Protein sizing by light scattering, molecular weight and polydispersity

Ulf Nobbmann <Ulf.Nobbmann@malvern.co.uk>

Malvern Instruments

Malvern, Worcestershire WR14 1XZ

Outline

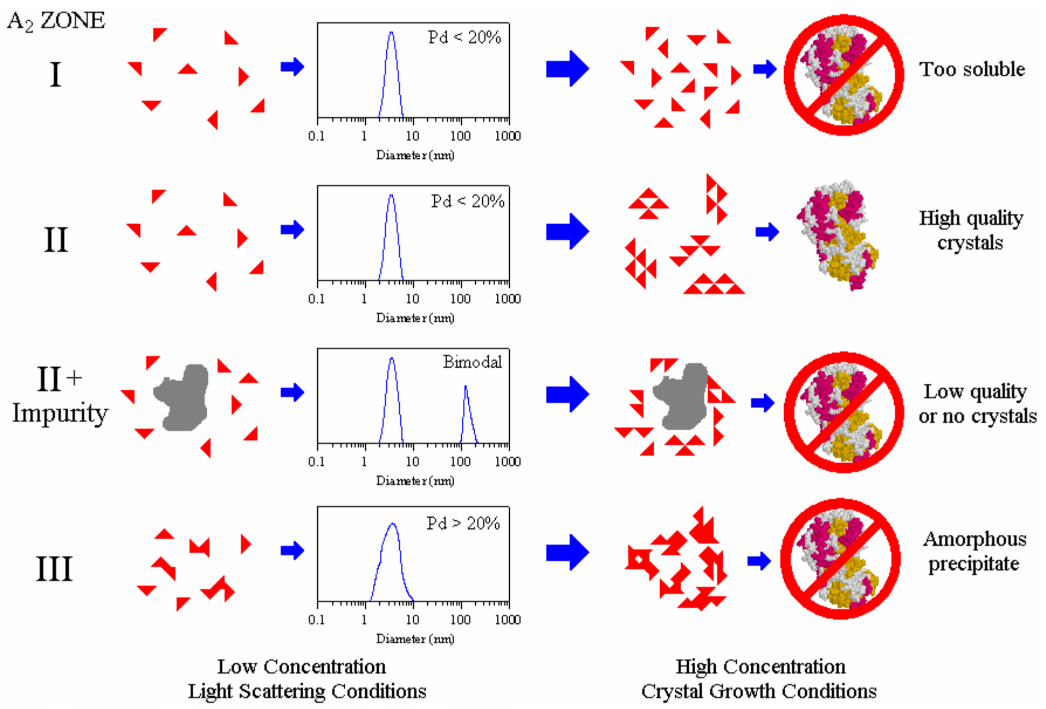
- Why light scattering?
  - Crystallization…
- Theory
  - Static light scattering (SLS) → molecular weight
  - Dynamic light scattering (DLS) → polydispersity
  - Electrophoretic light scattering (ELS) → zeta potential
- Application examples
  - Molecular weight
  - Sizing
  - Polydispersity
- Malvern 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



<u>Rule of thumb:</u>

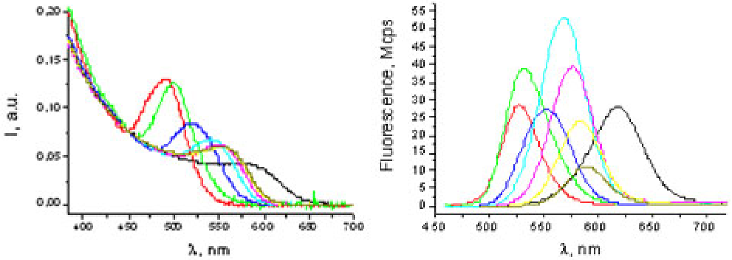
Low polydispersity ($Pd \leq 20\%$) for best chance of crystal growth

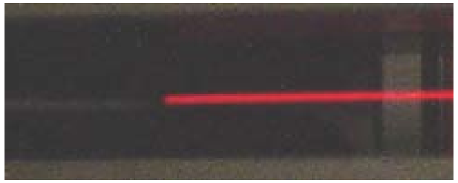
Light - Matter Interactions

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

- Transmission
- Absorption
- Fluorescence

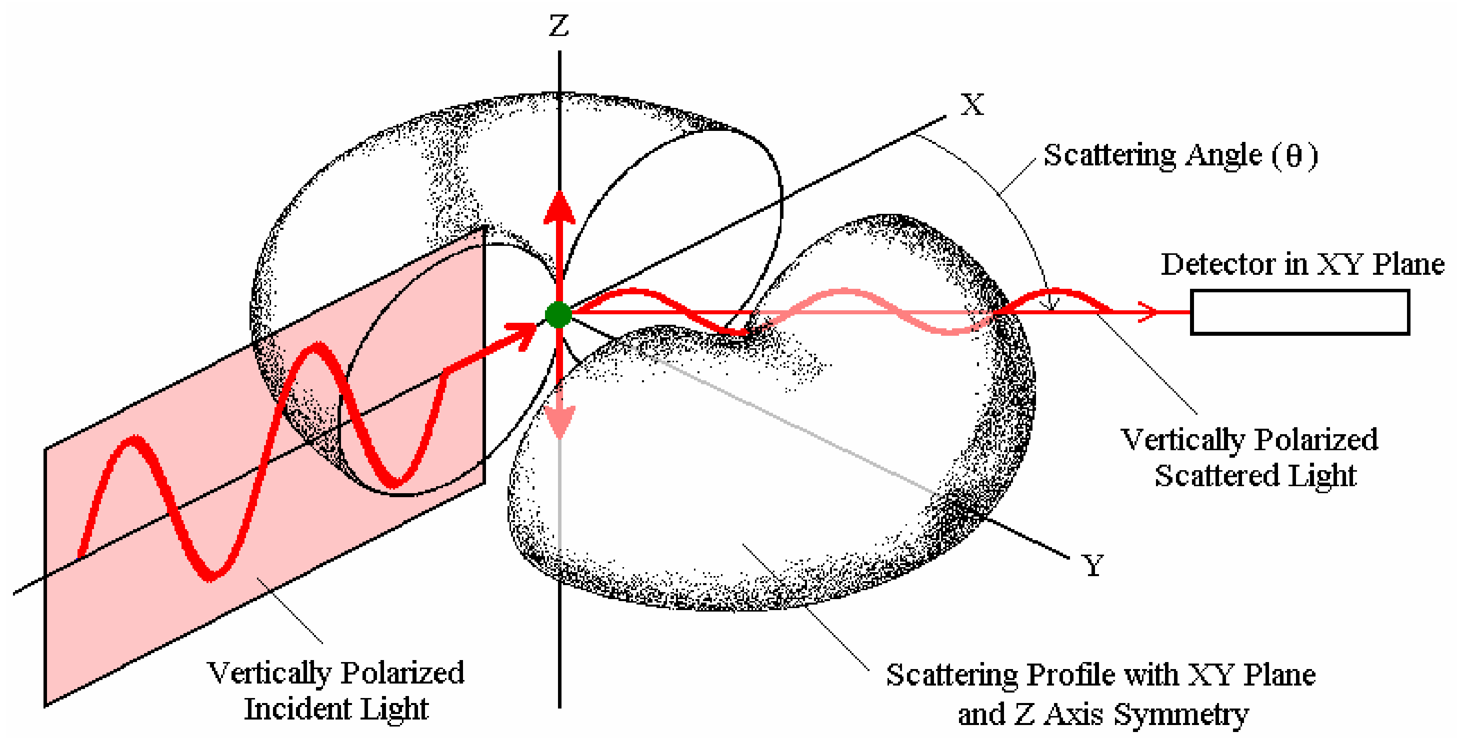
- **Scattering !**





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:

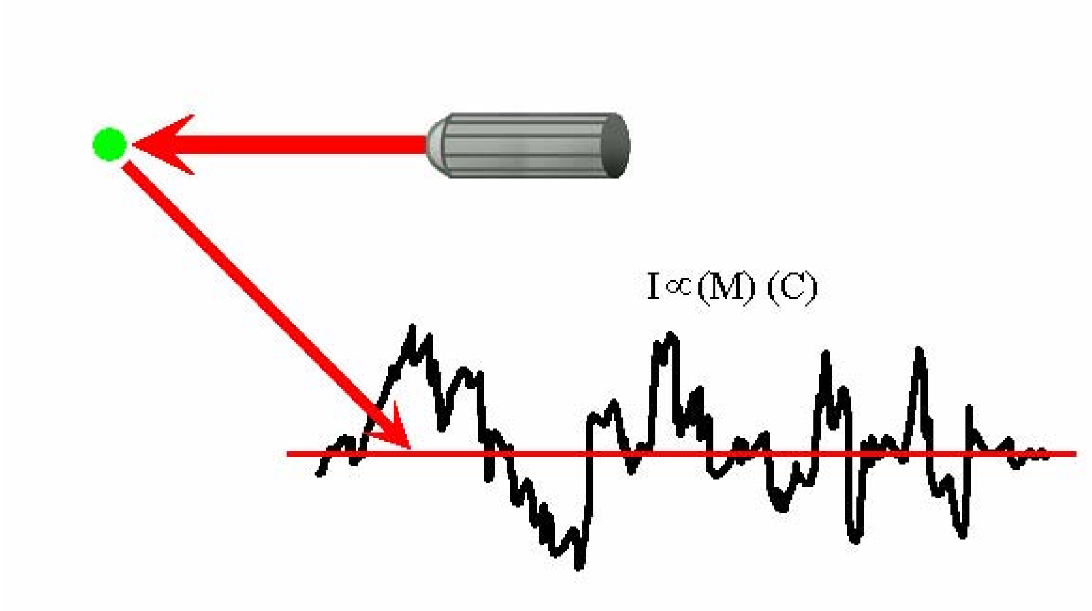
- Average signal strength: **static**, ‘classic’
- Fluctuations of signal: **dynamic**, quasi-elastic
- Shift of the signal: **electrophoretic**

Static Light Scattering

Molecular Weight Measurements

Static Light Scattering (SLS)

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



Rayleigh Equation

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

K = Optical constant
M = Molecular weight
A_2 = 2nd Virial coefficient

C = Concentration
R_θ = Rayleigh ratio
$P(\theta)$ = Shape (or form) factor

Static Light Scattering (SLS)

$$\frac{\boxed{K}C}{\boxed{R_\theta}} = \left(\frac{1}{M} + 2A_2C\right)\boxed{\frac{1}{P_\theta}}$$

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

λ_o = laser wavelength
N_A = Avogadros number
n_o = Solvent RI
dn/dc = 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 = Radius of gyration
θ = Measurement angle

$$R_\theta = \frac{I_A n_o^{\,2}}{I_T n_T^{\,2}} R_T$$

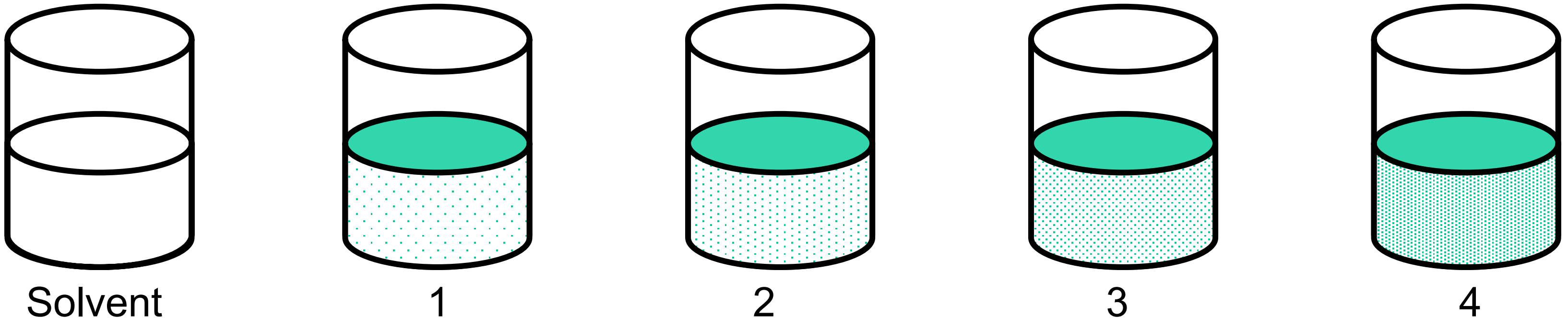
I_A = Intensity of analyte (sample I – solvent I)
n_o = Solvent RI
I_T = Intensity of standard (toluene)
n_T = Standard (toluene) RI
R_T = Rayleigh ratio of standard (toluene)

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 ($I \alpha (M_W)(C)$)
- 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(θ) = 1 and the equation is simplified to

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

Therefore a plot of KC/R_θ 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)

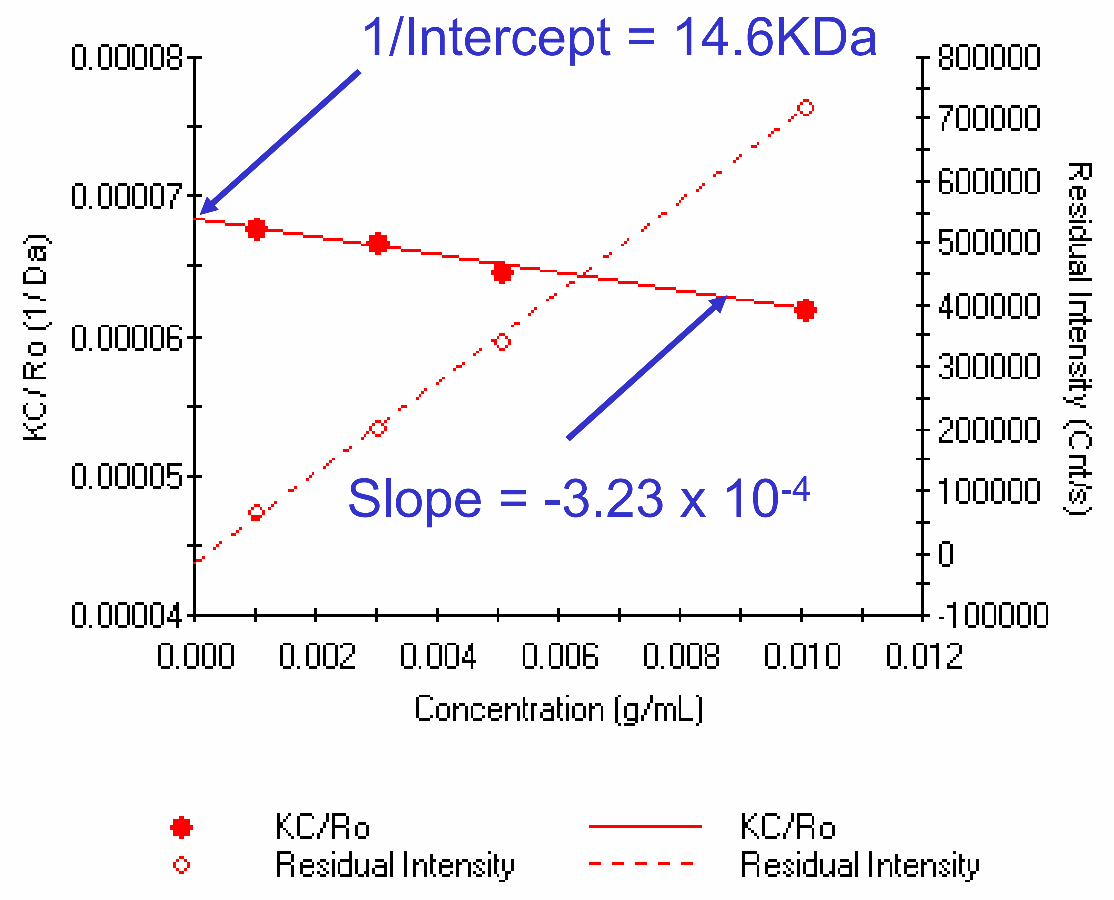
$$\frac{dn}{dc} = 0.185(mL/g)$$

I_{tol} = 192630 (counts/sec)

I_{sol} = 21870 (counts/sec)

Lysozyme Concentration (mg/mL)	Measured Intensity (counts/sec)	Intensity of Analyte (counts/sec)	KC/R_θ (1/Da)
1.006	87,830	65,960	6.1994×10^{-5}
3.018	222,900	201,030	6.4765×10^{-5}
5.029	366,770	344,900	6.6682×10^{-5}
10.059	742,570	720,700	6.7743×10^{-5}

Molecular Weight Example (Lysozyme in PBS)



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

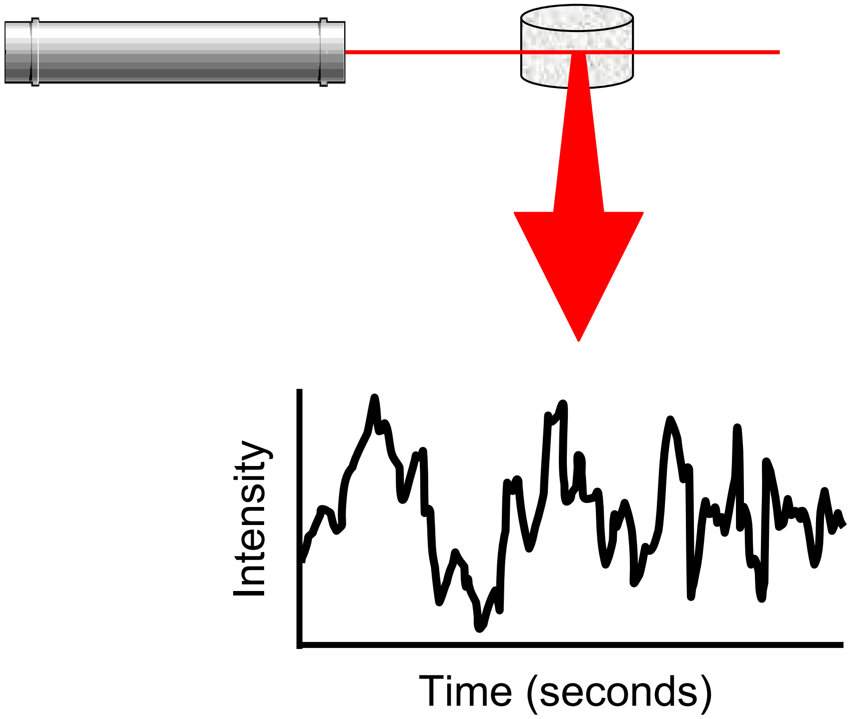
Molecular Size Measurements

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

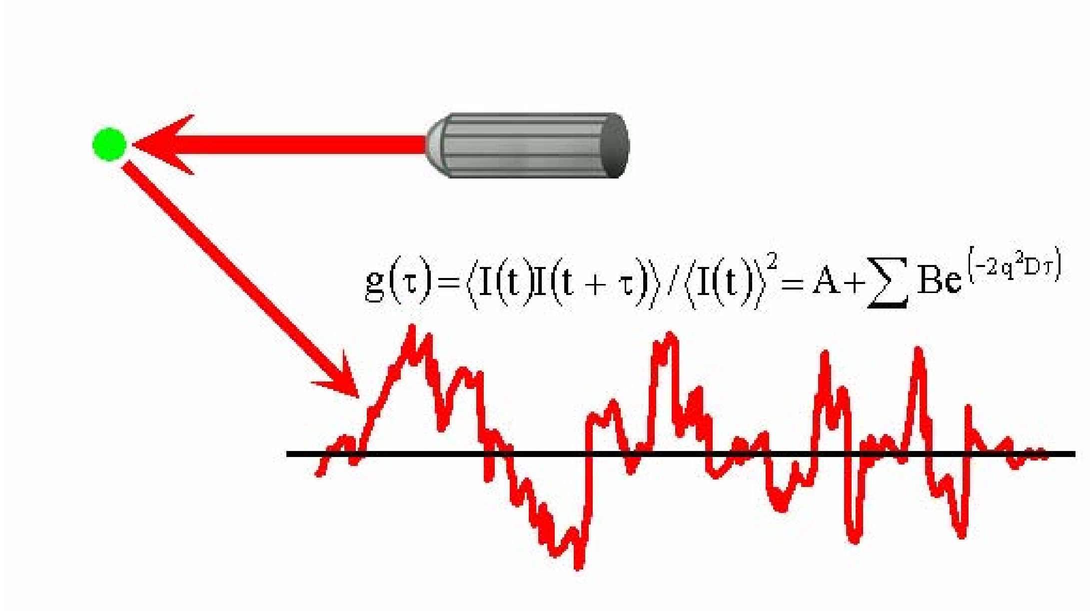
- Dynamic light scattering is a technique for measuring the size of molecules and nanoparticles
- DLS measures the time dependent fluctuations in the scattering intensity to determine the translational diffusion coefficient (D), and subsequently the hydrodynamic size



- 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.



$$g(\tau)=\langle I(t)I(t+\tau)\rangle/\langle I(t)\rangle^{2}=A+\sum Be^{(-2q^{2}D\tau)}$$

Stokes-Einstein

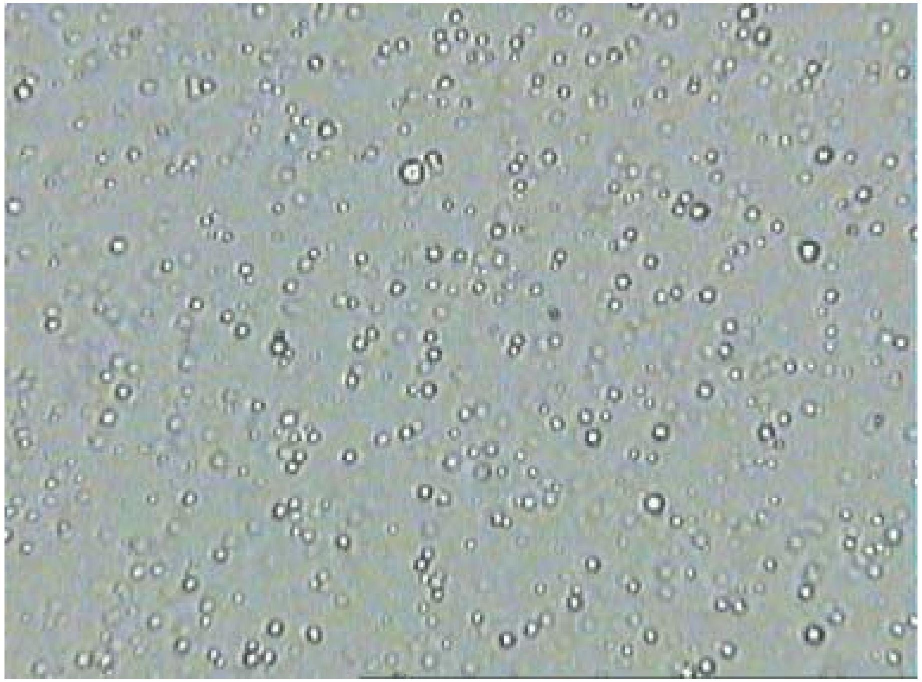
$$R_H=\frac{kT}{6\pi\eta D}$$

q = Scattering vector
R_H = Radius
T = Temperature

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

Brownian Motion

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

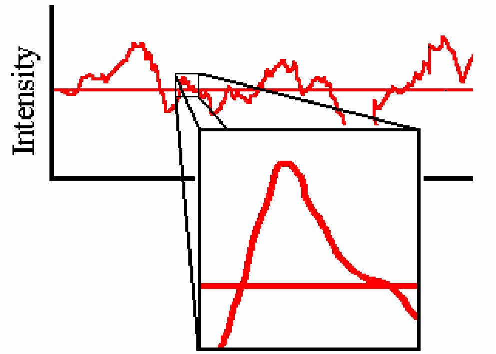


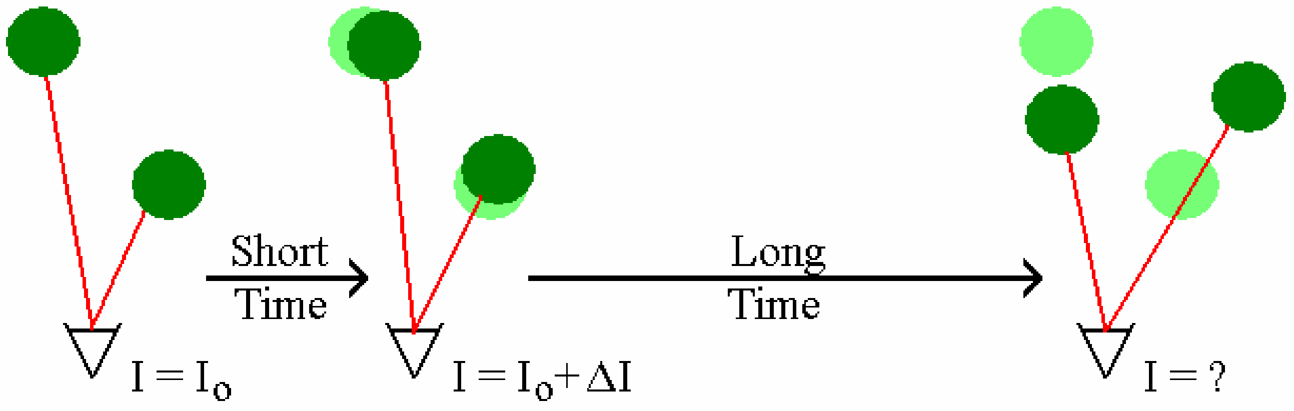
Brownian Motion

- 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

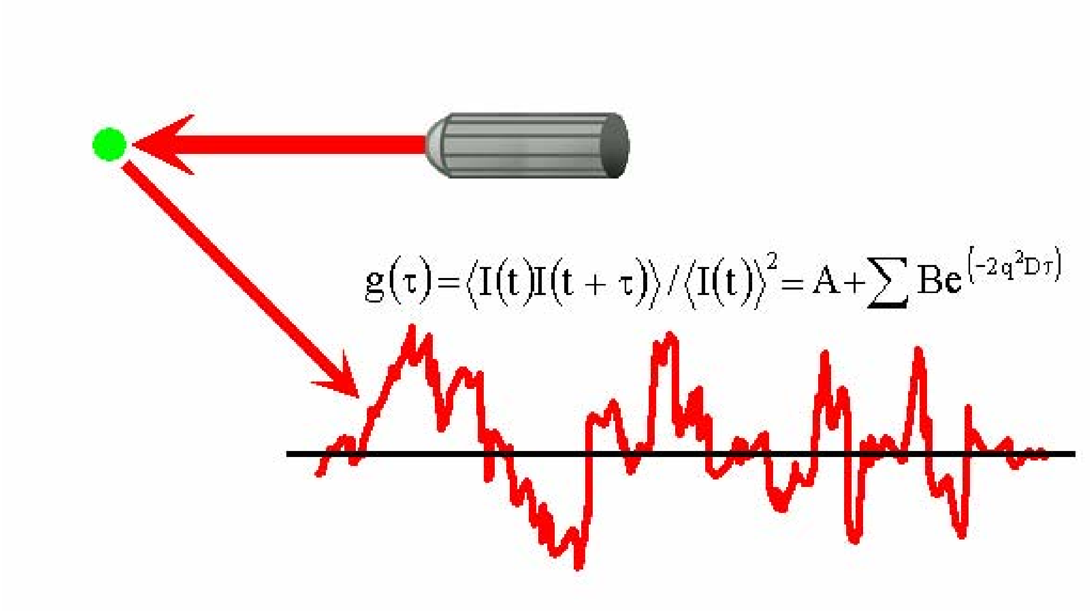
The non-randomness of the intensity trace is a consequence of the physical confinement of the particles to be in locations very near to their initial locations across very short time intervals.





Dynamic Light Scattering (DLS)

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



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

Stokes-Einstein Equation

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

where

R_H = hydrodynamic diameter

k_B = Boltzmann's constant

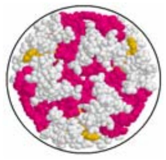
T = absolute temperature

η = viscosity

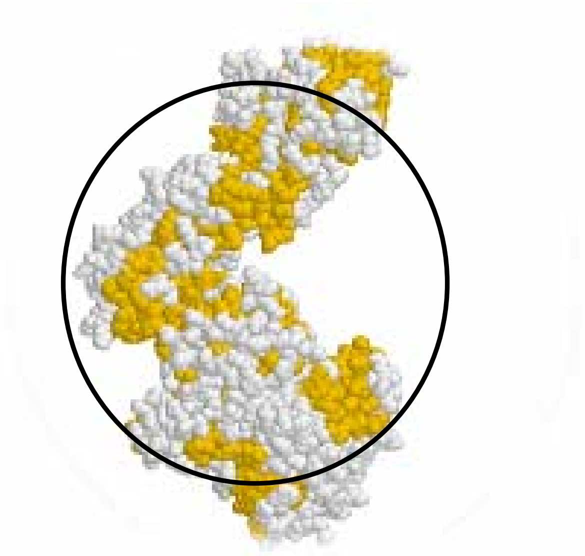
D_T = diffusion coefficient

Comparative Protein R_H Values

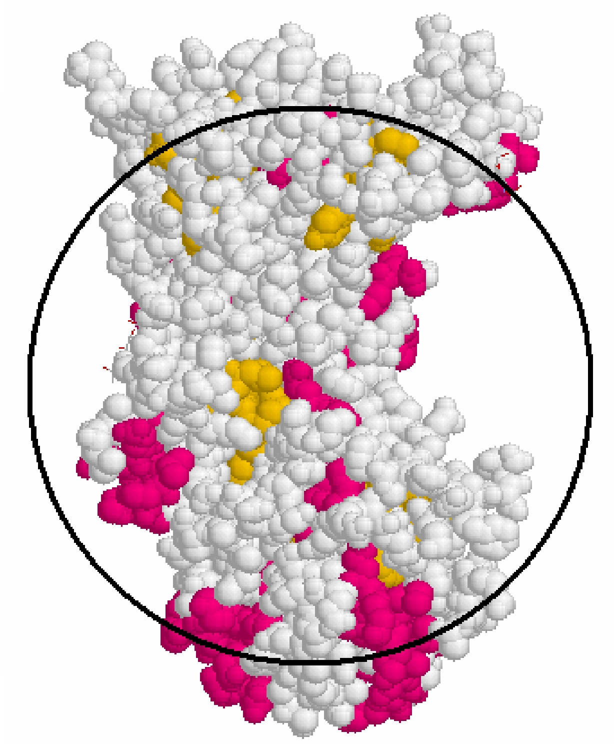
Lysozyme
M_W=14.5 kDa
R_H=1.9 nm



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



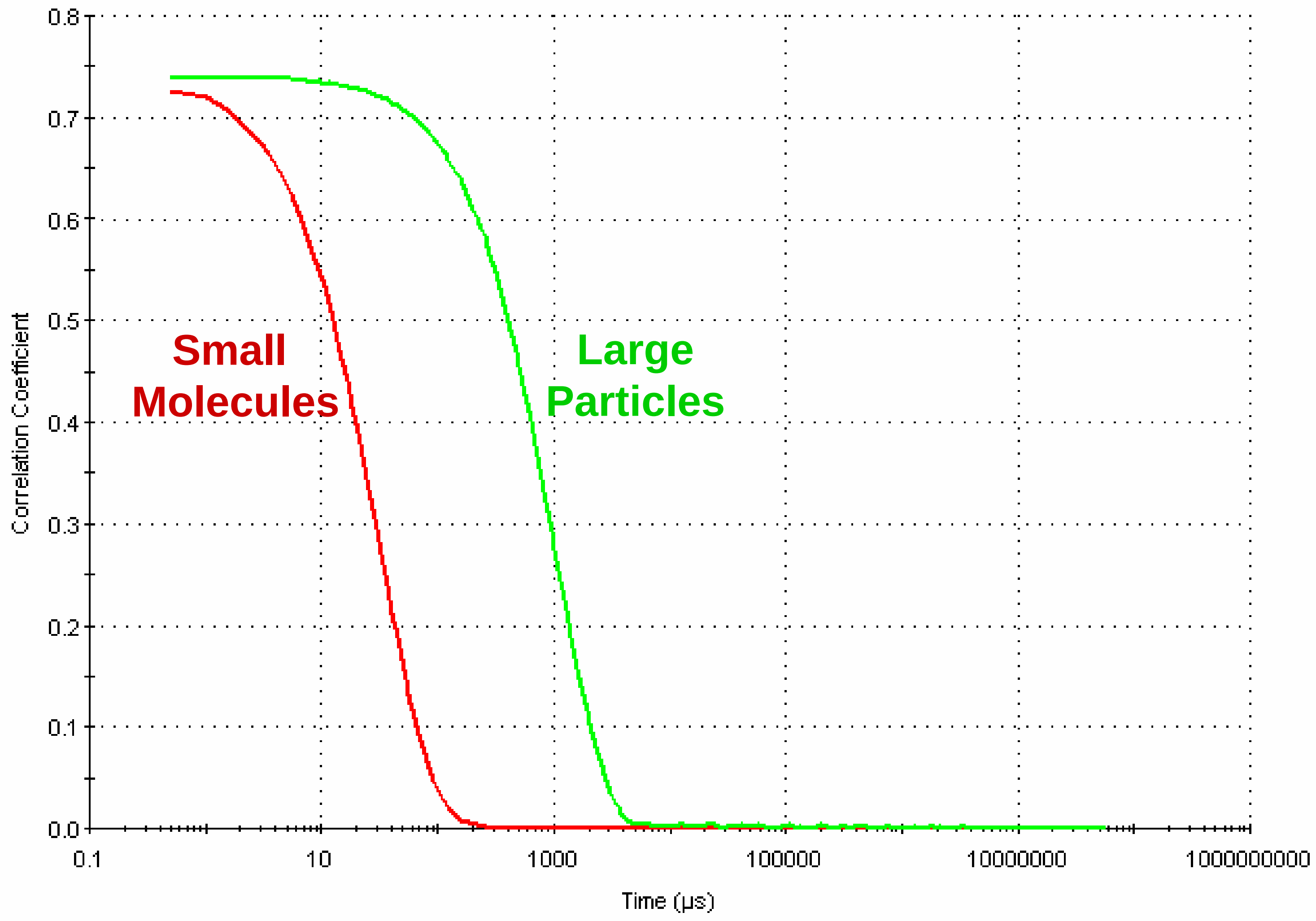
Immunoglobulin G
M_W=160 kDa
R_H=7.1 nm



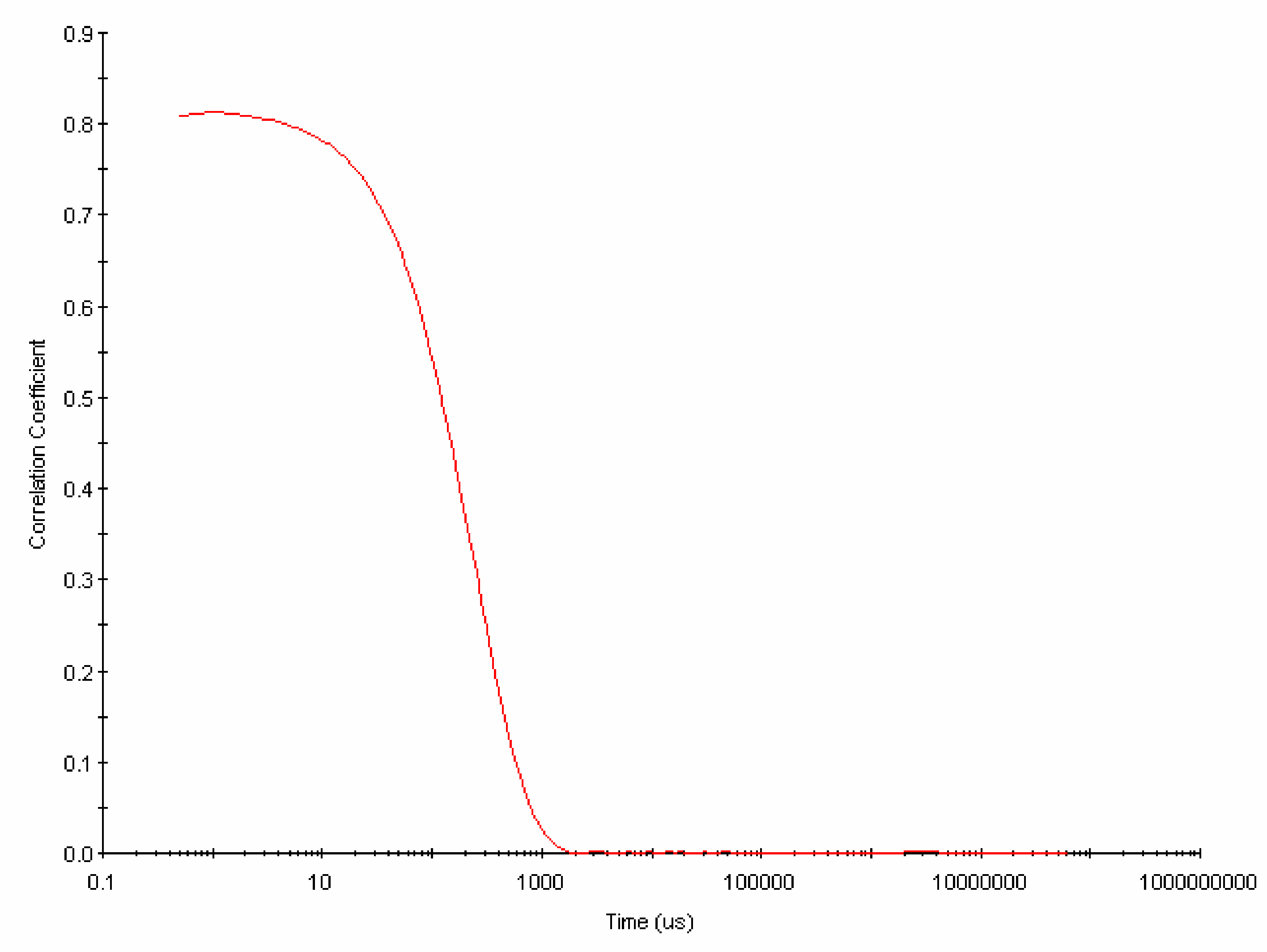
Thyroglobulin
M_W=650 kDa
R_H=10.1 nm

5 nm

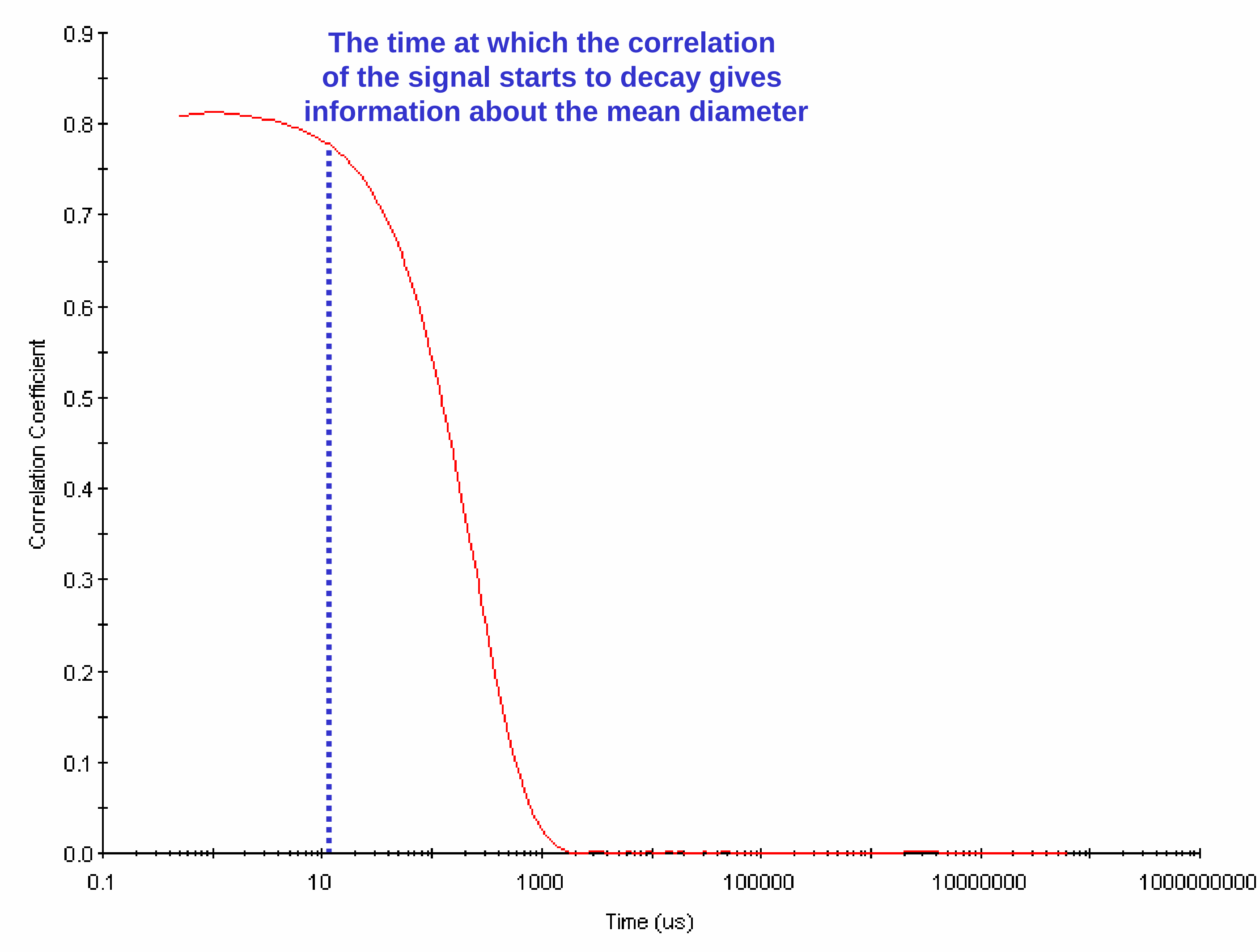
Correlogram Interpretation



Correlogram Interpretation

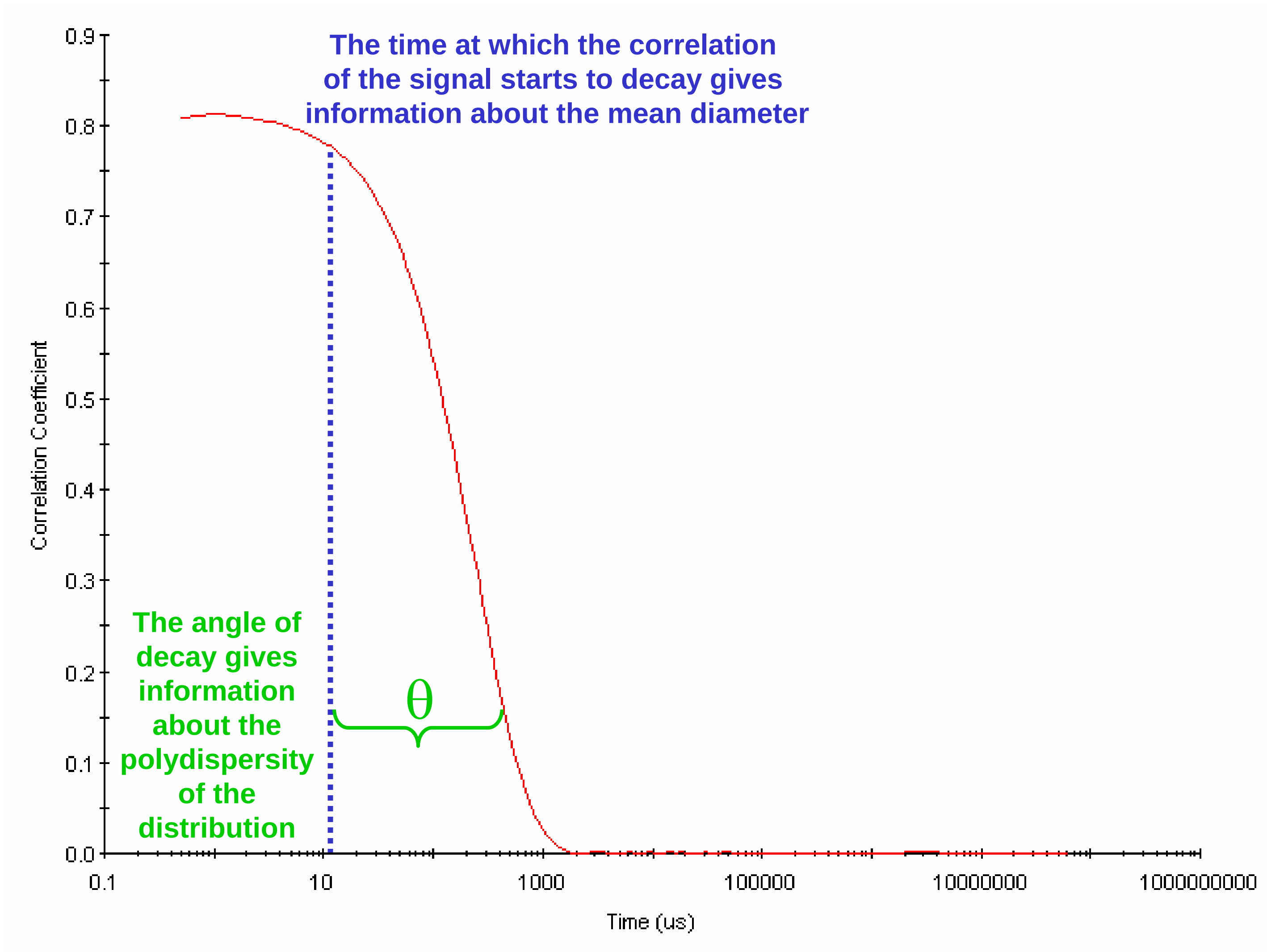


Correlogram Interpretation



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

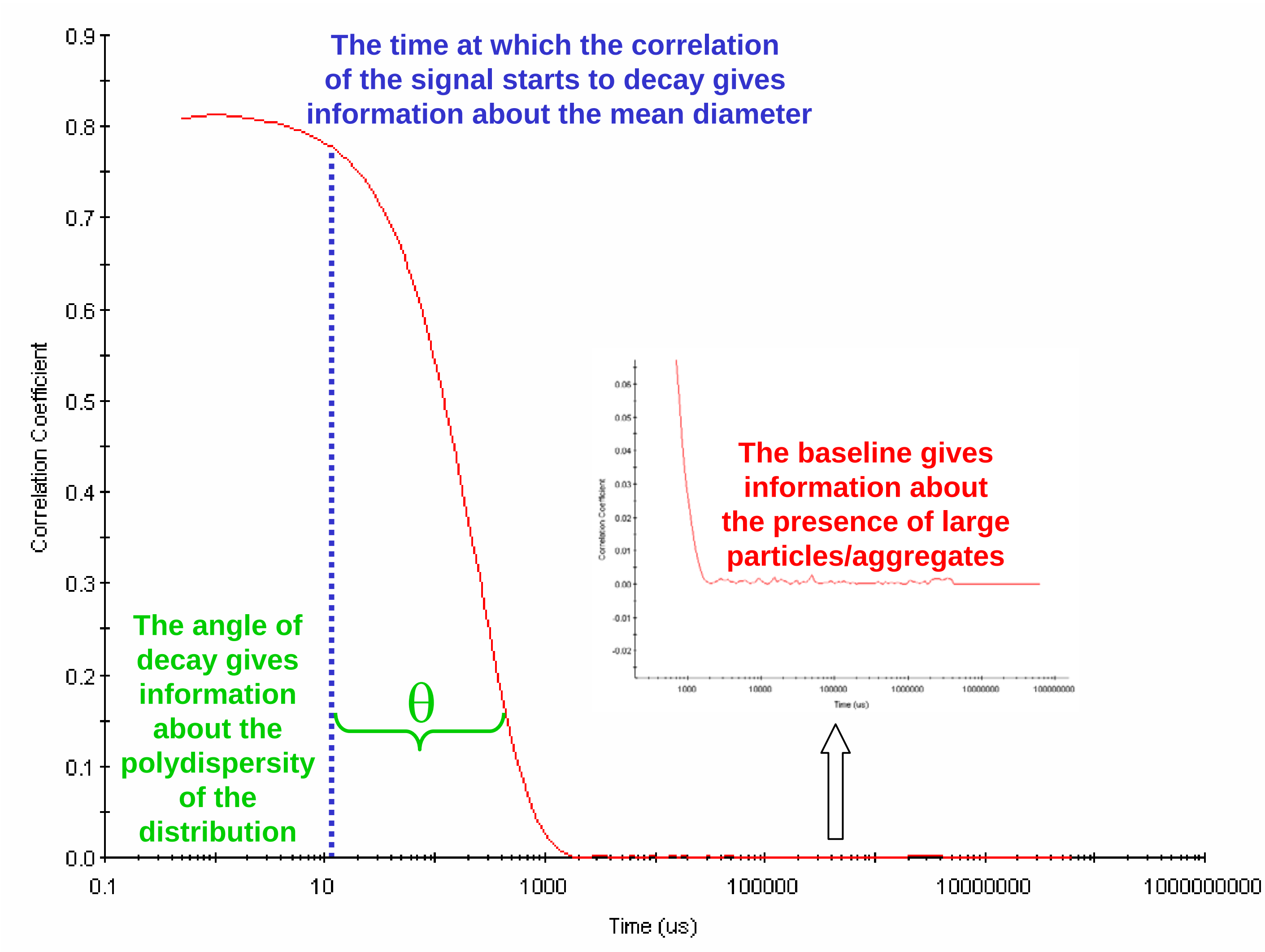
Correlogram Interpretation



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

Correlogram Interpretation



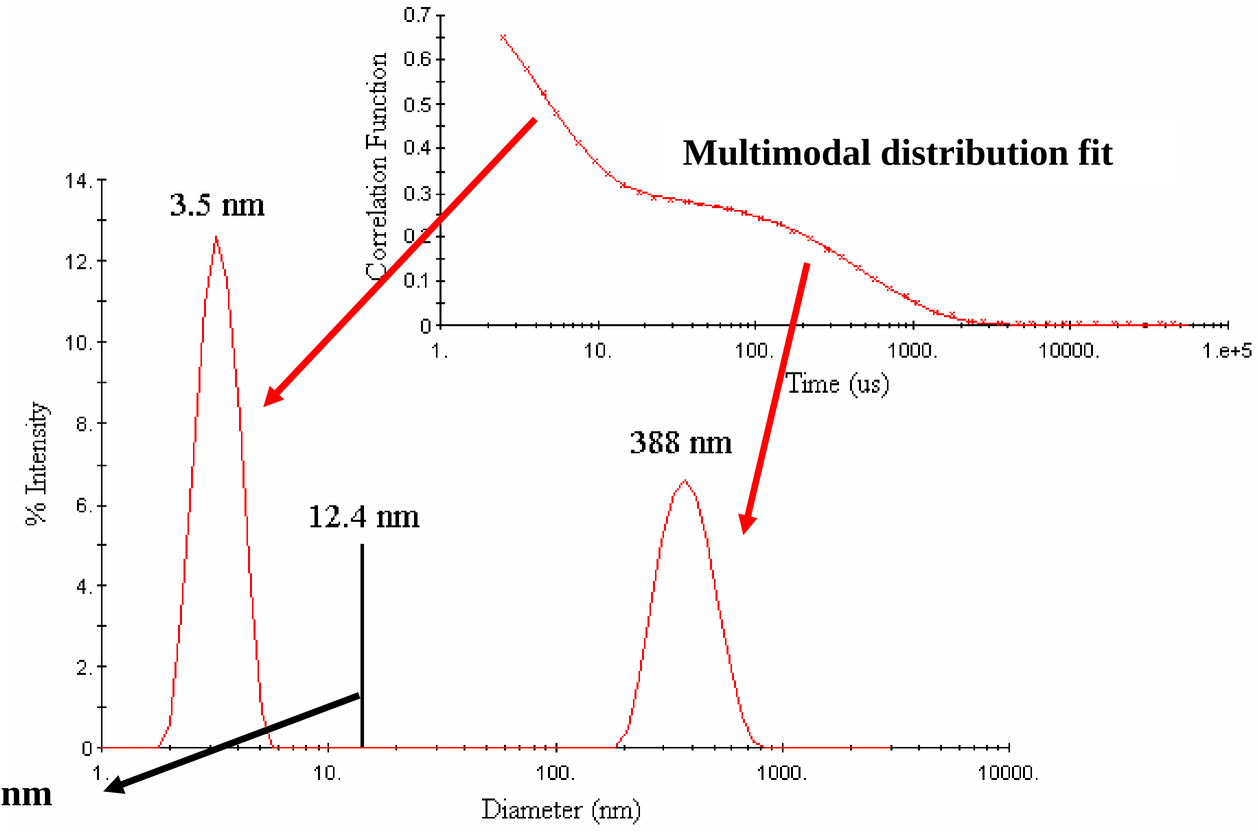
The time at which the correlation of the signal starts to decay gives information about the mean diameter

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 multi-modal distribution results.



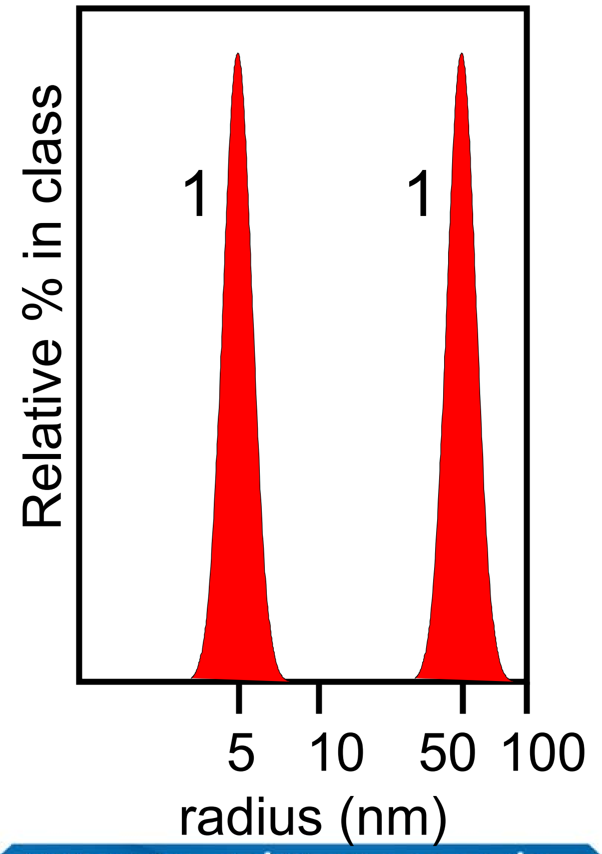
Z average: 12.4nm

(Single species assumption)

Intensity, Volume And Number Distributions

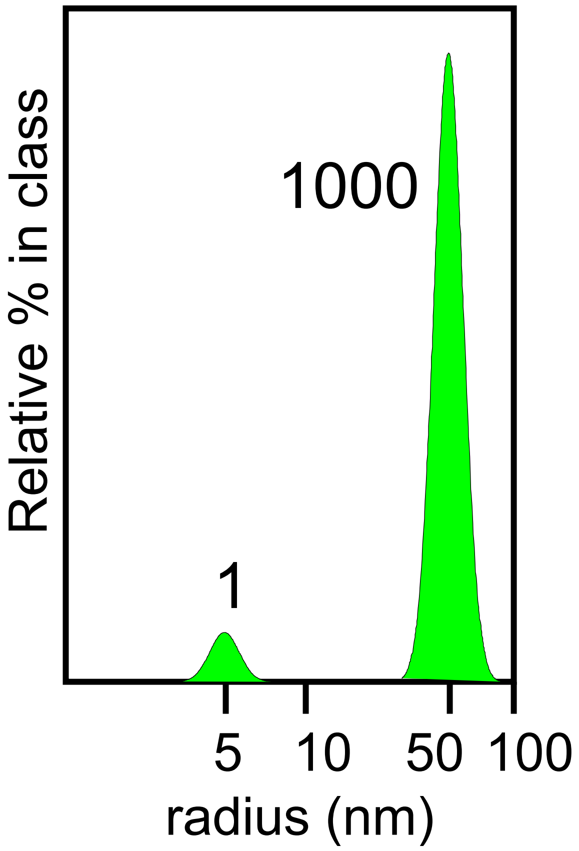
Mixture containing equal numbers of 5 and 50nm spherical particles

NUMBER



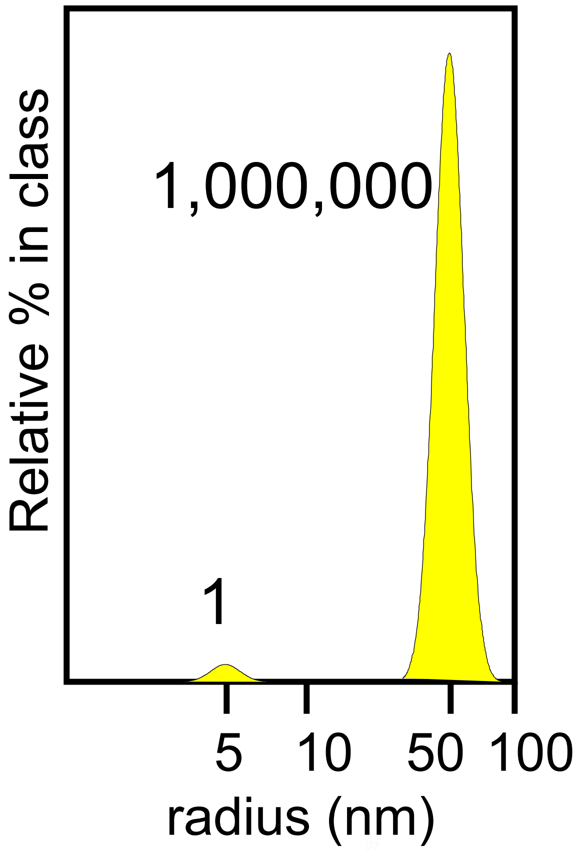
VOLUME

$$= \frac{4}{3}\pi r^3$$



INTENSITY

$$= r^6$$



Benefits Of Sizing By DLS

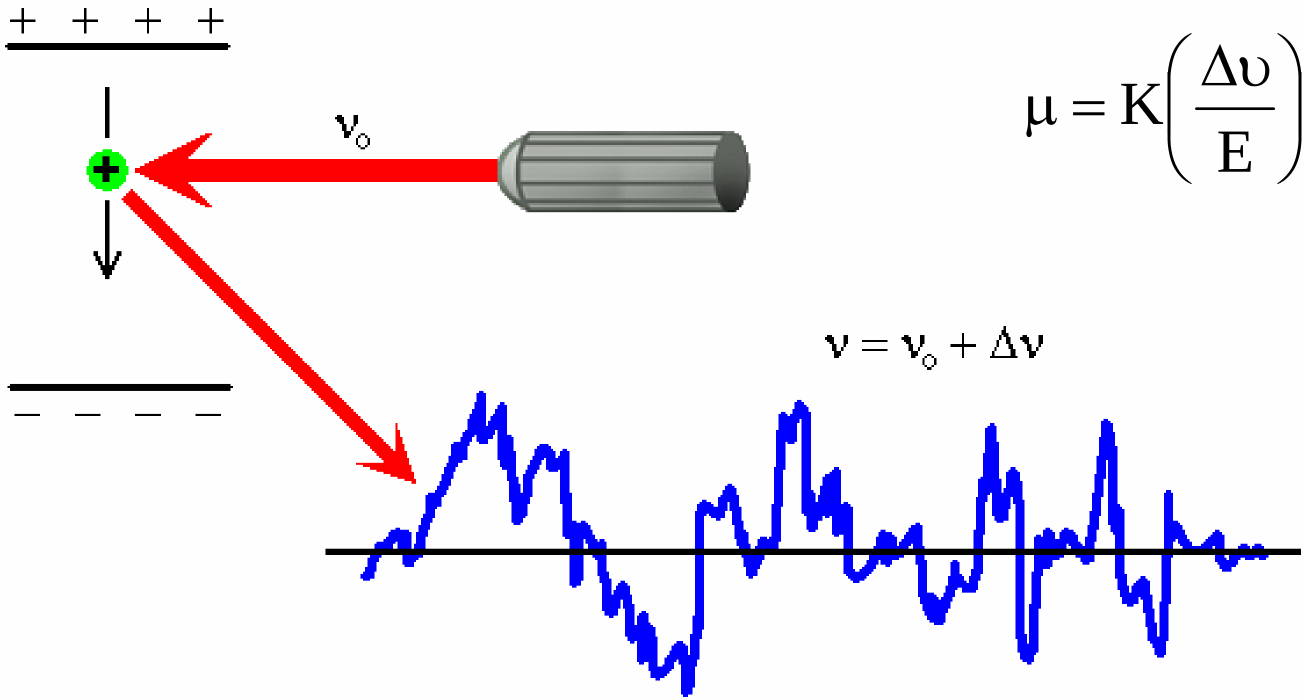
- Non-invasive
- High sensitivity (< 0.1 mg/mL for typical proteins)
- Low volume (12 µL)
- Scattering intensity is proportional to the square of the protein molecular weight, making the technique ideal for identifying the presence of trace amounts of aggregate.

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.

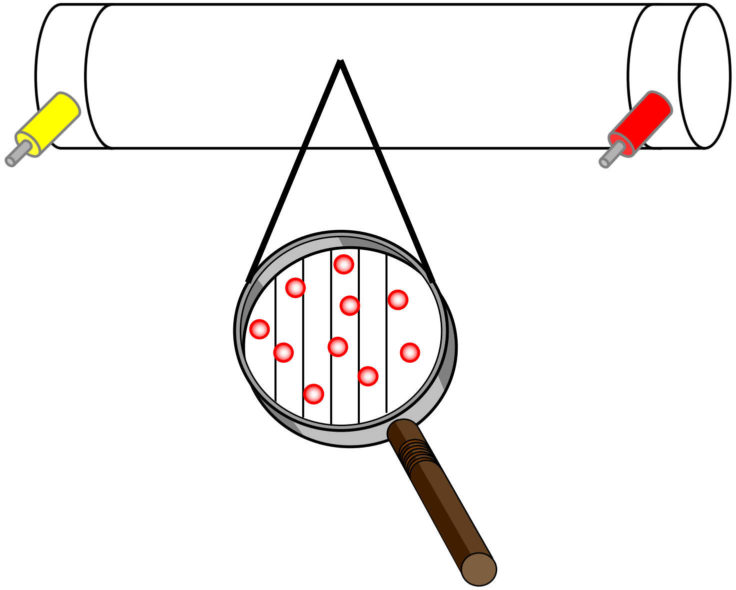


$$\mu = K\left(\frac{\Delta\upsilon}{E}\right)$$

μ 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

Electrophoresis in a

Closed Capillary Cell

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 Δf is equal to:

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

v = the particle velocity
λ = laser wavelength
θ = scattering angle

…measure phase instead

Mixed Mode Measurement (M3)

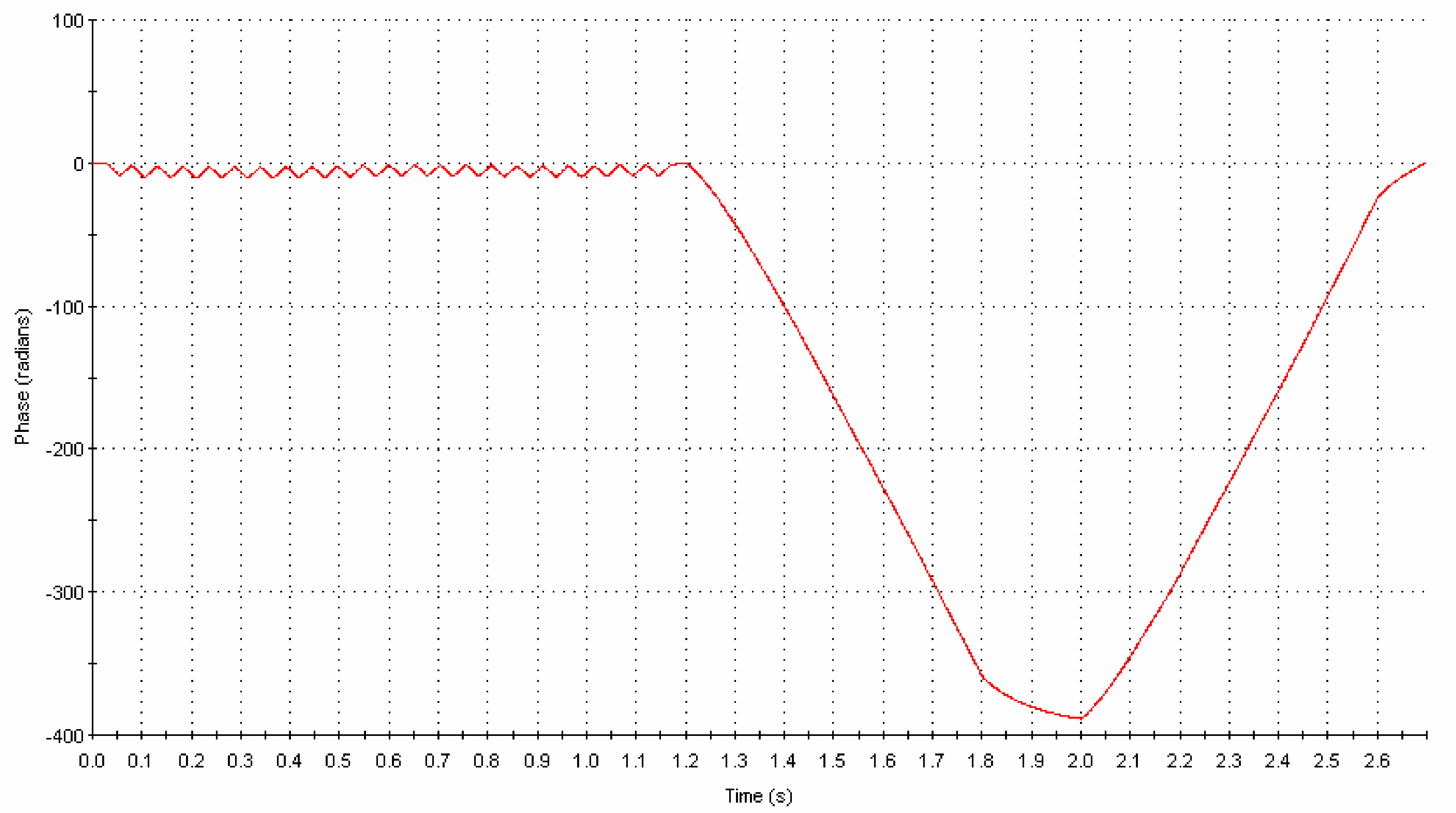
- Mixed mode measurement (M3) is a patented method that allows measurement at any point in a capillary cell
- It eliminates electroosmosis by reversing the applied field at a high frequency
- Malvern have combined M3 with PALS to improve the measurement sensitivity and accuracy (M3-PALS)

Phase Analysis Light Scattering

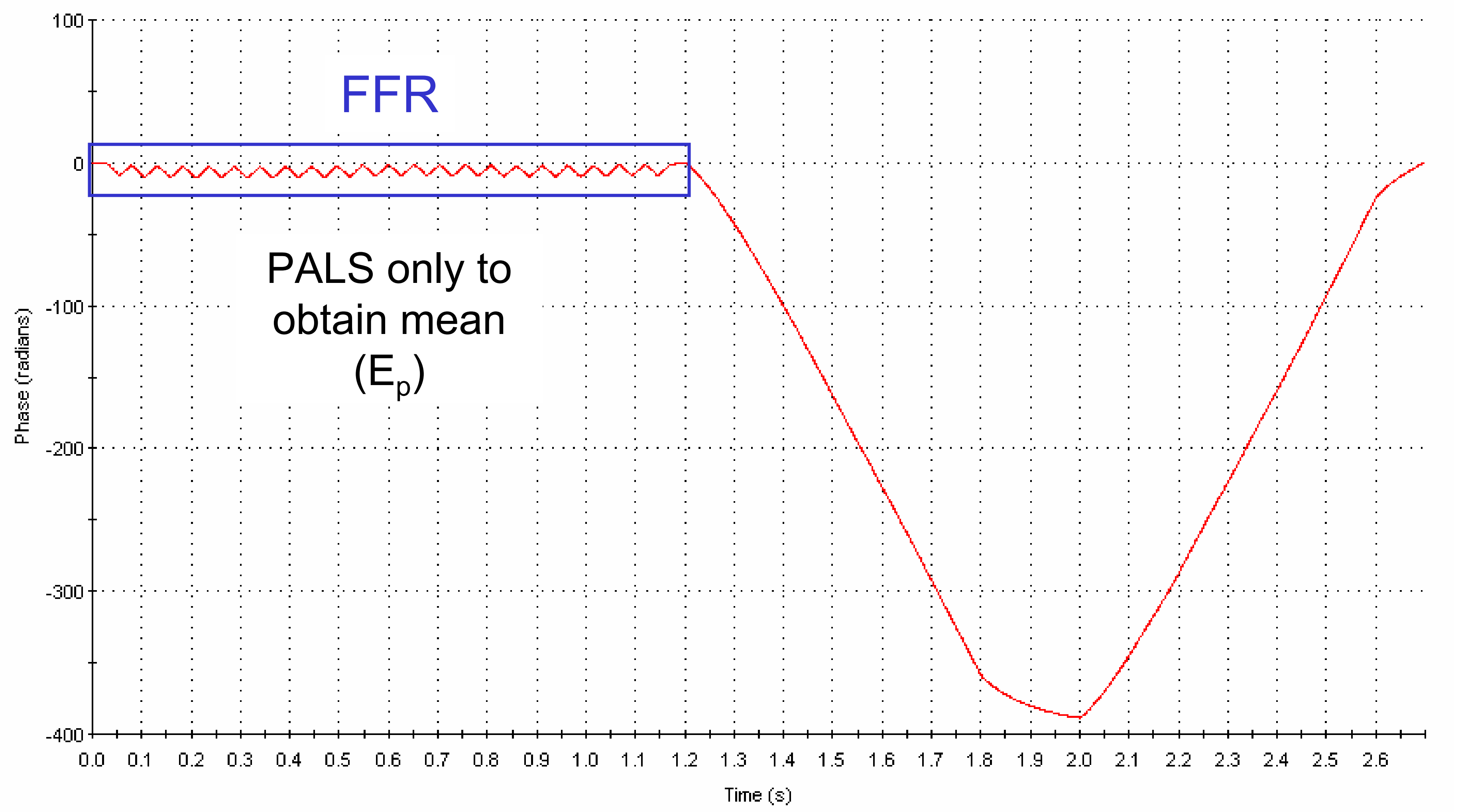
Phase Difference Demonstration

▶

Phase Plot From General Purpose



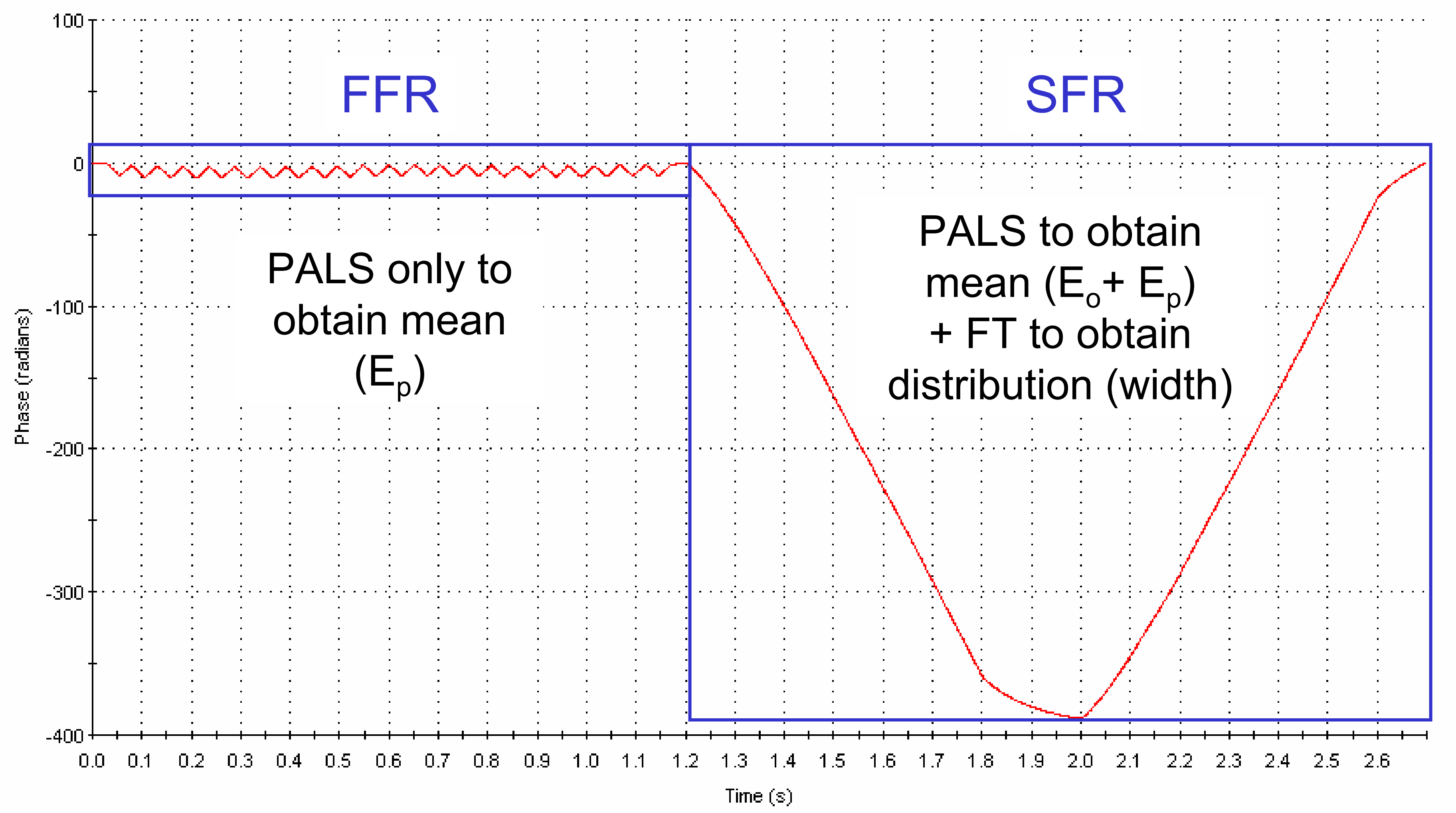
Phase Plot From General Purpose



FFR

PALS only to obtain mean (E_p)

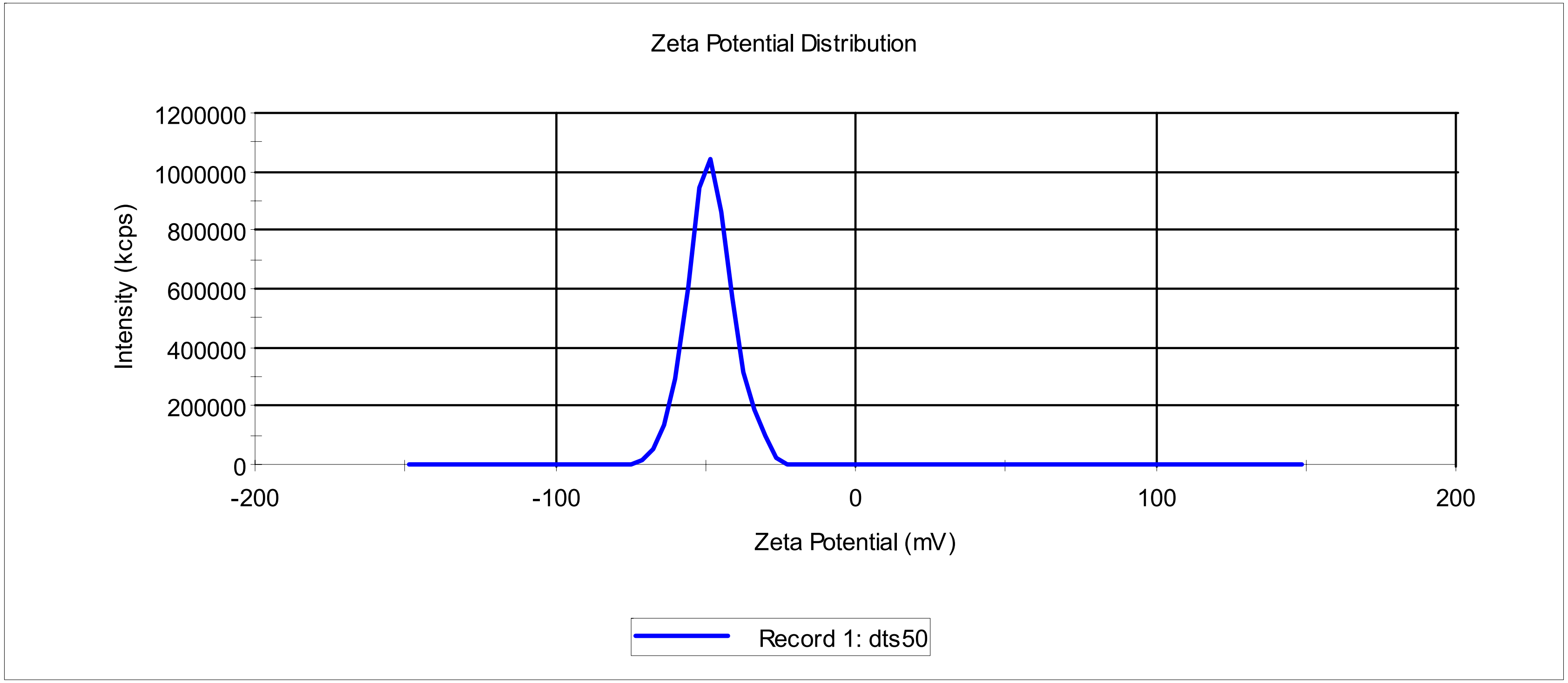
Phase Plot From General Purpose



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



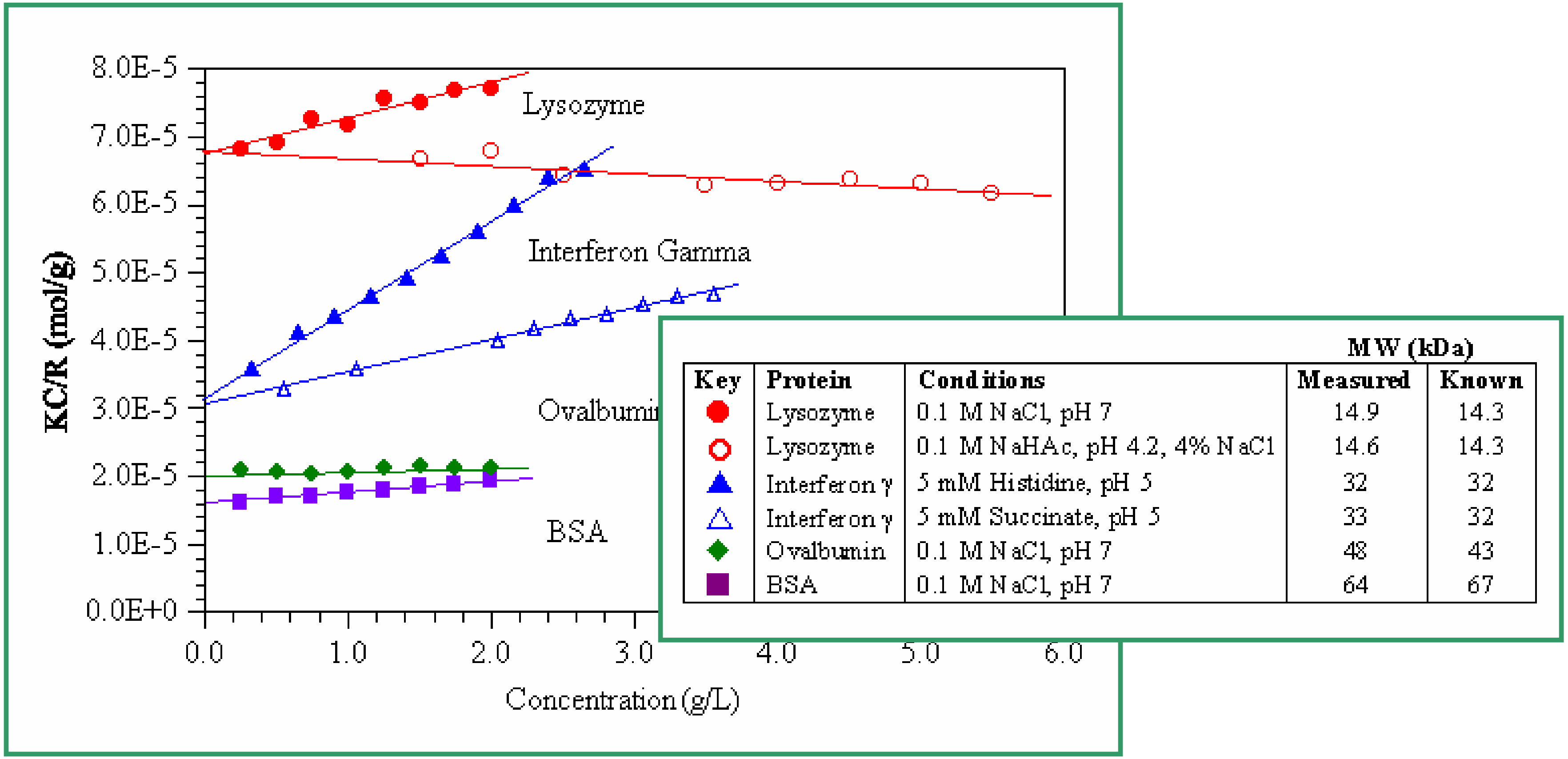
Light Scattering Return

- ☐ 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

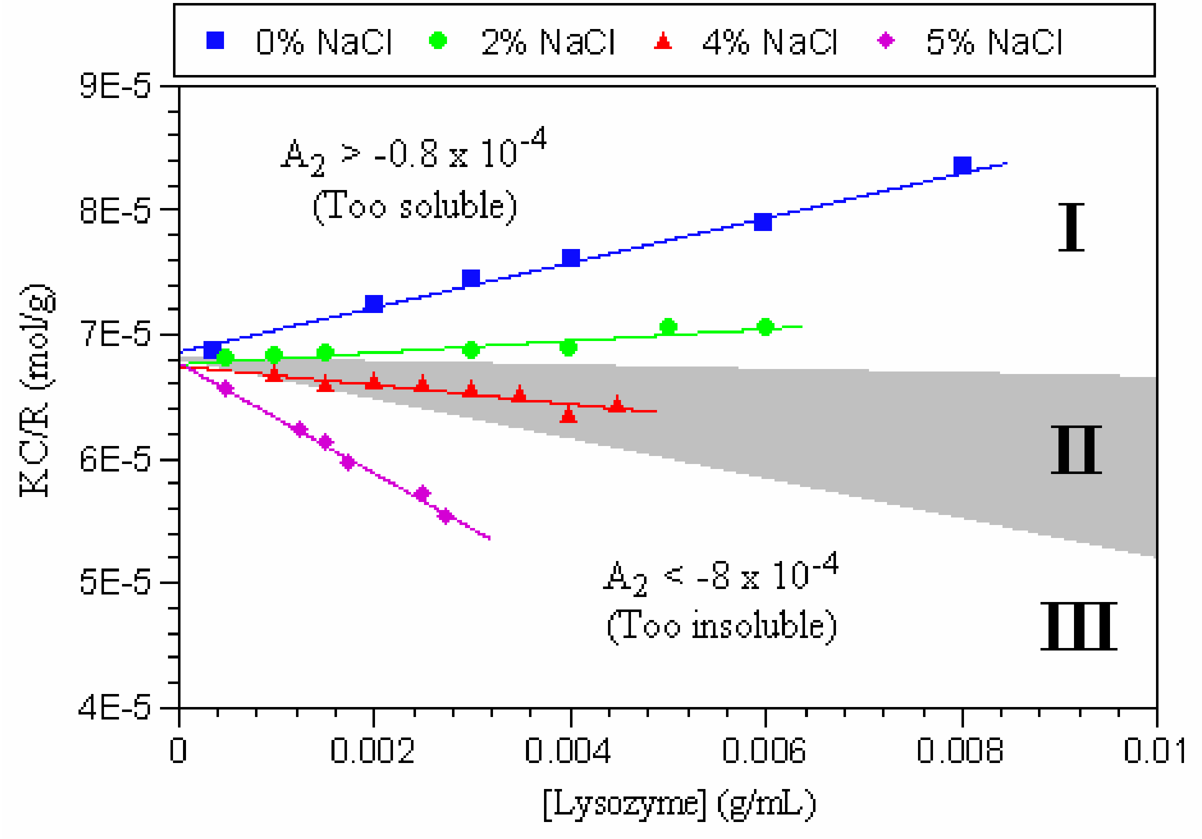


			MW (kDa)	
Key	**Protein**	**Conditions**	**Measured**	**Known**
●	Lysozyme	0.1 M NaCl, pH 7	14.9	14.3
○	Lysozyme	0.1 M NaHAc, pH 4.2, 4% NaCl	14.6	14.3
▲	Interferon γ	5 mM Histidine, pH 5	32	32
△	Interferon γ	5 mM Succinate, pH 5	33	32
◆	Ovalbumin	0.1 M NaCl, pH 7	48	43
■	BSA	0.1 M NaCl, pH 7	64	67

Jones, Mattison “Automated dynamic and static light scattering measurements”, BioPharm 2001, 14(8): 56.

A_2 Crystallization Window

Crystallization Window: $-0.8 > A_2 > -8$ ($\times 10^{-4}$ mol mL / g^2)*



Rayleigh Equation

$$\frac{KC}{R_\theta} = \frac{1}{M} + 2A_2C$$

*George, A; Wilson, W.W. "Predicting protein crystallization from a dilute solution property", *Acta Crystallogr* **1994**, *D50*, 361-365.

Application Example

Molecular Size Measurements

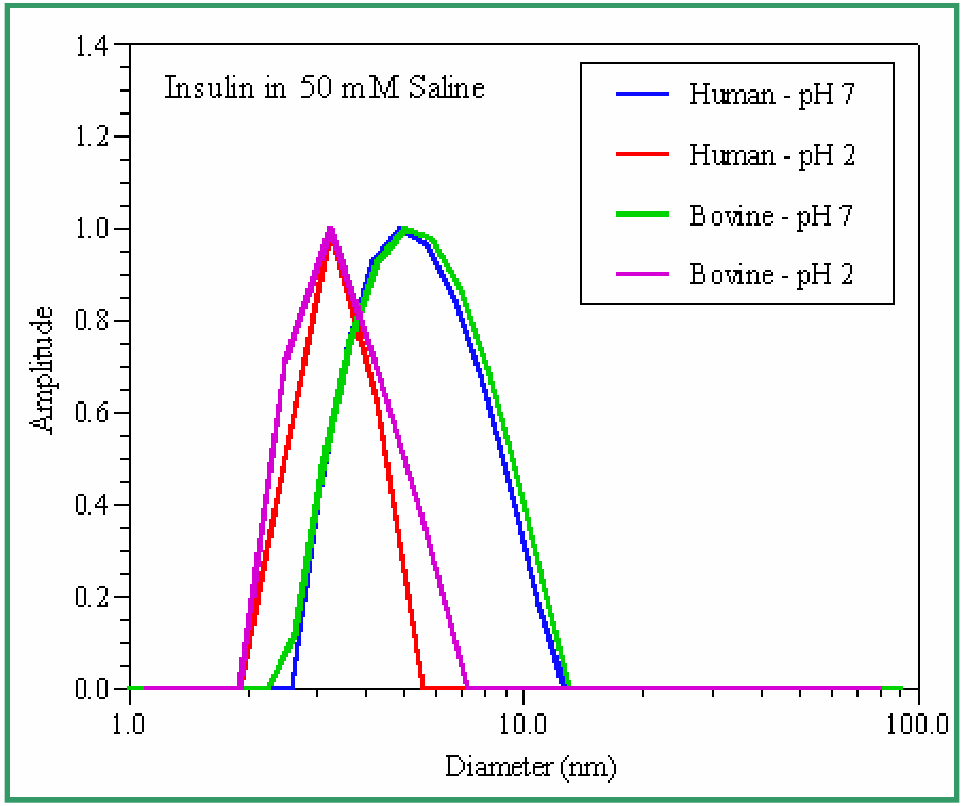
Common Proteins

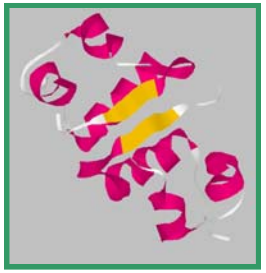
Comparison of known to estimated molecular weight for common proteins.

Protein	Dia (nm)	Known MW (kDa)	Estimated MW (kDa)	%Error
Lysozyme	3.8	14.7	15.1	-2.7
Chymotrypsinogen	4.8	25	26.1	-4.4
Carbonic Anhydrase	5.2	29	31.5	-8.6
Human Insulin (pH 7)	5.4	34.2	34.4	-0.6
Ovalbumin	6	43	44	-2.3
Hexokinase sub-unit	6.6	51	55	-7.8
Hemoglobin	7	65	63	3.1
Bovine Serum Albumin	7.1	67	65.3	2.5
Horse Alcohol Dehydrogenase	7.4	80	71.9	10.1
Amyloglucosidase	7.8	99	81.3	17.9
Hexokinase	8.6	102	102.2	-0.2
Yeast Alcohol Dehydrogenase	9.8	150	138.7	7.5
Apoferritin	16.4	443	462.7	-4.4
Thyroglobulin	20.2	669	753.5	-12.6

Identifying Quaternary Structure

DLS results indicate an insulin structure that is dimeric at pH 2 and hexameric at pH 7, consistent with crystallographic data.





Human Insulin - pH 2
R_H = 1.73
Est MW = 12.1 kDa
Act MW = 11.4 kDa
Dimer

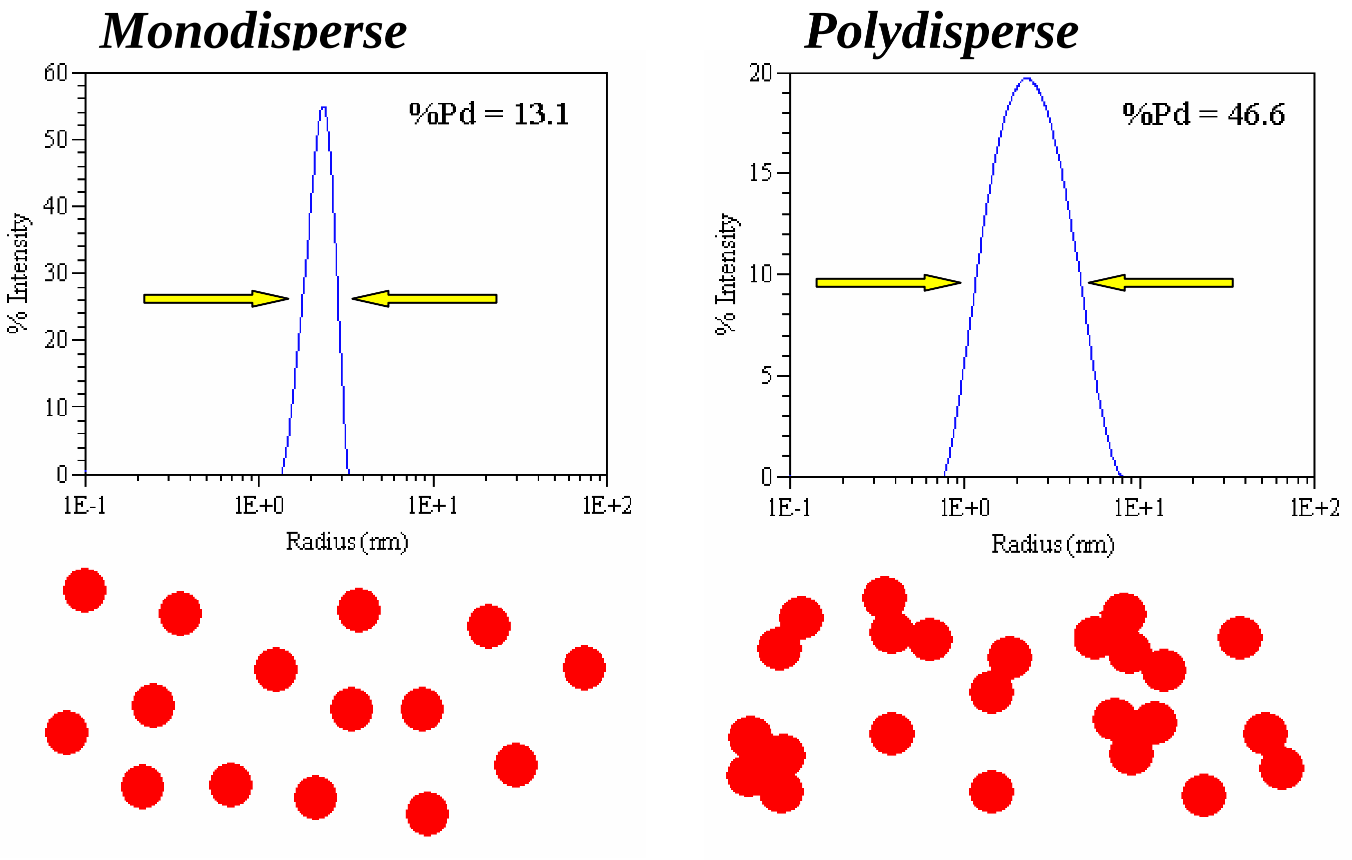
Human Insulin - pH 7
R_H = 2.69
Est MW = 34.1 kDa
Act MW = 34.2 kDa
Hexamer

Application Example

Polydispersity Measurements

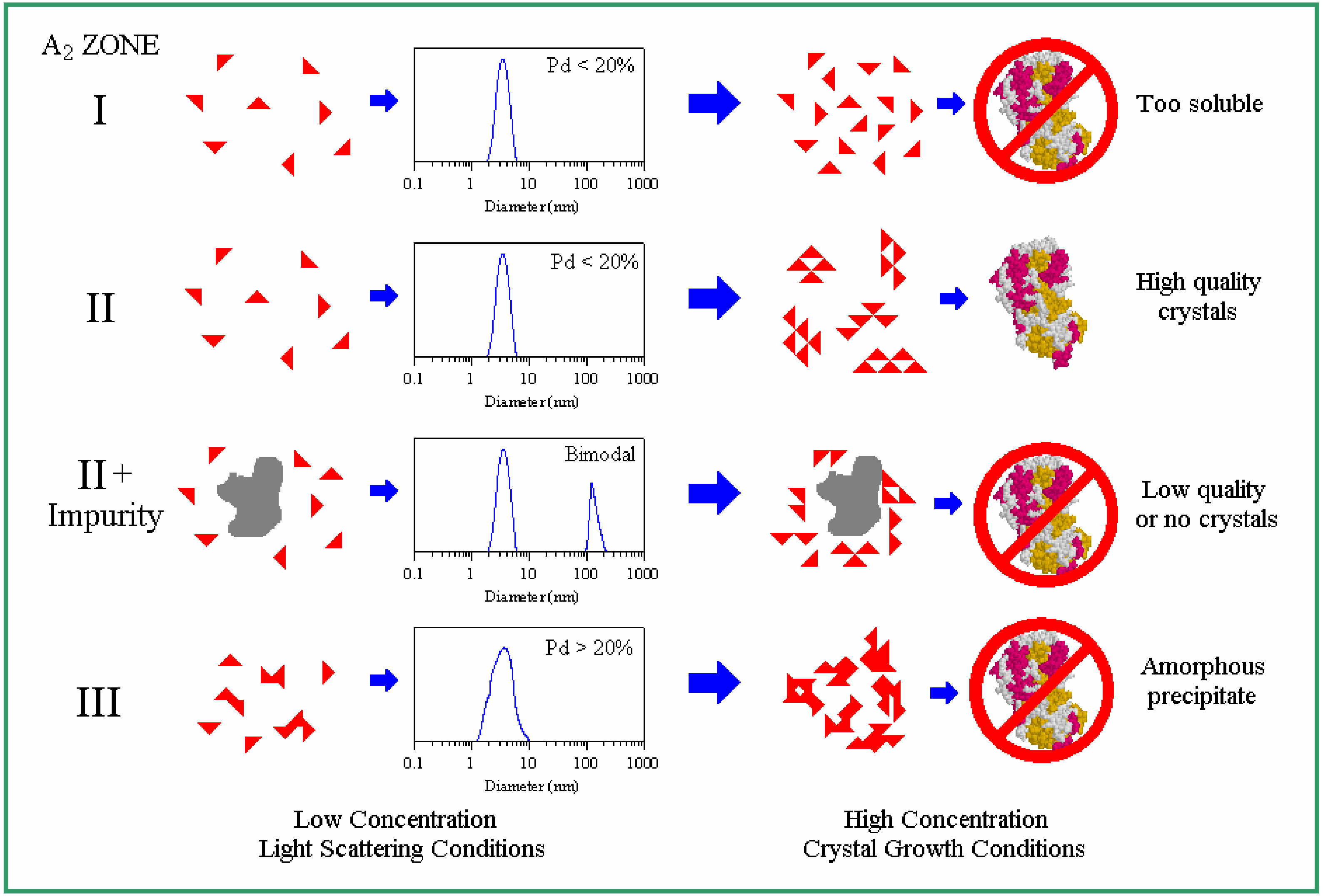
Polydispersity (Pd) From DLS

Pd is representative of the particle size distribution width.



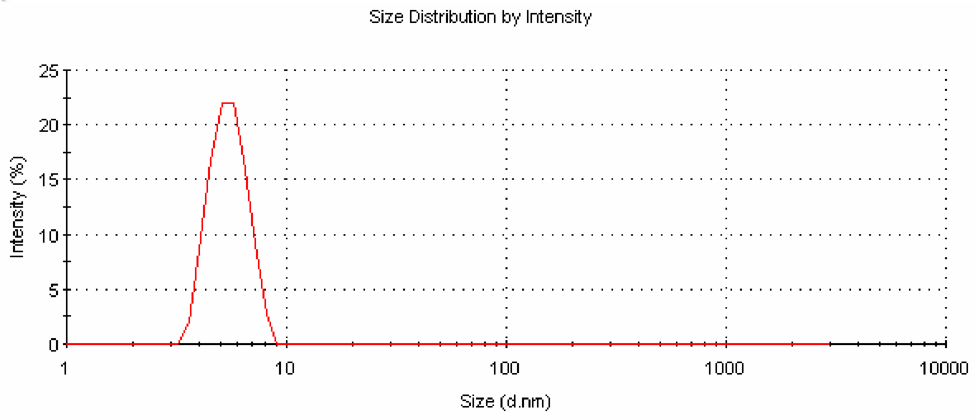
60 second measurement

Crystal Screening Using DLS

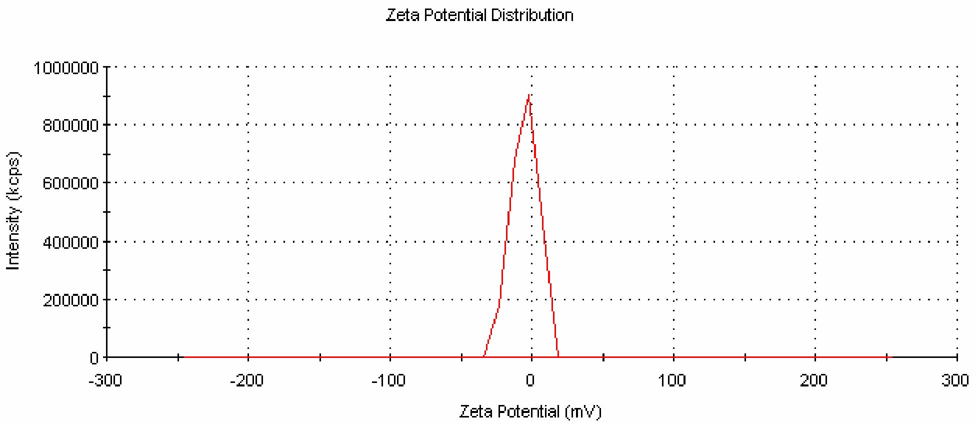


Monoclonal Antibody Fragment

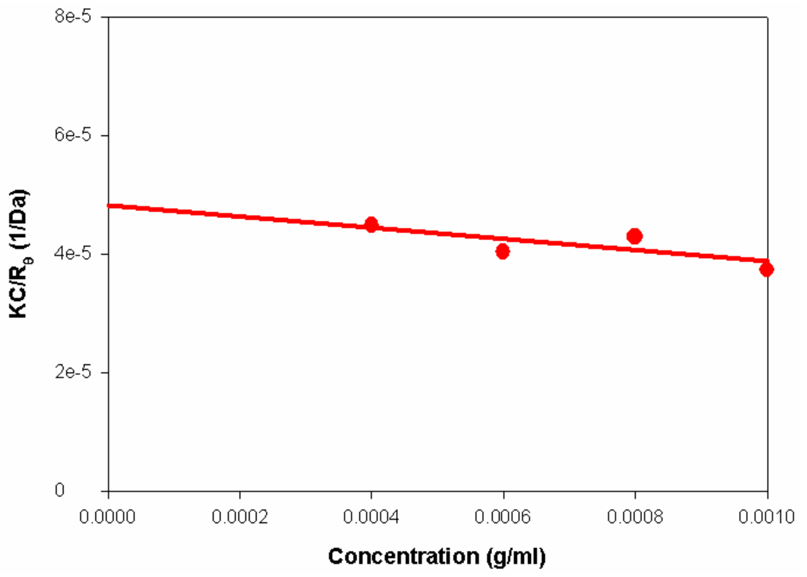
Size, zeta potential and molecular weight



Mean diameter = 5.1nm



Mean zeta potential = -7.6mV



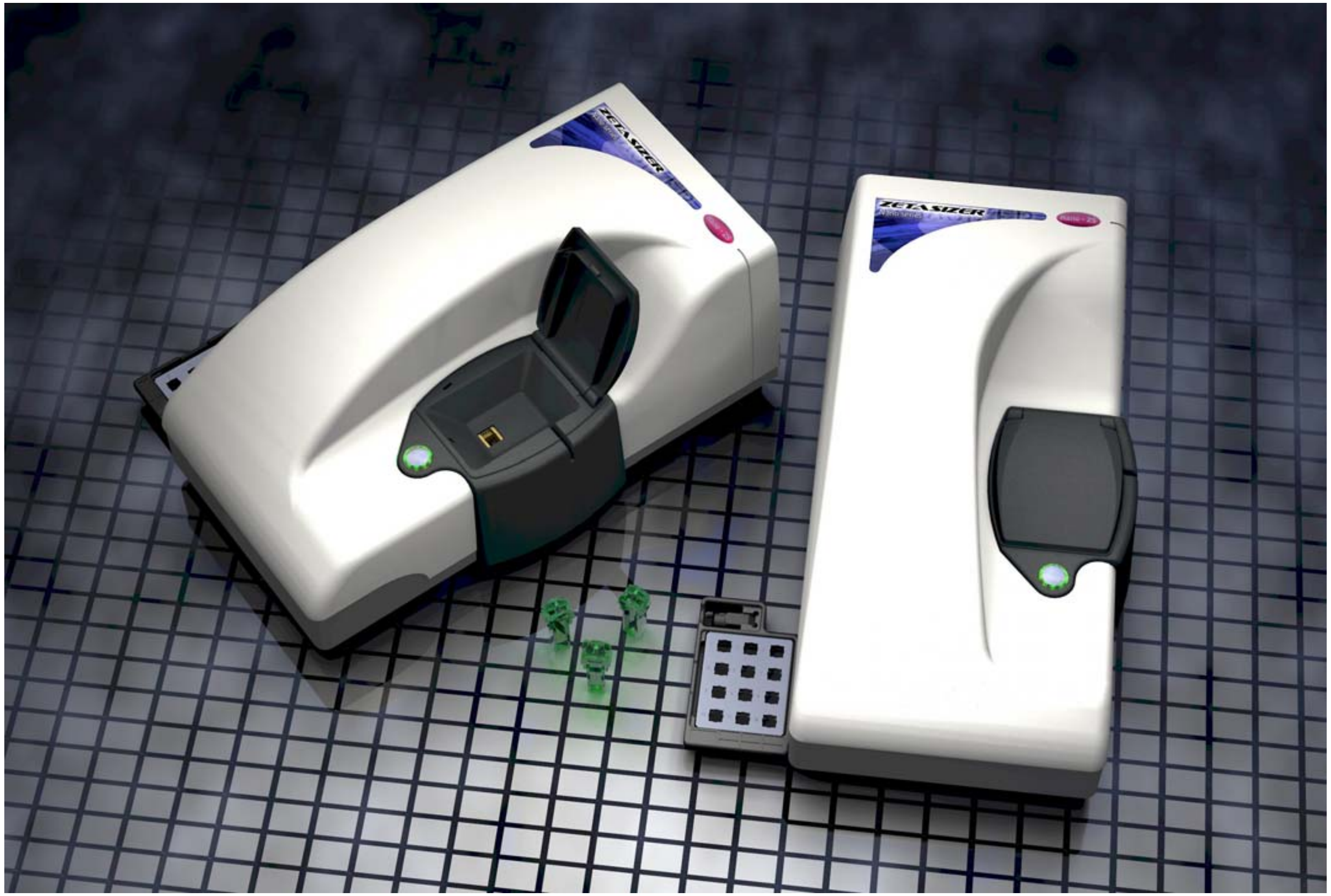
$M_W = 20.7KDa$

$A_2 = -0.0049$ ml mol/g^2

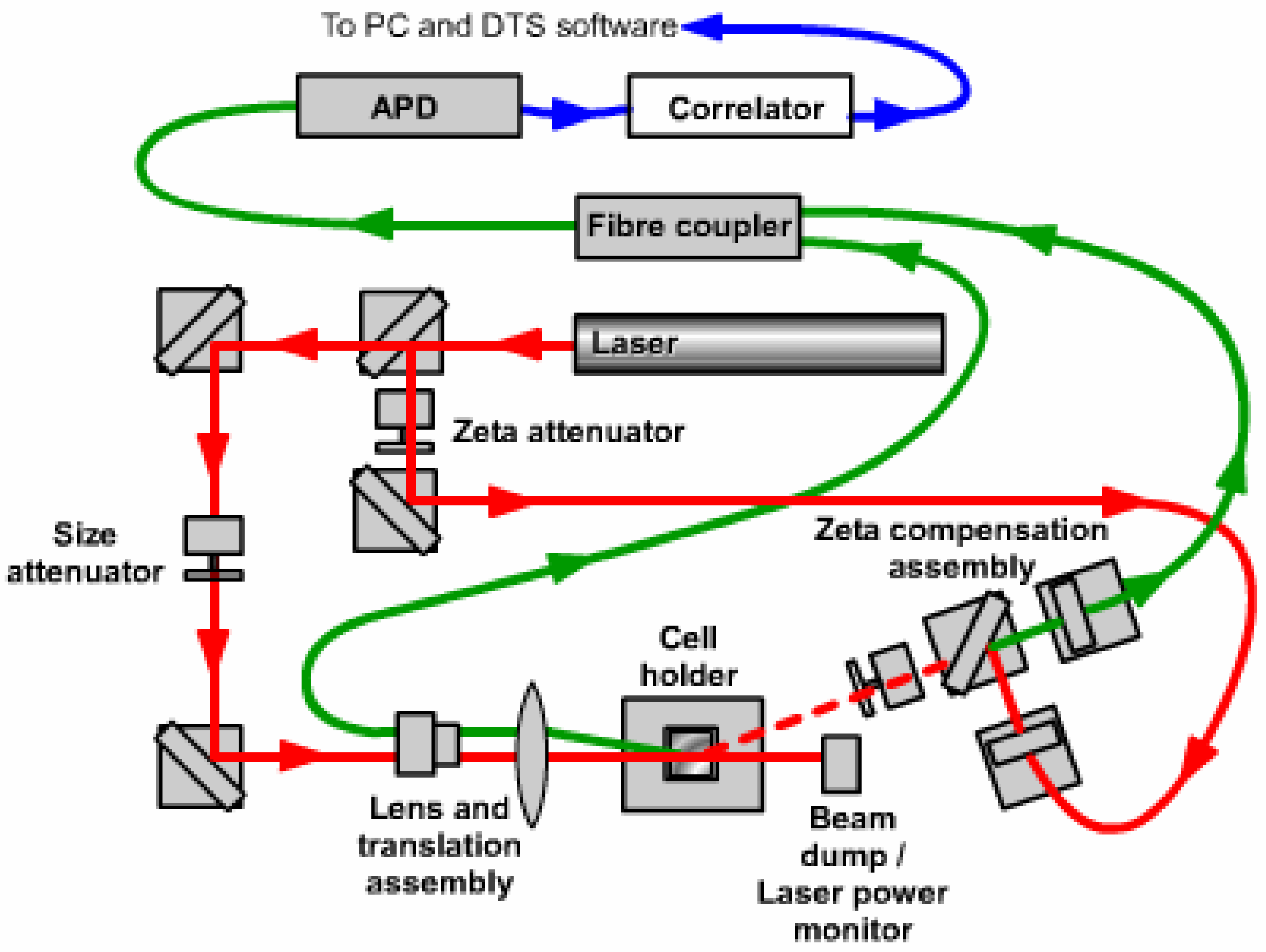
Malvern Instruments

& the Zetasizer Nano

Zetasizer Nano

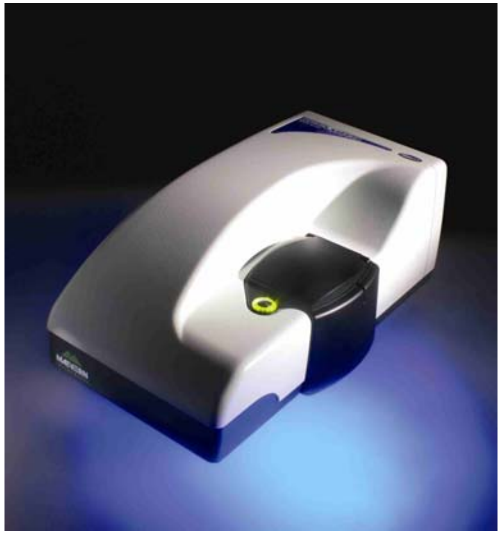


Optics of the Zetasizer Nano



Zetasizer Technical Specifications

Parameter	**Value**
Sizing range	0.6 nm to 6 µm Diam
Concentration range	0.1 mg/mL Lys to 30w%
Min sizing sample volume	12 µL
Min zeta sample volume	0.75 mL
Temperature control	2 to 90 ^{o}C
Conductivity range	0 to 200 mS/cm
Laser	3 mW 633 nm HeNe
Detector	APD



- 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