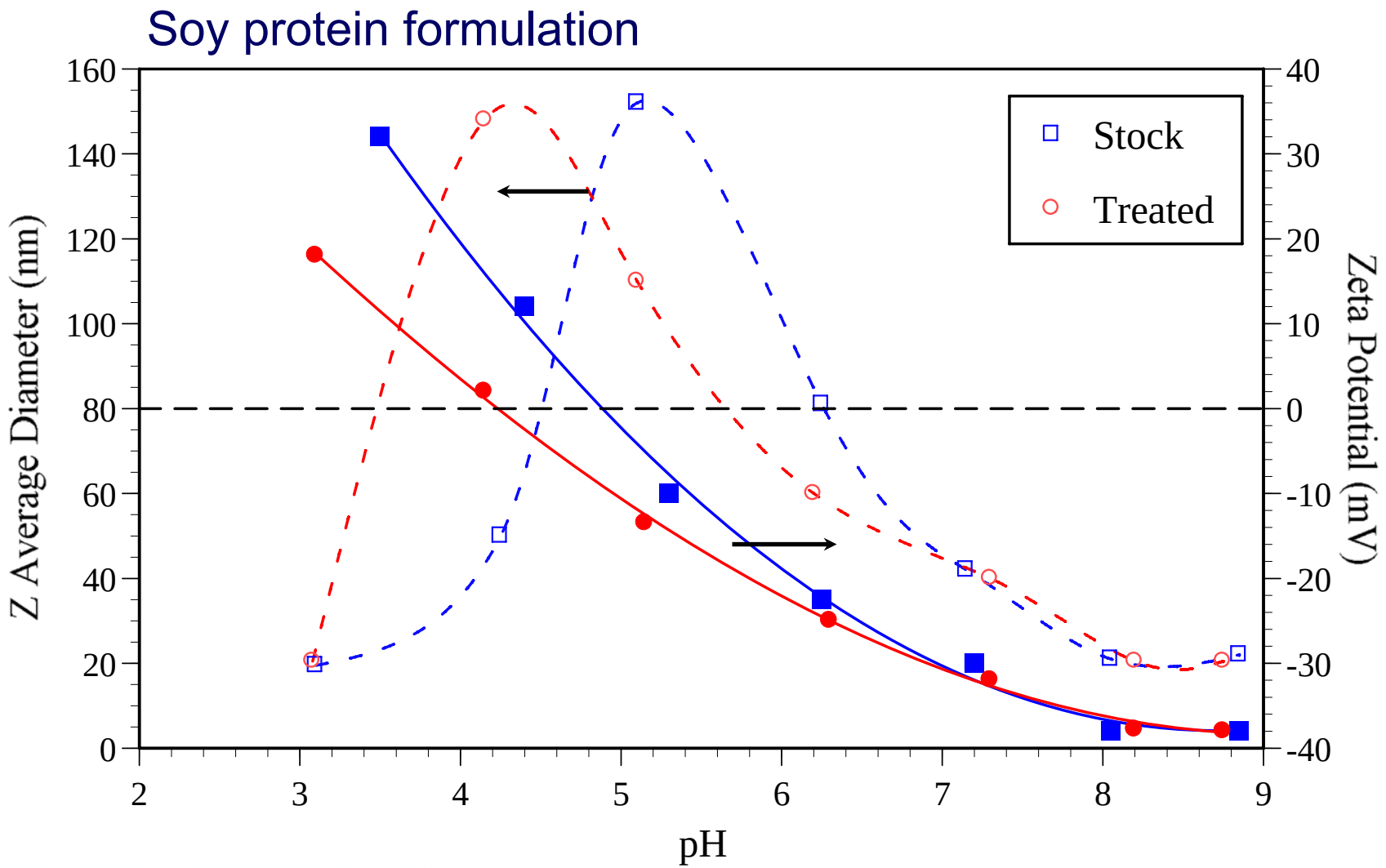
## Formulation A




## Formulation B

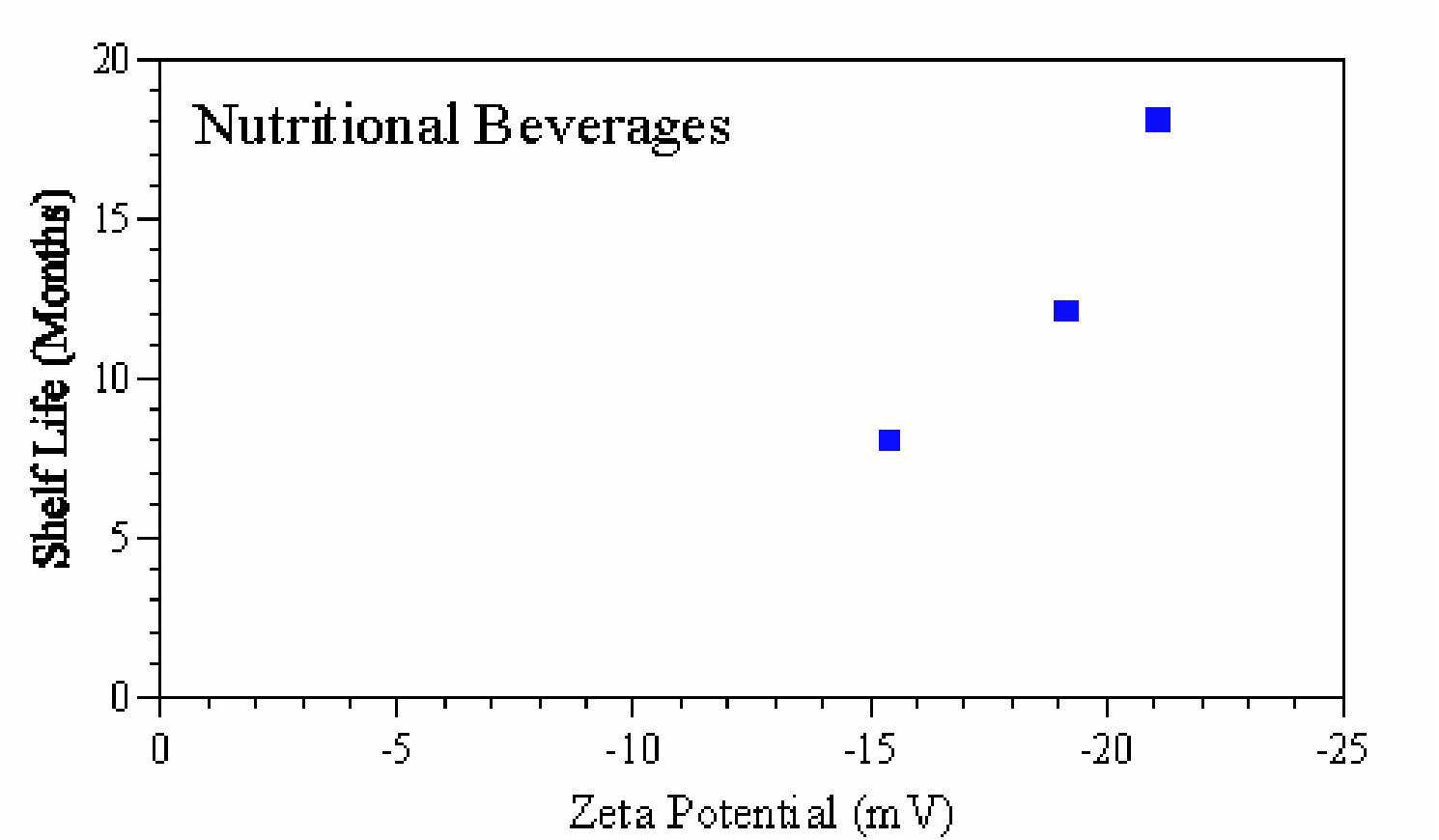


# Prediction & Observation



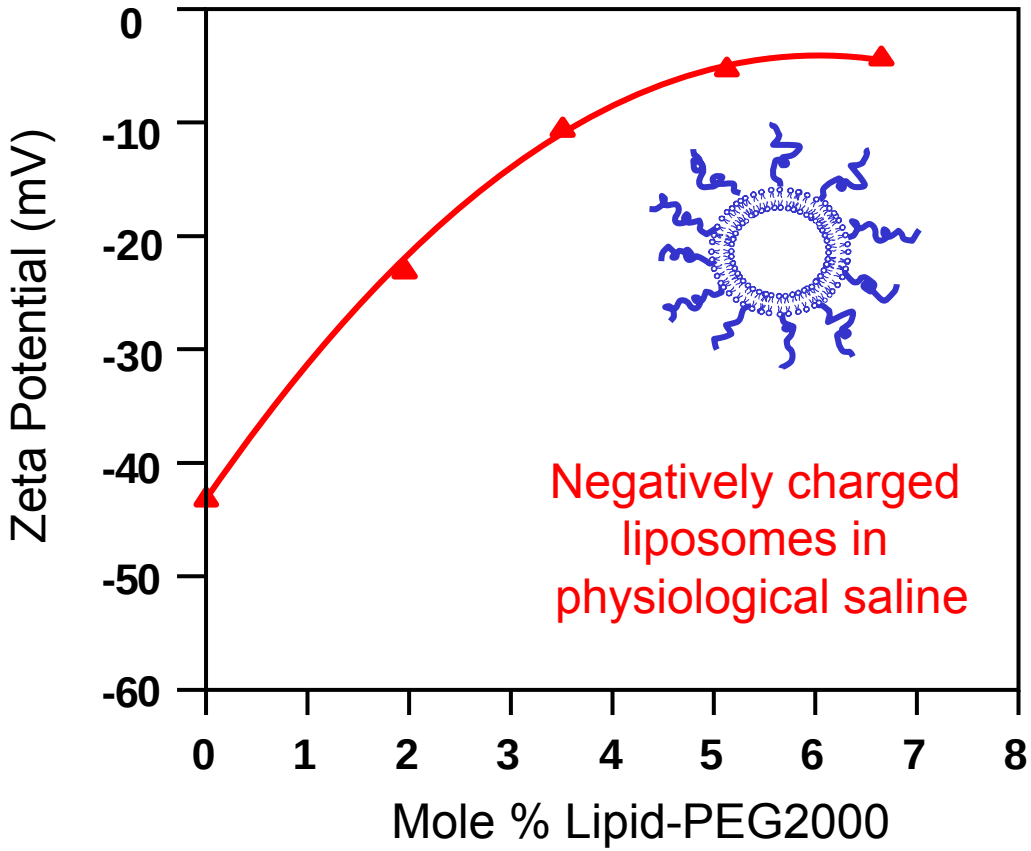
# Zeta Potential - Shelf Life Correlation

Formulations with varying soy fiber content.



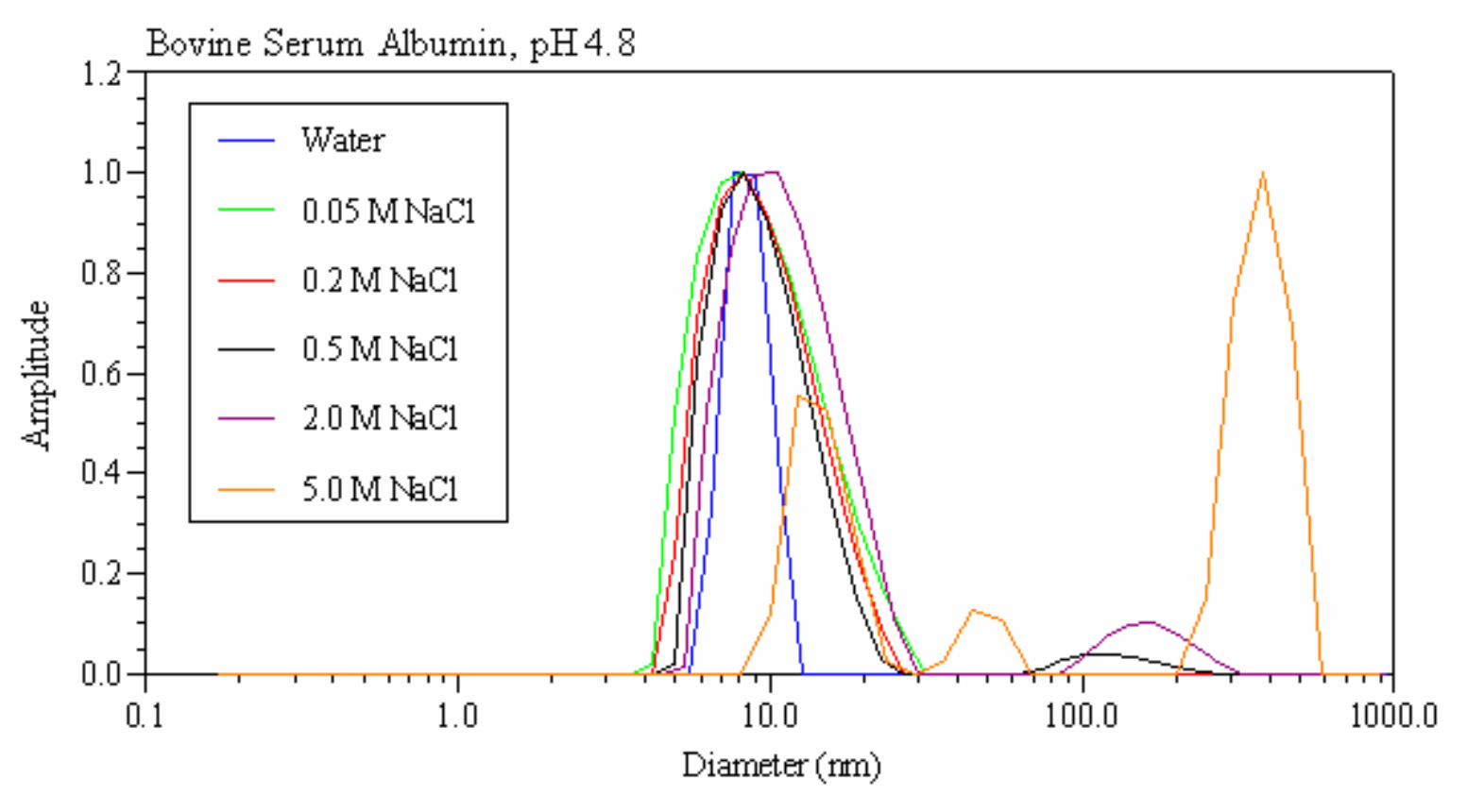
# Electrostatic vs. Steric Stability

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



# Salt Effects On Aggregation

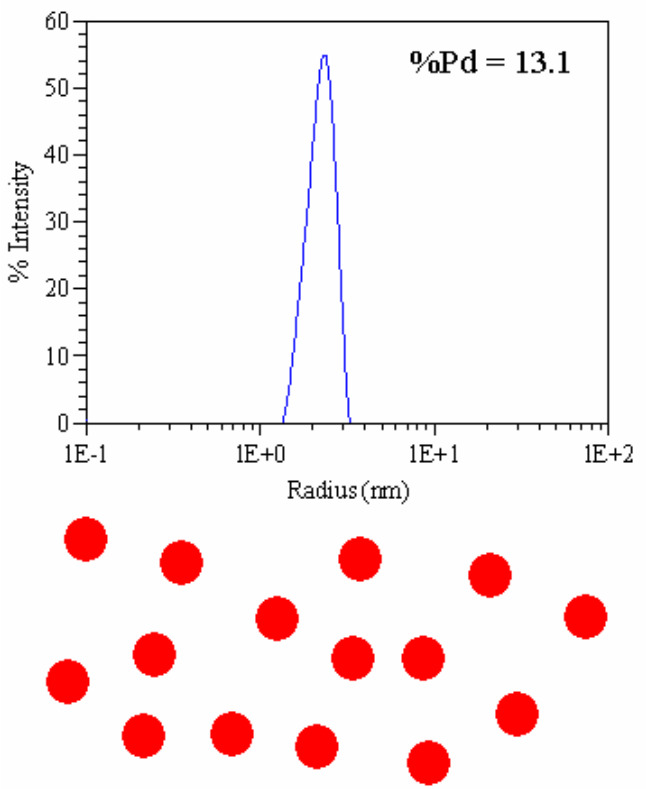
Electrostatic shielding enhances BSA aggregation for NaCl concentrations ≥ 500 mM at the isoionic point. But the aggregation is reversible, suggesting that it is “non-denaturing”.



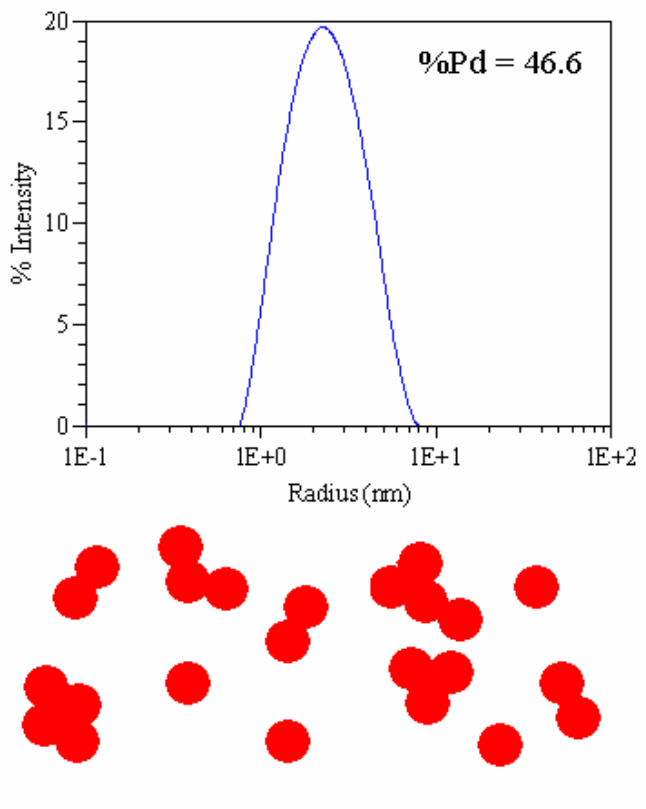
# Polydispersity (Pd) From DLS

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

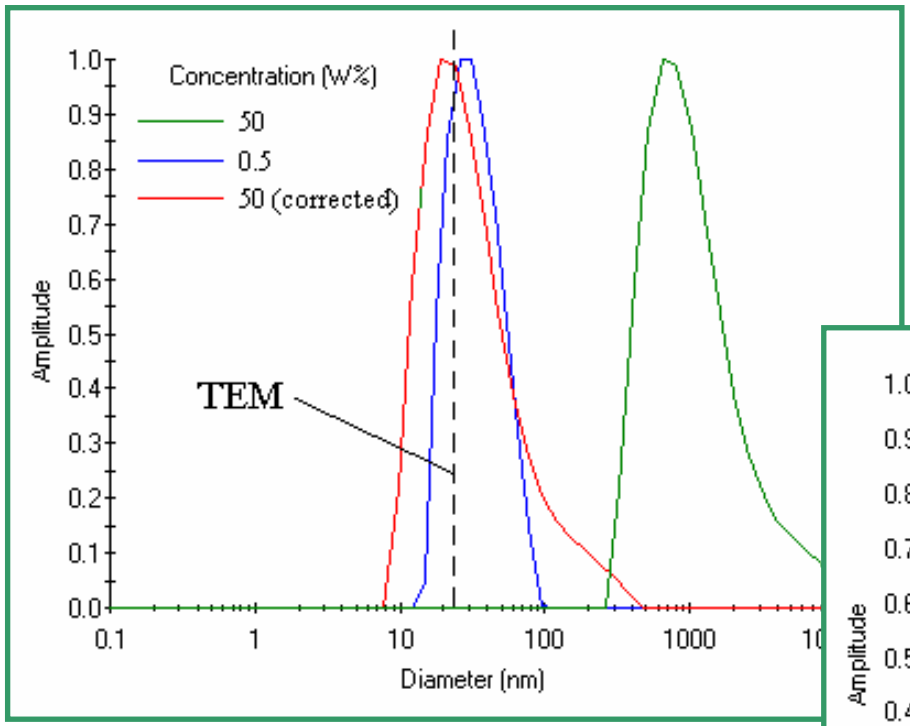
***Monodisperse***




***Polydisperse***



# High Concentration Sizing

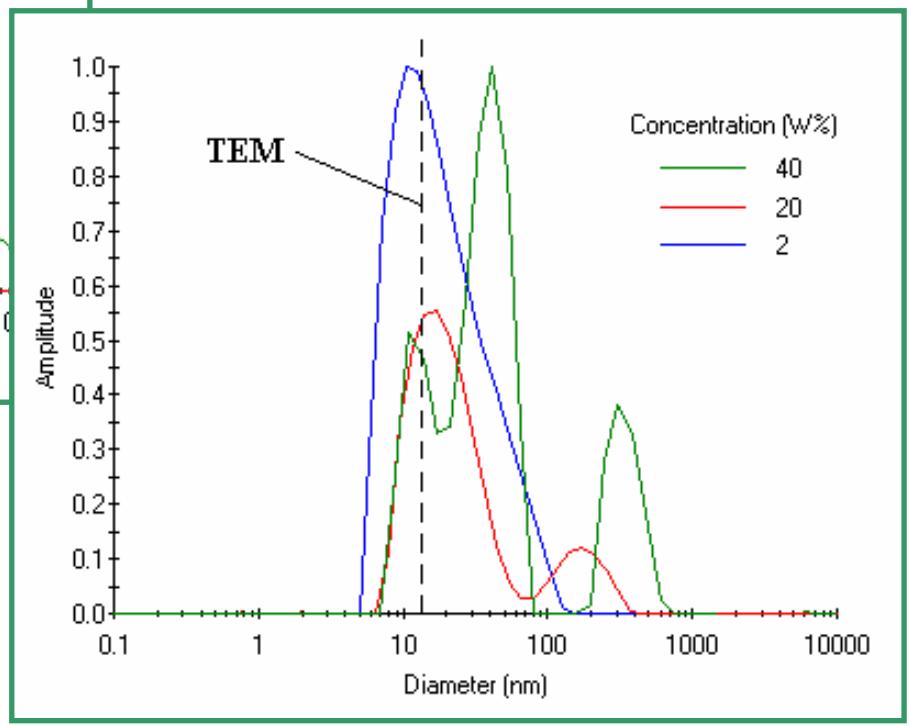



***Restricted Diffusion***

*Stable – Long Shelf Life*

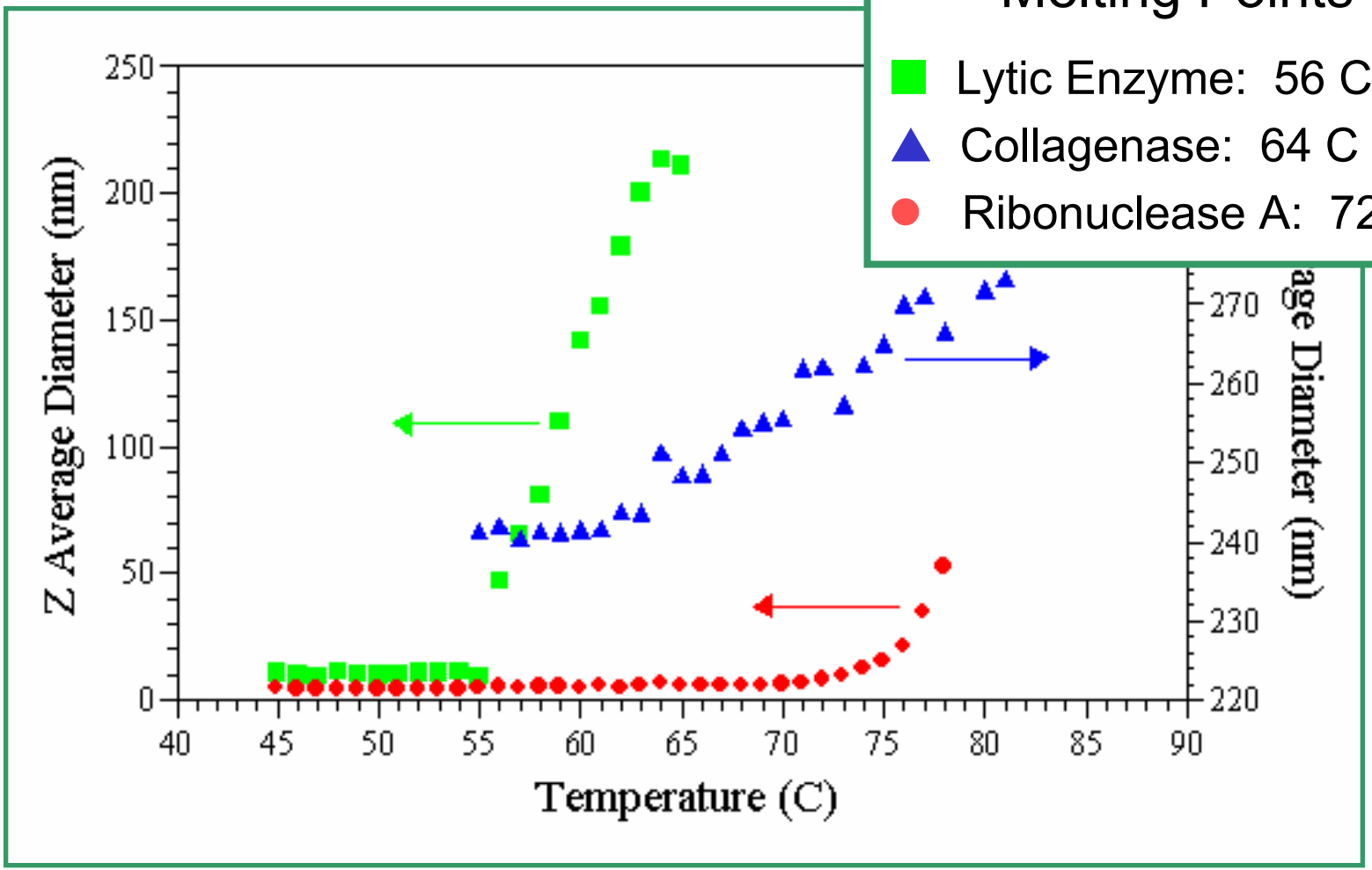
***Equilibrium Aggregation***

*Unstable – Short Shelf Life*



# Using $T_M$ As A Stability Predictor

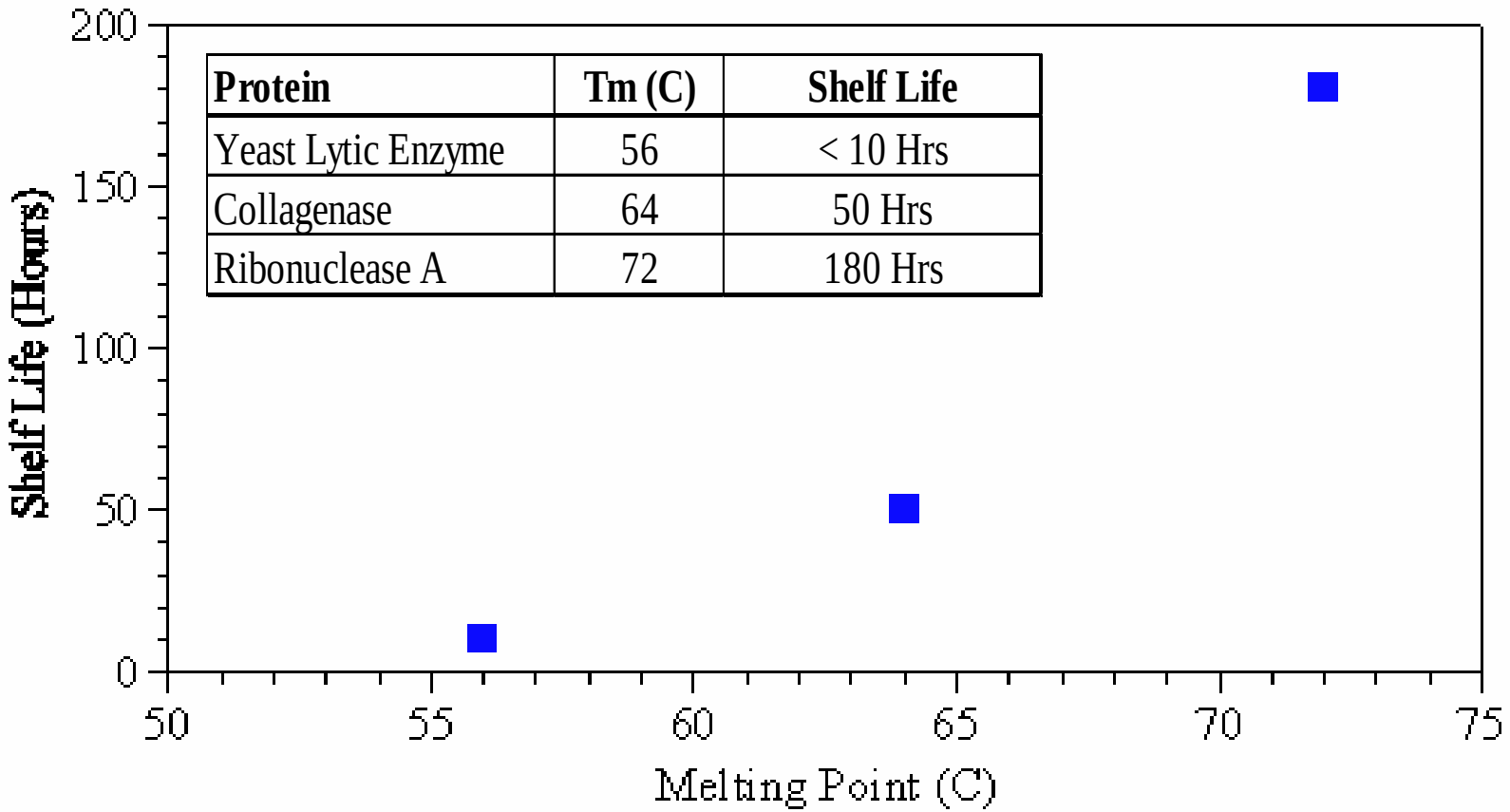
In phosphate buffered saline



# Melting Point - Shelf Life Correlation

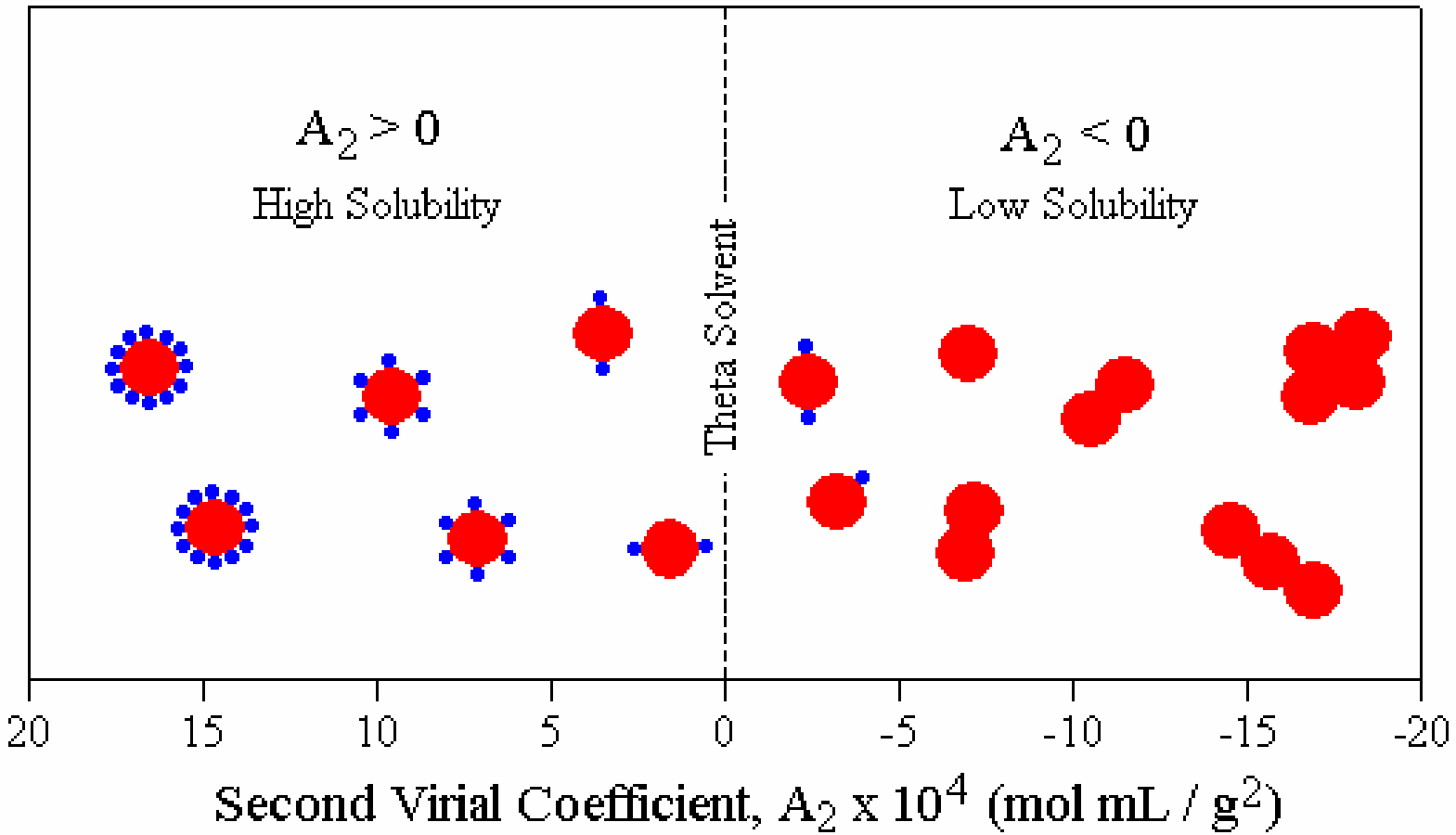
Samples prepared in phosphate buffered saline at 1.0 mg/mL protein concentration.

| Protein | Tm (C) | Shelf Life |
|---|---|---|
| Yeast Lytic Enzyme | 56 | < 10 Hrs |
| Collagenase | 64 | 50 Hrs |
| Ribonuclease A | 72 | 180 Hrs |



# Using $A_2$ As A Stability Predictor

$A_2$ is closely correlated with sample solubility



# 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%)

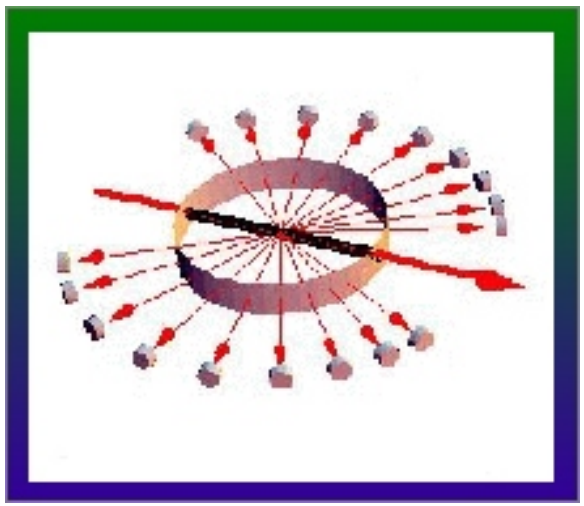
**Missing information: how much and/or what size?**

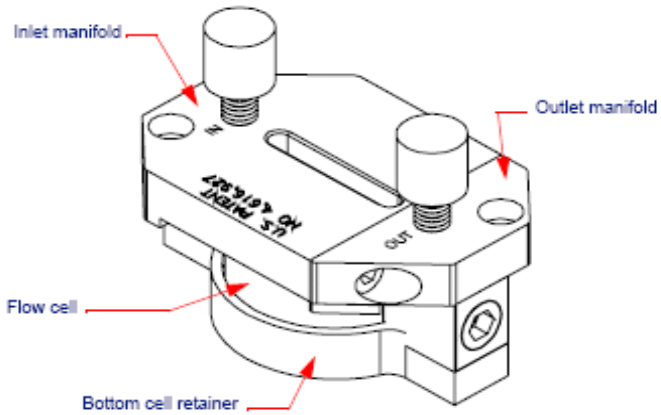
## Solutions

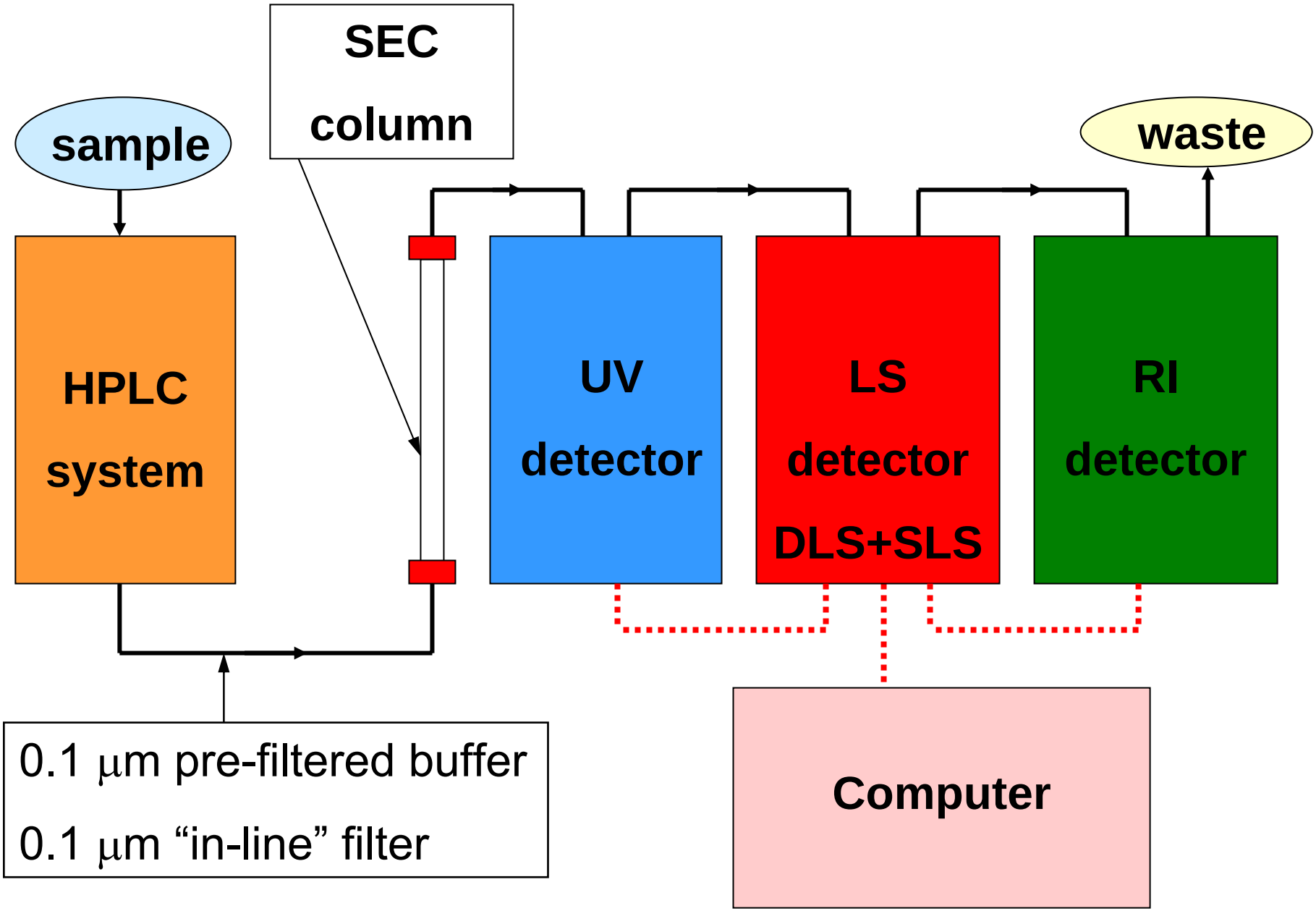
- Sample fractionation followed by batch measurements
- **Column separation with simultaneous LS characterization**

# Light Scattering Applications

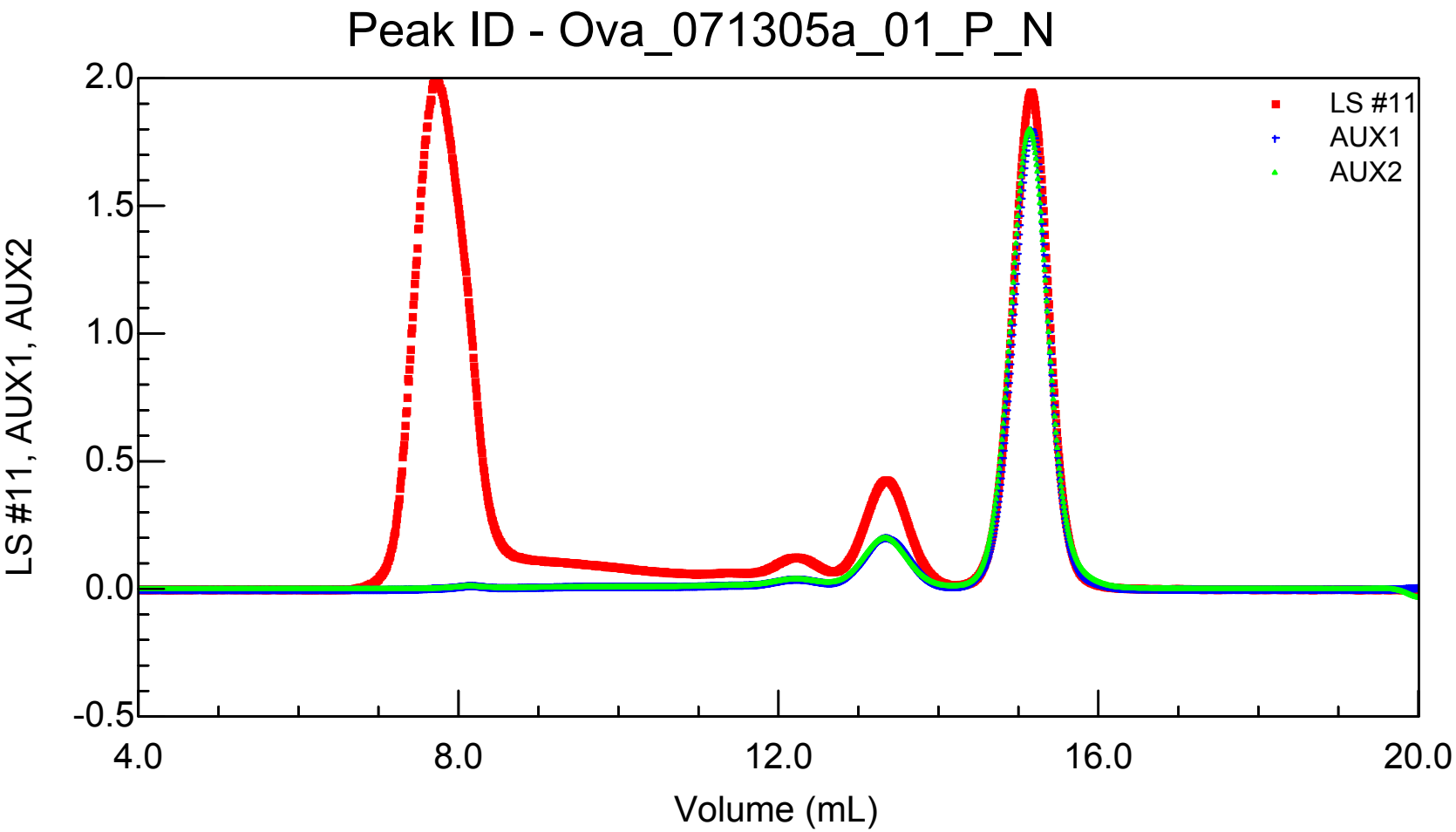
## Flow Mode Measurements







SEC
column
sample
waste
HPLC
system
UV
detector
LS
detector
DLS+SLS
RI
detector
0.1 μm pre-filtered buffer
0.1 μm “in-line” filter
Computer

# Three Detector monitoring




UV at 280 nm  RI  LS at 90°

# Ovalbumin 43 kDa

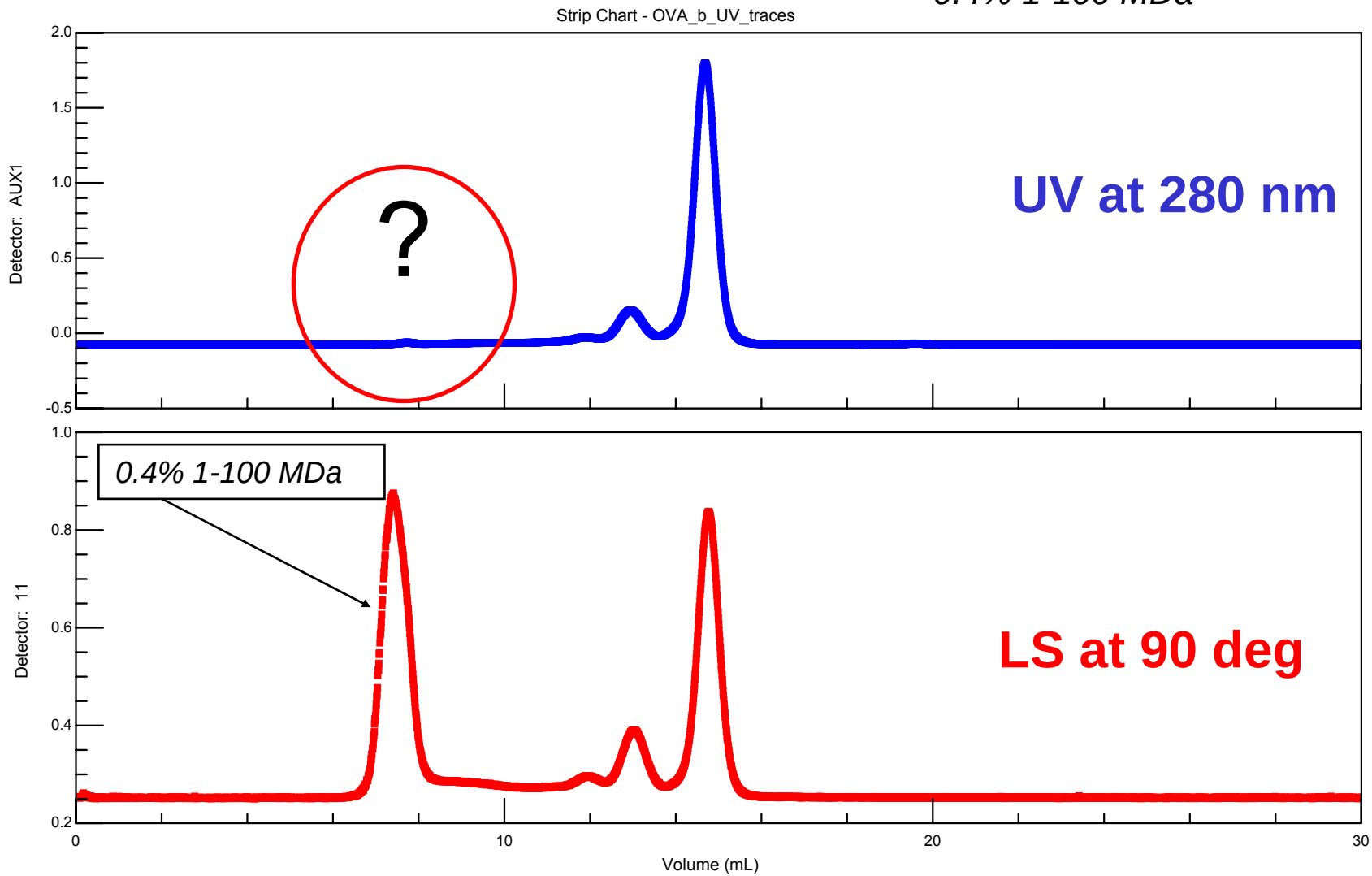
*88% monomer*

*8% dimer*

*1.5% trimer*

*3% aggregates < 1MDa*

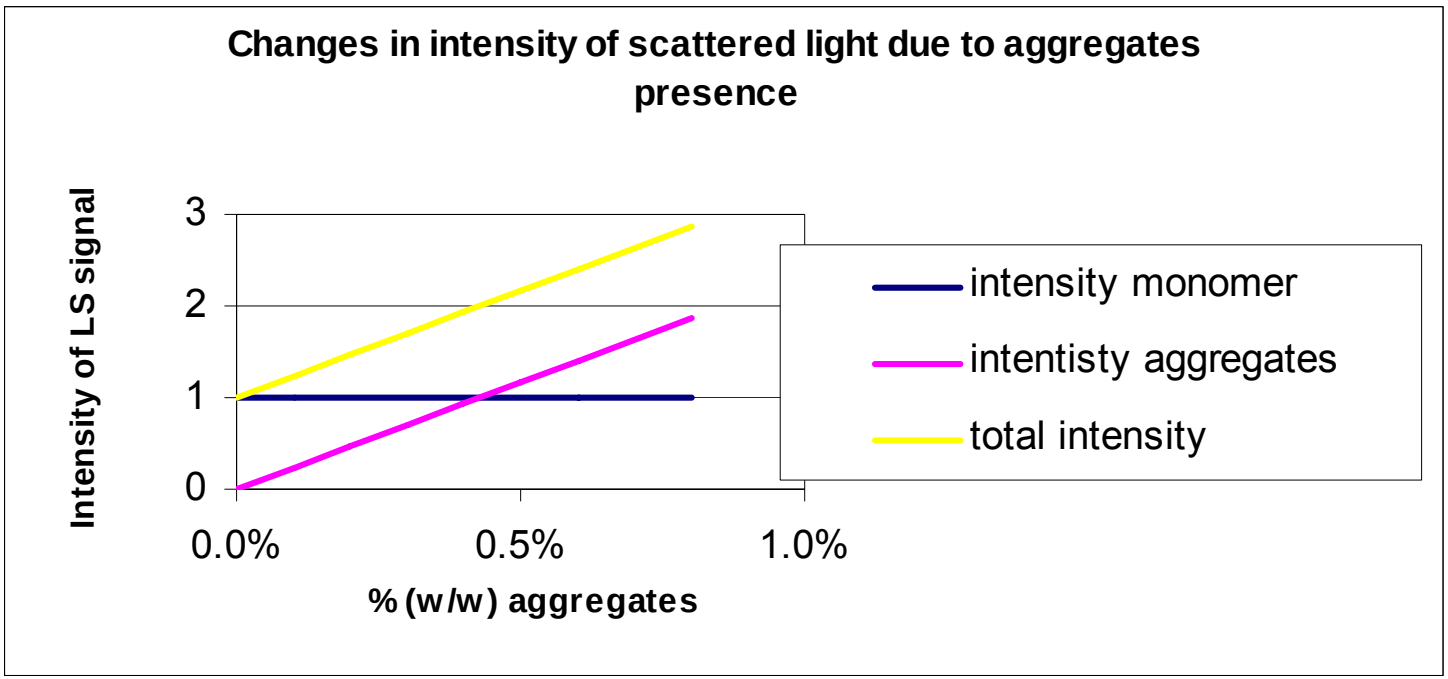
*0.4% 1-100 MDa*



Intensity of scattered light ~ $M_w$*c

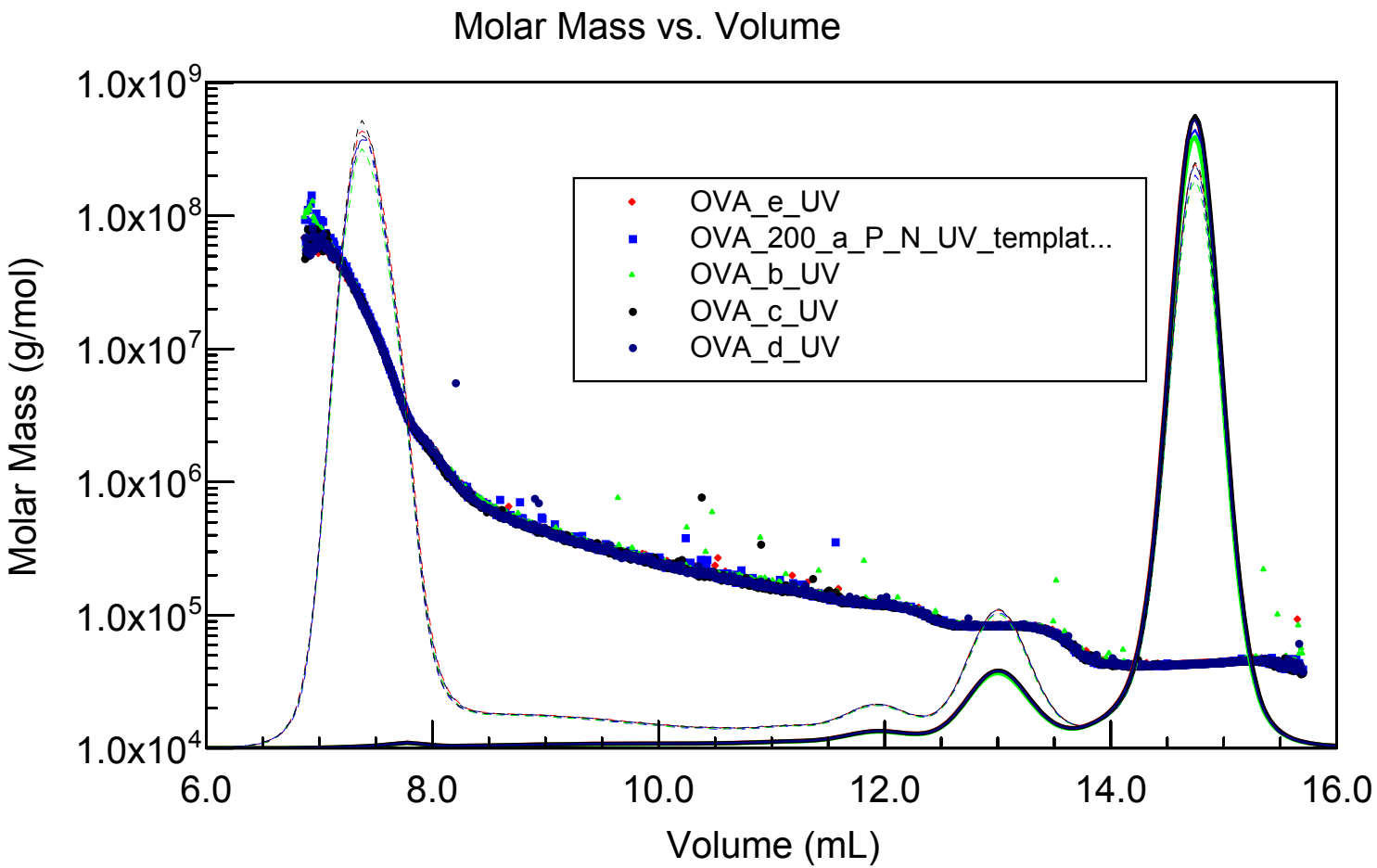
due to their high Mw aggregates scatter very strongly

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

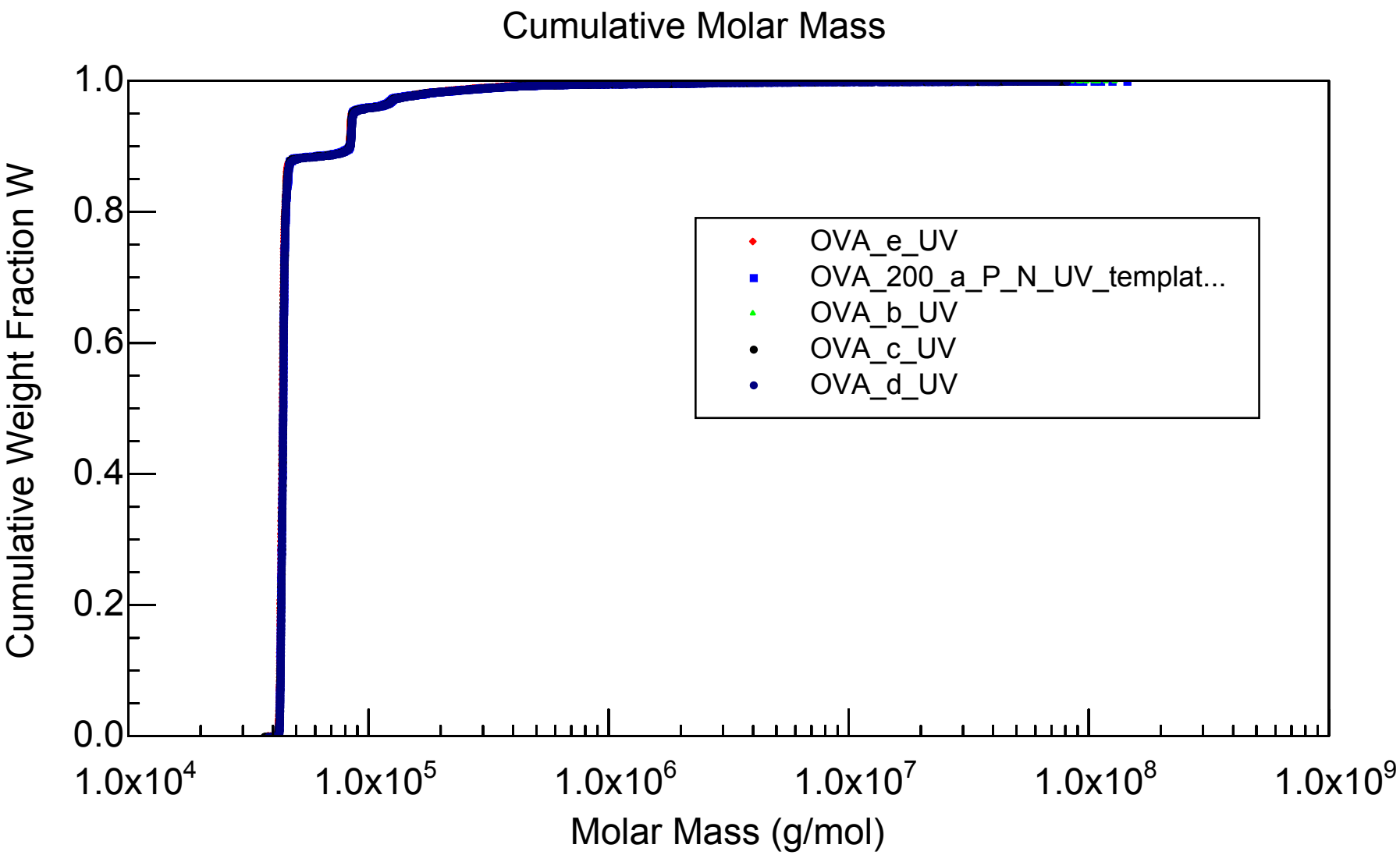


# Molar mass distribution for multiple analyses

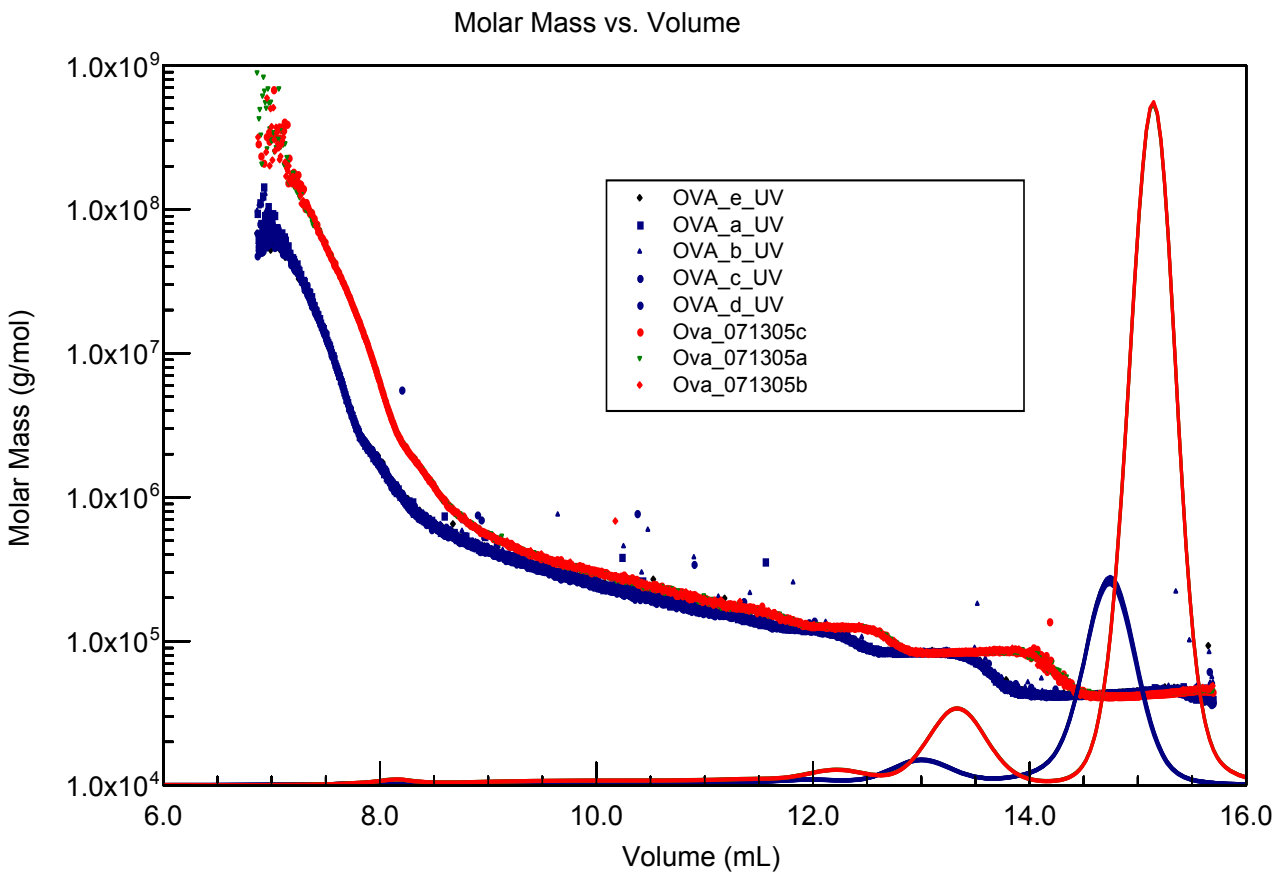
**Ovalbumin 43 kDa** automated template processing of five data sets



# 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)**

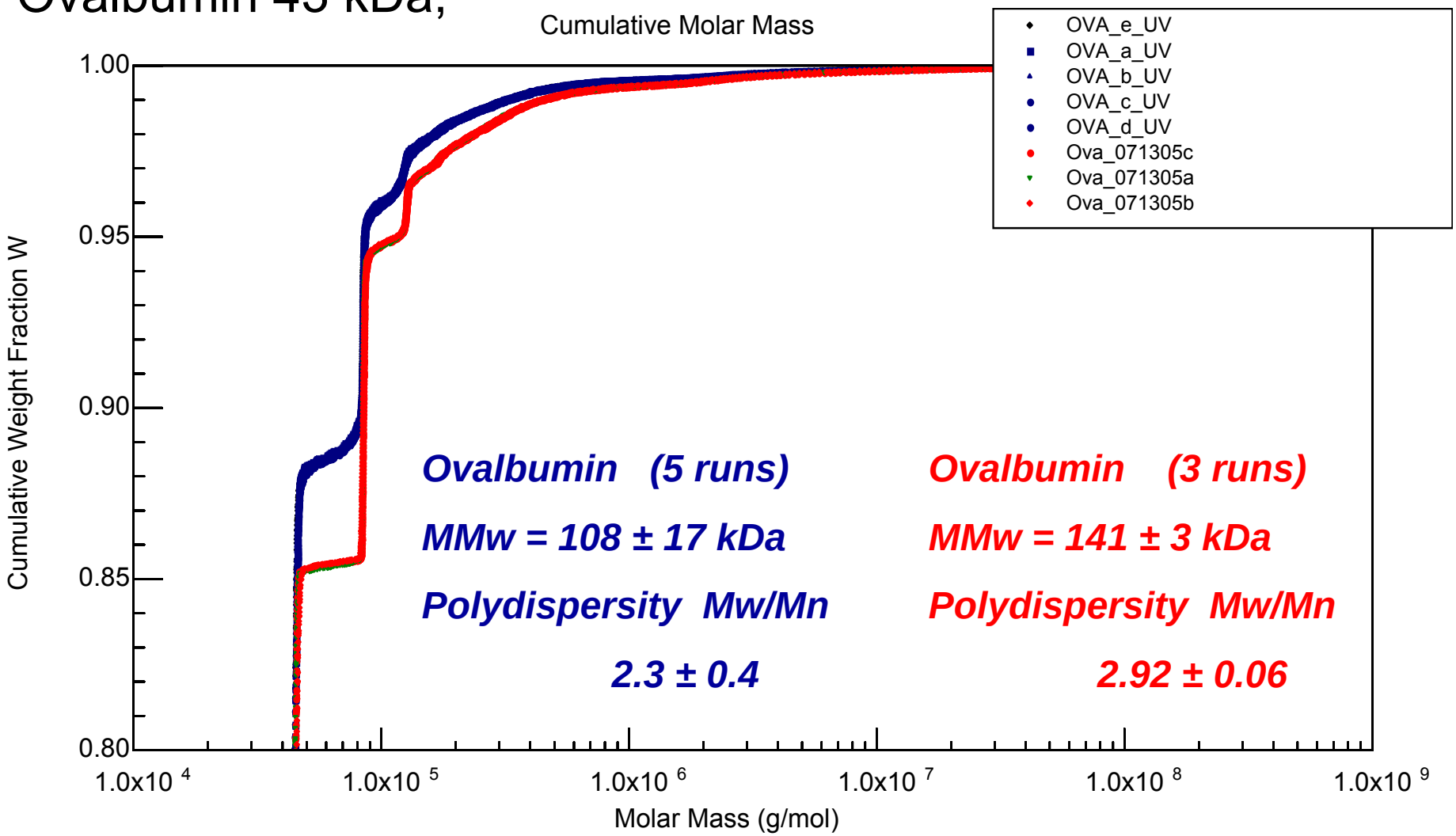
**Mw = 141 ± 3 kDa**

**Polydispersity Mw/Mn**

**2.92 ± 0.06**

# Differences in population based on molar mass distribution

Ovalbumin 43 kDa,



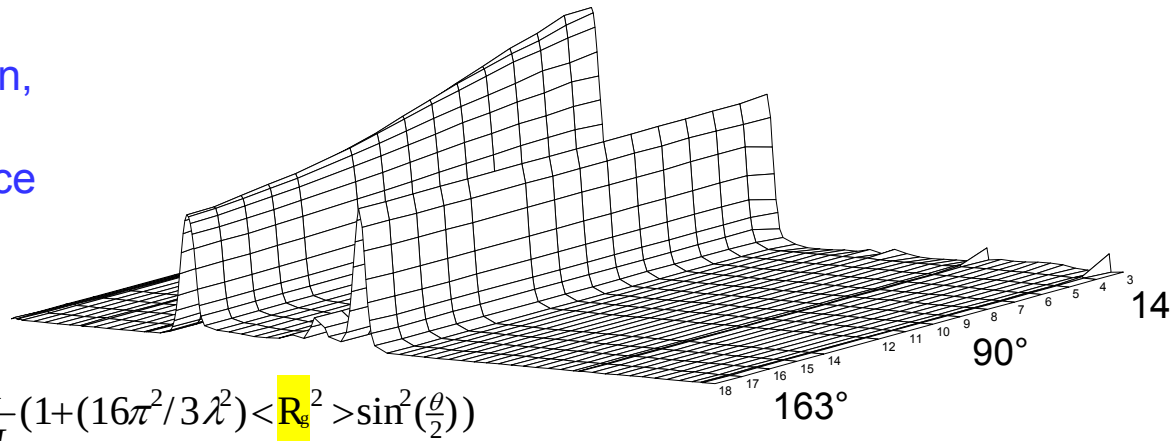
# Differences in population based on molar mass distribution

## Ovalbumin 43 kDa

| Oligomeric state | Average Mw ± SD [kDa] (5 analyses) | Average Mw ± SD [kDa] (3 analyses) | Fraction of Mass [% of total] (5 analyses) | Fraction of Mass [% of total] (3 analyses) |
|---|---|---|---|---|
| | Mw = 108 ± 17 | Mw = 141 ± 3 | Mw = 108 ± 17 | Mw = 141 ± 3 |
| Mono (20-50 kDa) | 43.0 ± 0.1 | 42.80 ± 0.02 | 88.1 ± 0.1 | 85.23 ± 0.06 |
| Di (50-96 kDa) | 82.7 ± 0.4 | 84.1± 0.2 | 7.68 ± 0.04 | 9.4 ± 0.0 |
| Tri (96-130 kDa) | 114 ± 4 | 121.8 ± 0.7 | 1.54 ± 0.05 | 1.9 ± 0.0 |
| Agg. (0.13 –1 MDa) | 270 ±10 | 284 ± 2 | 2.18 ± 0.08 | 2.87± 0.06 |
| Agg. (1 –100 MDa) | $10\pm1 \times 10^3$ | $10.9\pm0.4 \times 10^3$ | 0.4 ± 0.0 | 0.6 ± 0.0 |

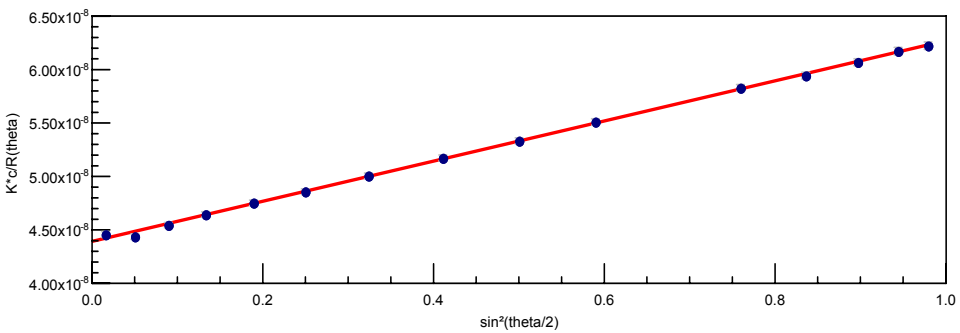
# Morphology of aggregates from angular dependence of LS signal; size determination- Rg

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



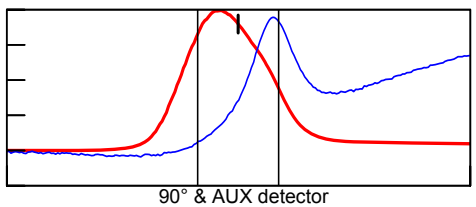

Zimm Plot

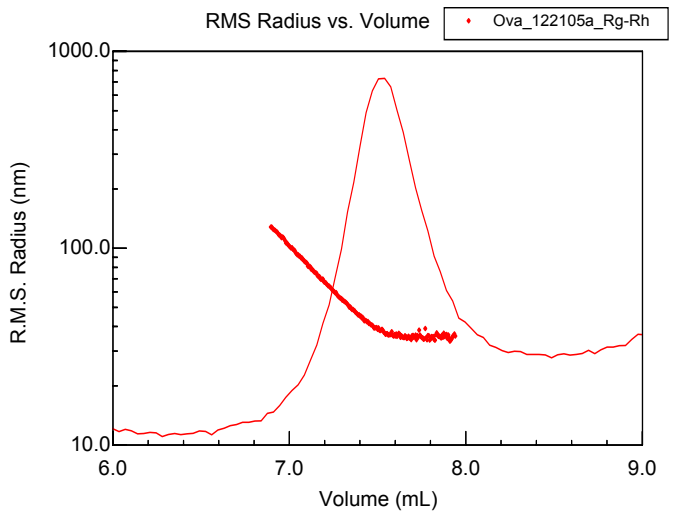
$$\frac{K^*c}{R(\theta)} = \frac{1}{M_w}\left(1+\left(16\pi^2/3\lambda^2\right)<R_g^2>\sin^2\left(\frac{\theta}{2}\right)\right)$$




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

Radius: 46.8±0.2 nm



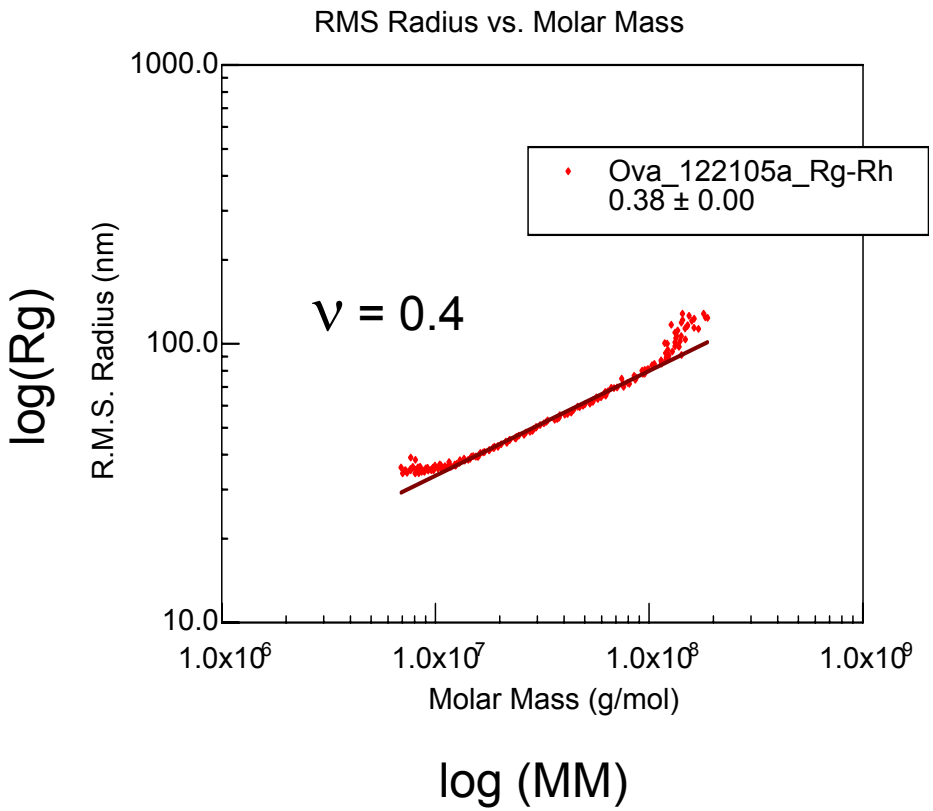
# Inferring conformational information from the relationship between molecular size (Rg) and molecular weight (Molar Mass)

$$R_g \sim M^{\nu}$$

$\log(R_g)$ *versus* $\log(MM)$

Slope = $\nu$

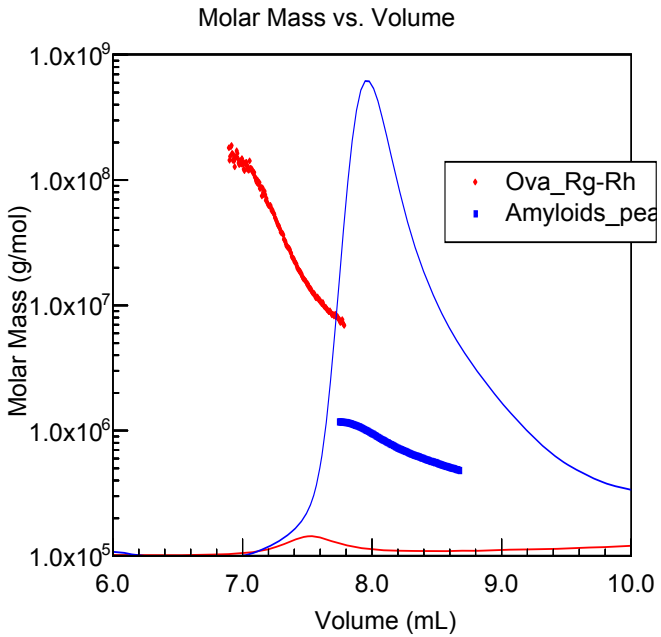
| For | $\nu$ |
|---|---|
| Sphere | 0.33 |
| Coil | 0.5 |
| Rod | 1 |

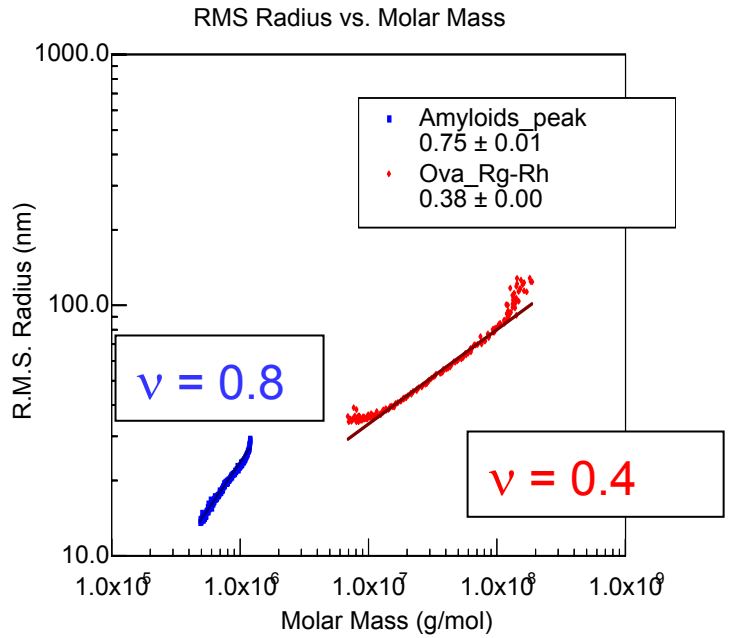



Rollings, J.E. (1992) in *"Laser Light Scattering in Biochemistry"*, Eds. S.E. Harding, D. B. Sattelle and V. A. Bloomfield; p. 275-293

# Shape analysis: **log(Rg) *versus* log(MM)**

Aggregates of Ovalbumin *vs.* "amyloid-type" fibers




| For | ν |
|---|---|
| Sphere | 0.33 |
| Coil | 0.5 |
| Rod | 1 |

Ova_aggr ν = 0.4 Sphere/Coil

Amyloids ν = 0.8 Coil/Rod

# Shape analysis: **shape factor $\rho = R_g/R_h$**

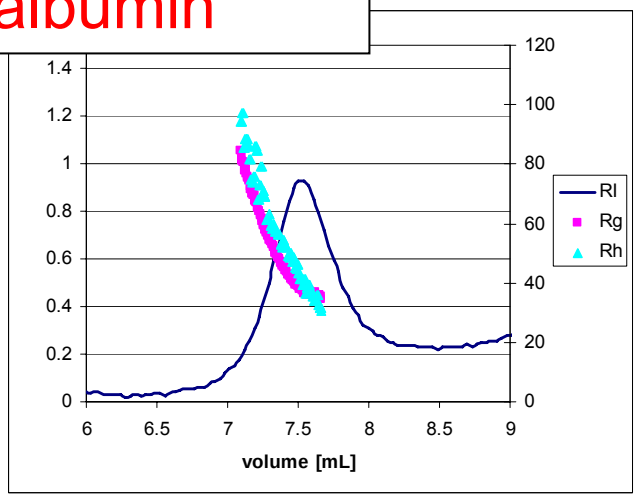
Aggregates of Ovalbumin *vs.* amyloid fibers

Shape factor: $\rho = R_g/R_h$

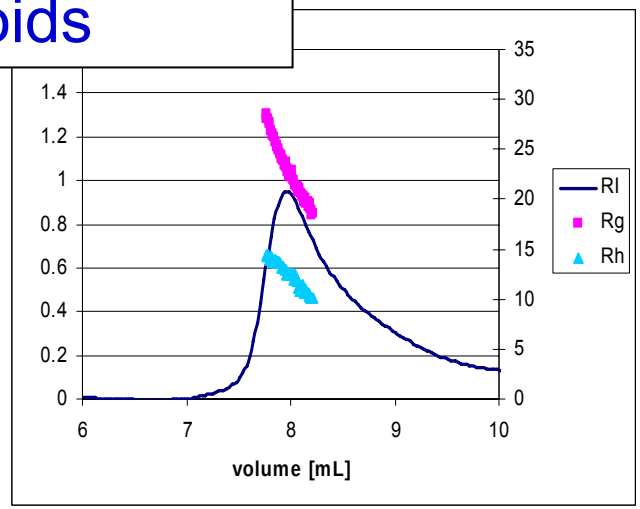
| For | $\rho = R_g/R_h$ |
|---|---|
| Sphere | 0.774 |
| Coil | 0.816 |
| Rod | 1.732 |

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

Ovalbumin



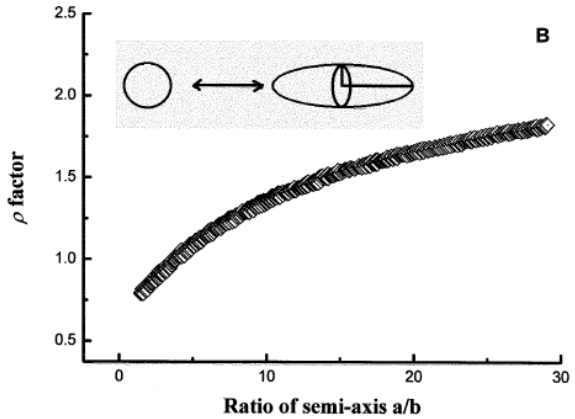

Amyloids



# Shape analysis: **shape factor $\rho = R_g/R_h$**

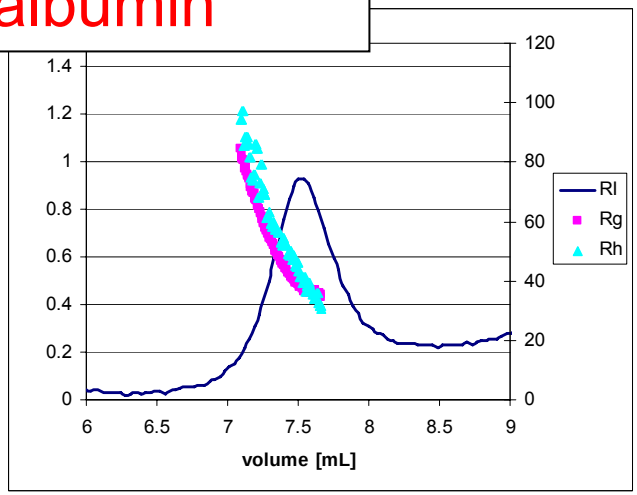
Aggregates of Ovalbumin *vs.* amyloid fibers

Shape factor: $\rho = R_g/R_h$

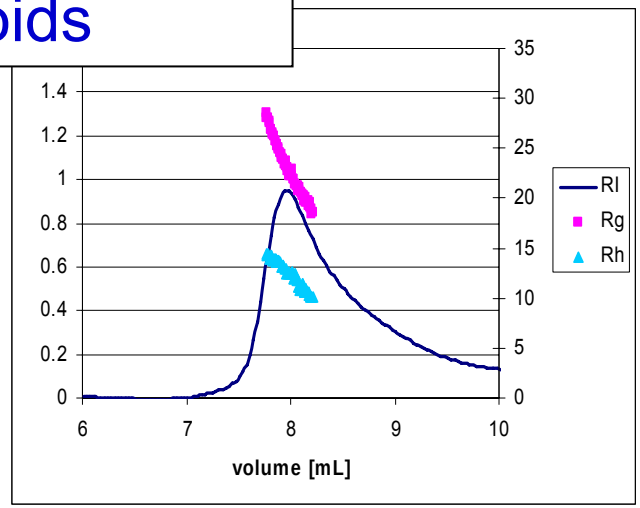



**Combination of MALS ($R_g$) and DLS ($R_h$)**

Ovalbumin




Amyloids

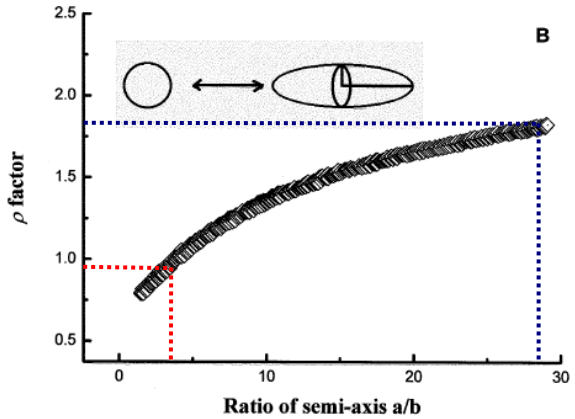



[B] Rizos, A. K., Spandidos, D. A., and Krambovitis, E., (2003) *Int. J. Med.,* 12, 559-563

# Shape analysis: **shape factor $\rho = R_g/R_h$**

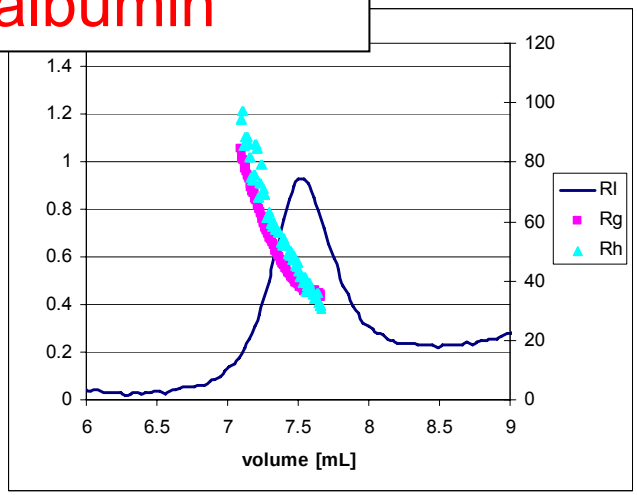
Aggregates of Ovalbumin *vs.* amyloid fibers

Shape factor: $\rho = R_g/R_h$



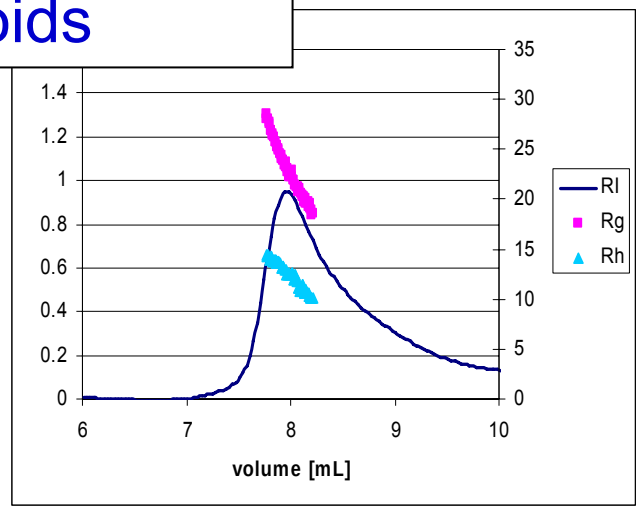

**Combination of MALS ($R_g$) and DLS ($R_h$)**

## Ovalbumin




$R_g/R_h$ = 0.91 Coil

## Amyloids




$R_g/R_h$ = 1.84 Rod

# Shape analysis: **shape factor $\rho = R_g/R_h$**

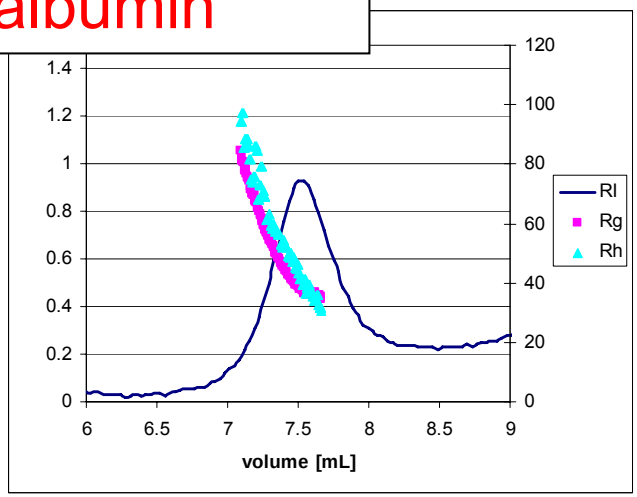
Aggregates of Ovalbumin *vs.* amyloid fibers

Shape factor: $\rho = R_g/R_h$

| For | $\rho = R_g/R_h$ |
|---|---|
| Sphere | 0.774 |
| Coil | 0.816 |
| Rod | 1.732 |

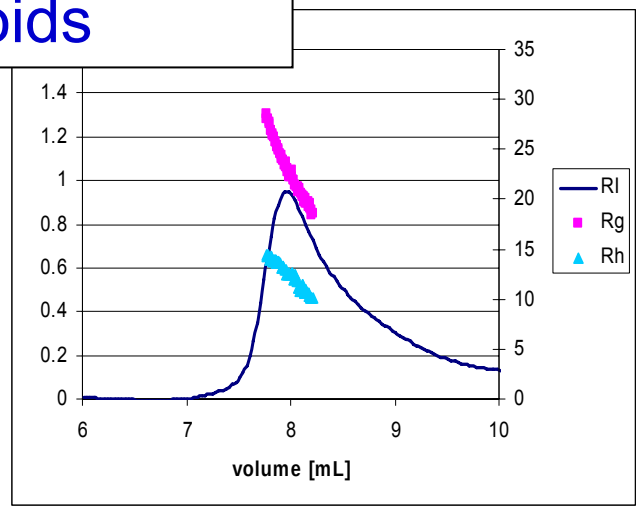
**Combination of MALS ($R_g$) and DLS ($R_h$)**

## Ovalbumin




$R_g/R_h$ = 0.91 Coil

## Amyloids




$R_g/R_h$ = 1.84 Rod

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

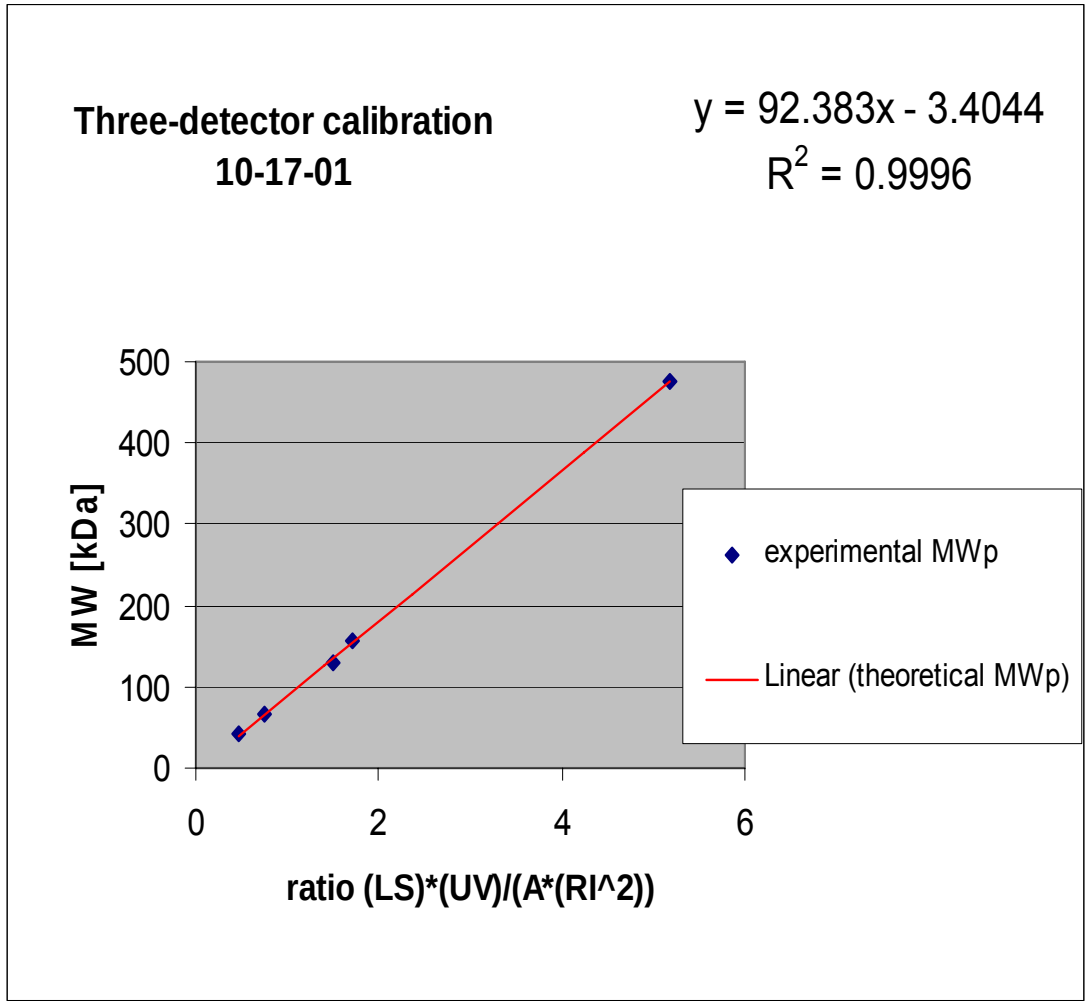
$$MW_p = \frac{k*(LS)(UV)}{\varepsilon(RI)^2}$$

| | |
|---|---|
| $MW_p$ | Molecular Weight (polypeptide) |
| ε | extinction coefficient |
| LS | light scattering intensity |
| UV | absorbance (ε) |
| RI | refractive index change |
| k | calibration constant |

Yutaro Hayashi, Hideo Matsui and Toshio Takagi (1989) Methods Enzymol,172:514-28
Jie Wen, Tsutomu Arakawa and John S. Philo (1996) Anal Biochem, 240(2):155-66

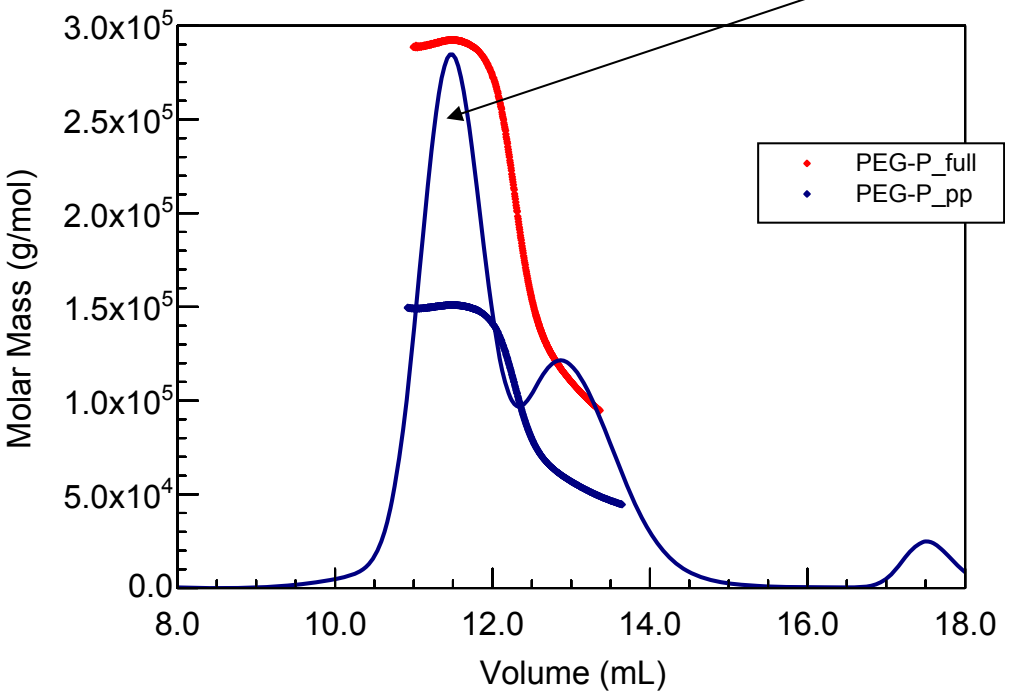
# MWp = 91.39 x [(LS)*(UV)/(ε*(RI^2))]

| Protein | MW (kDa) |
|---|---|
| Ova | 43 |
| BSA(1) | 66 |
| BSA(2) | 132 |
| Ald | 156 |
| Apo-Fer | 475 |



PEG-ylated protein: 75 kDa

36 kDa polypeptide + 39 kDa PEG




Polypeptide: 146 kDa

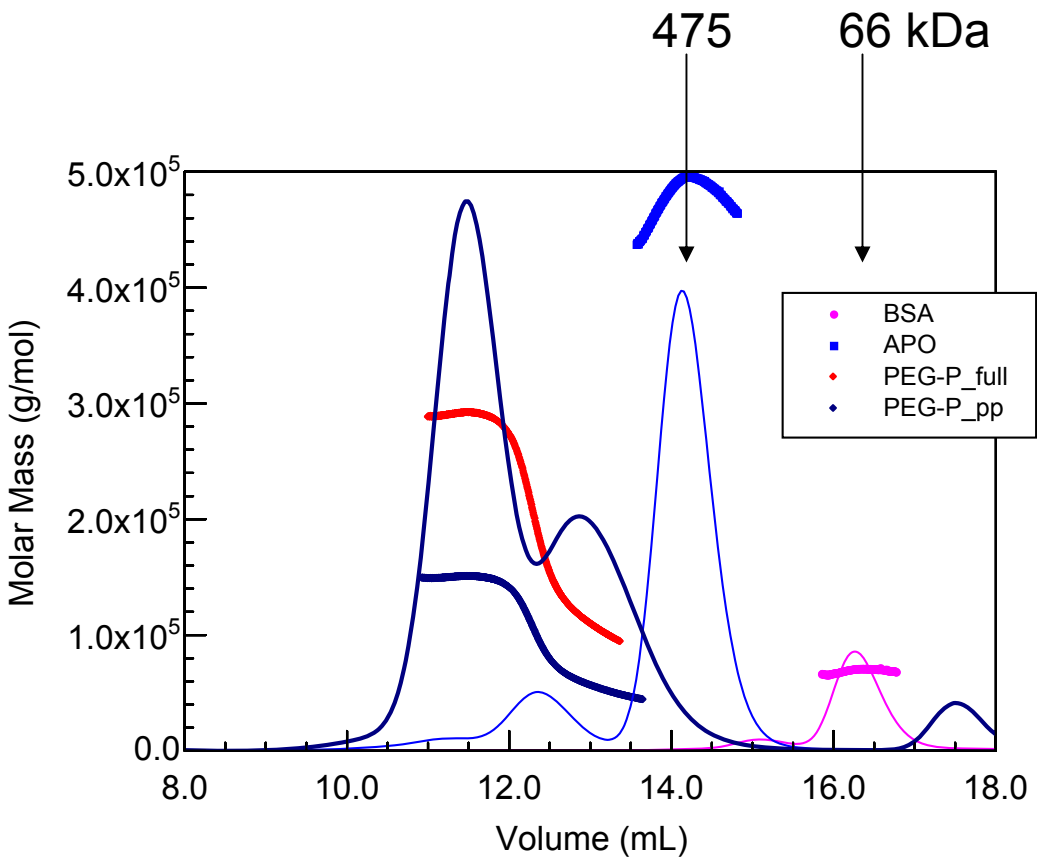
(tetramer: 144 kDa)

Full protein: 291 kDa

(tetramer: 300 kDa)

PEG-ylated protein: 75 kDa

36 kDa polypeptide + 39 kDa PEG




Polypeptide: 146 kDa

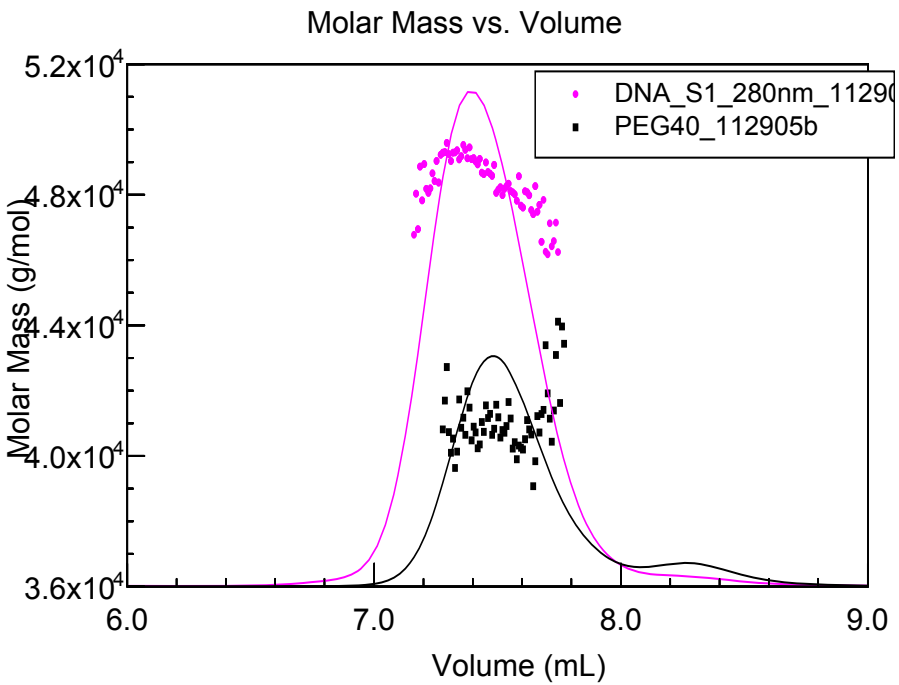
(tetramer: 144 kDa)

Full protein: 291 kDa

(tetramer: 300 kDa)

PEG-ylated oligo: 48.3 kDa

8.3 kDa oligo + 40 kDa PEG



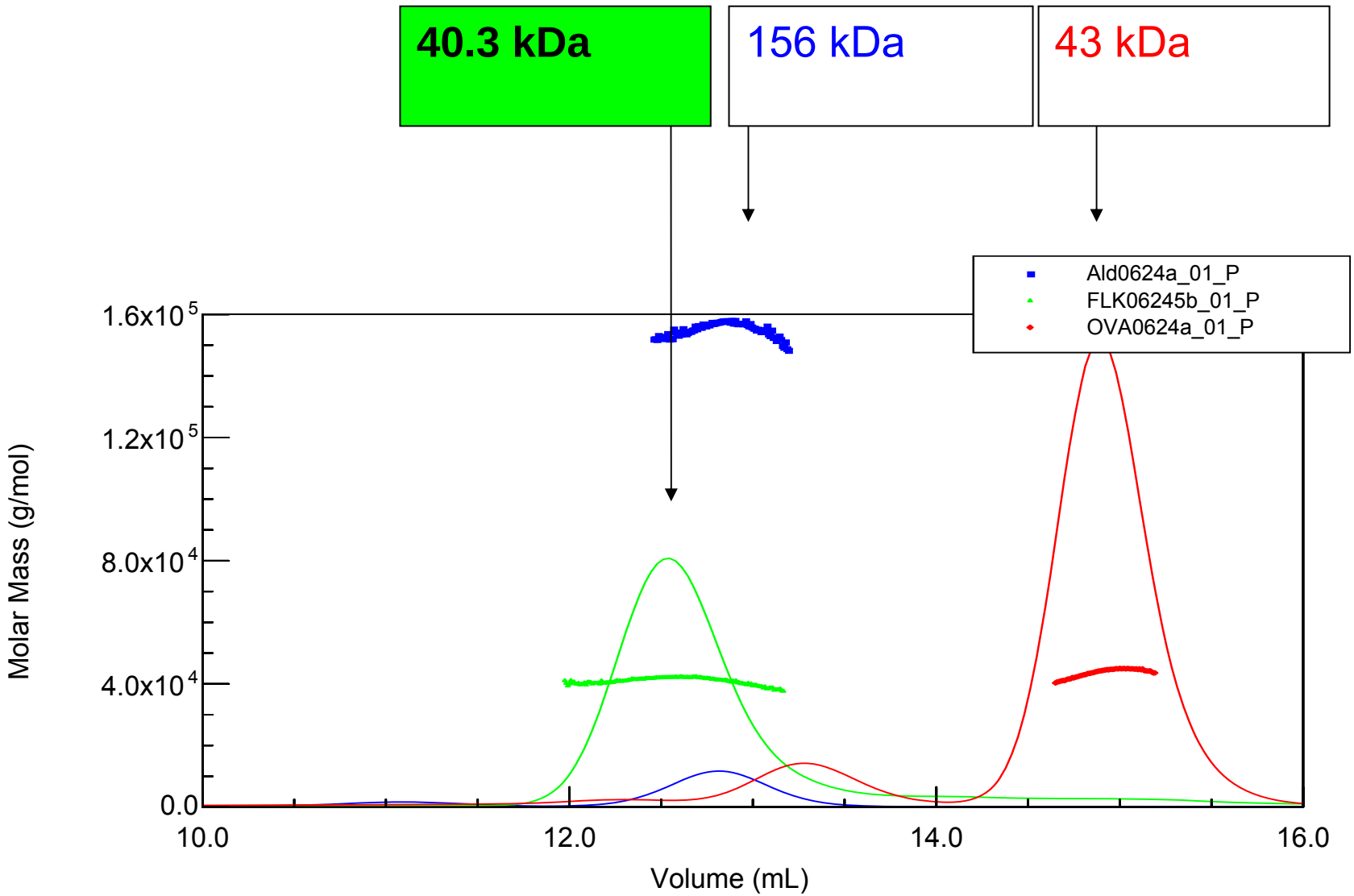

**PEG (40K) MM = 41.0 kDa**

Polydispersity= 1.001

40K PEG + 8.3 kDa oligo

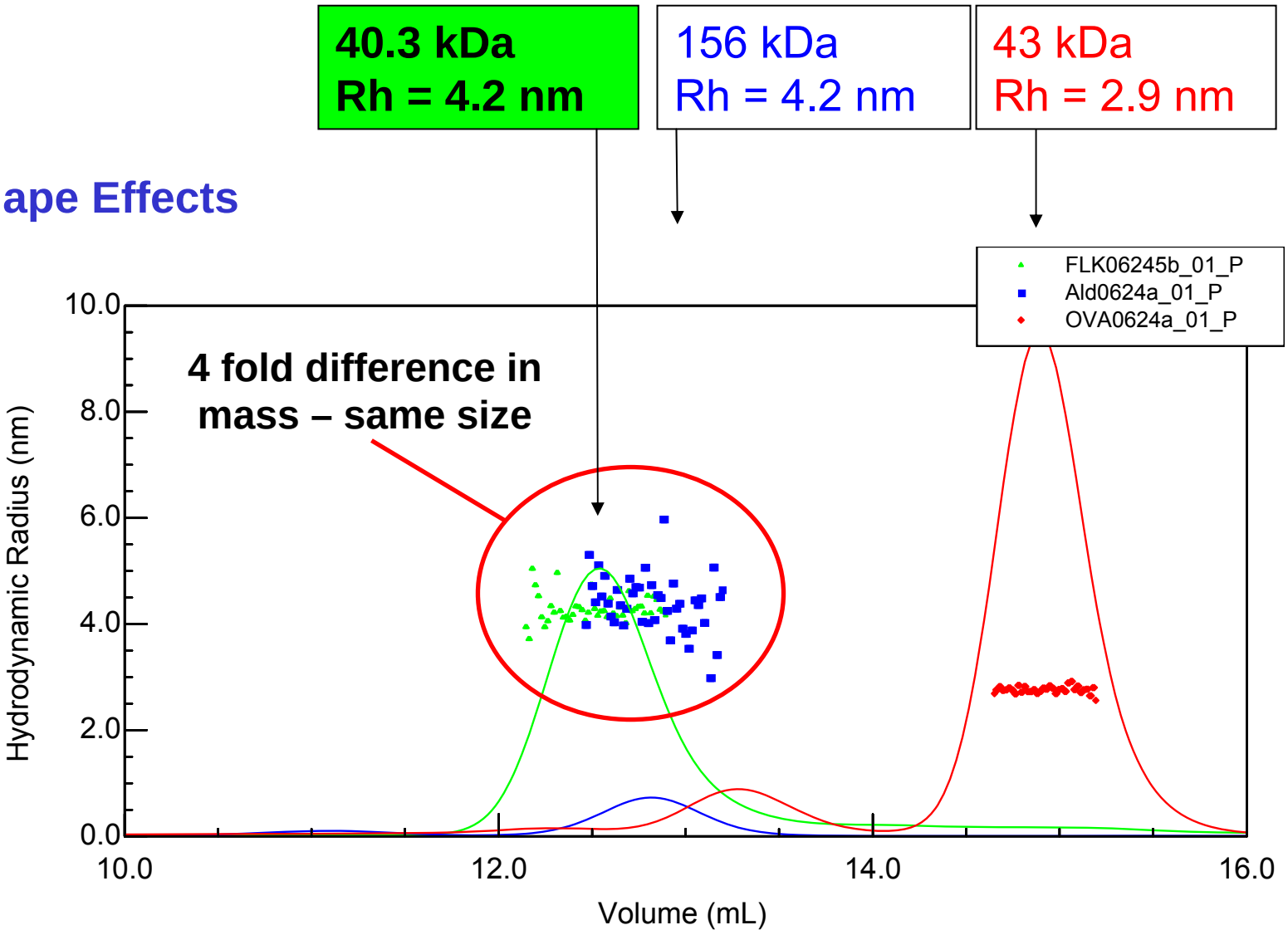
**PEG-oligo MM = 48.5 kDa**

# Protein “F”


40.3 kDa
156 kDa
43 kDa
Ald0624a_01_P
FLK06245b_01_P
OVA0624a_01_P
1.6x10^5
1.2x10^5
8.0x10^4
4.0x10^4
0.0
Molar Mass (g/mol)
10.0
12.0
14.0
16.0
Volume (mL)

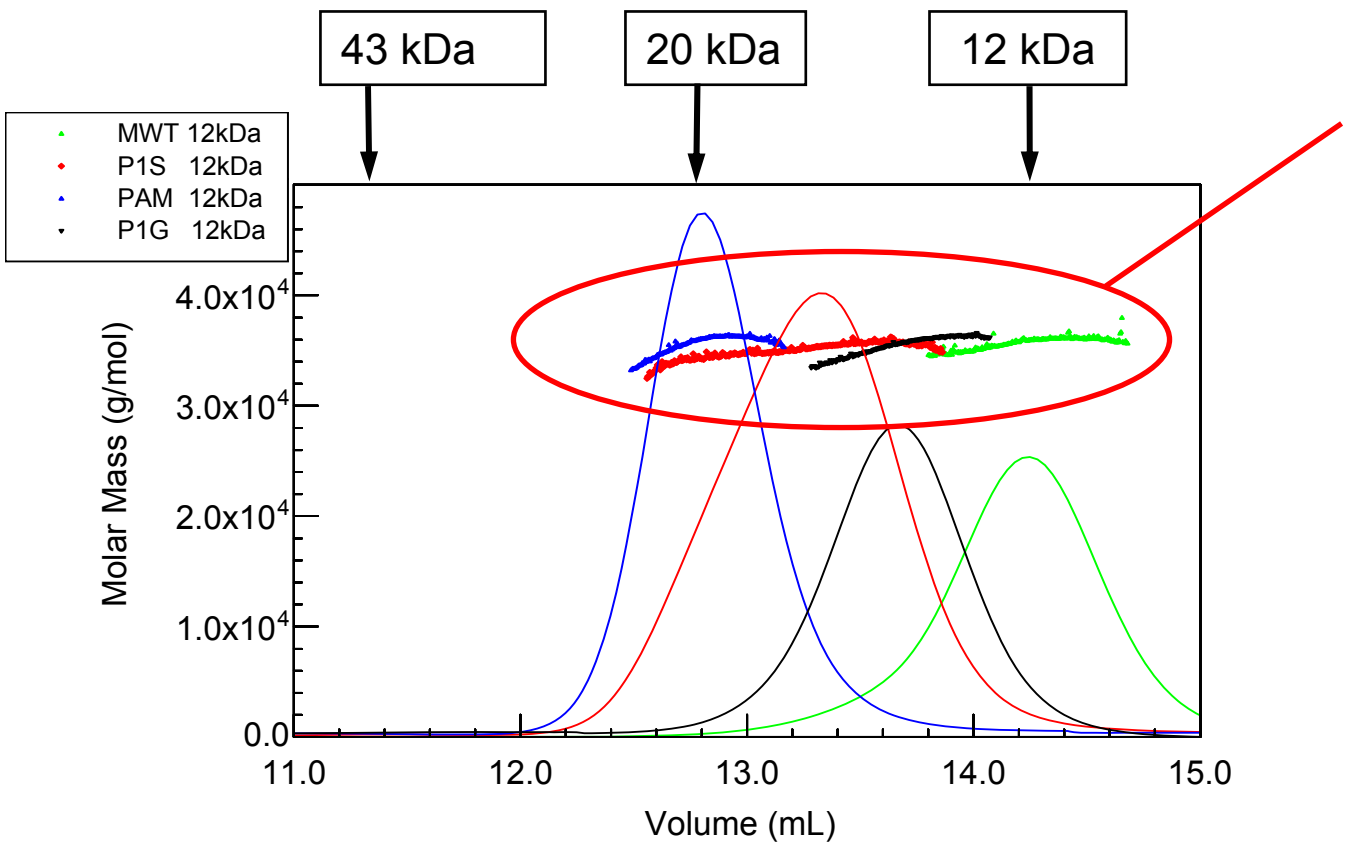
Protein “F” frictional ratio $R_h/R_s$ = 1.85 non-spherical shape

## Shape Effects



# 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

| Parameter | Specification |
|---|---|
| Size range - DLS | 0.6 nm to 6 um Diam |
| Size range - Zeta Potential | 10 nm to 20 um Diam |
| Concentration range | 0.1 mg/mL (Lys) to 30w% |
| Minimum sample volume | 2 uL |
| Temperature control | -4 to 130 $^{o}$C |
| **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