Technical note NT-Molecular Weight_KIN_EN_V1

Molecular weight characterization using Light Scattering measurements

Key Words: *Average molecular weight, Debye plot, Static Light Scattering (SLS)*

Abstract
This note presents the basic concepts of average molecular weight determination based on light scattering measurements as well as some results obtained on standard materials using the "VASCO Kin" Instrument and its "NanoKin" software.

Introduction

Absolute molecular weight (Mw) measurement is of primary importance for anyone who deals with macromolecules such as proteins or polymers. Indeed molecular weight determines many important physical properties of matter like stiffness, strength, viscosity, transition temperature from liquids to gels, *etc*[1].
Today, there are many known ways to estimate or determine averaged molecular weight such as Gel Permeation Chromatography (GPC), Membrane Osmometry, Sedimentation Velocity/Diffusion, and optical method such Static Light Scattering (SLS). Actually, the intensity of light scattered by a specie depends directly on its overall polarizability, which is itself, related to its molecular weight (or its size). The instrument "VASCO Kin" allows to measure the average molecular weight using such an optical method and features a dedicated calculation tool included in its software "NanoKin". We present its principle hereafter illustrated by some results obtained with standard proteins and polymers.

SLS Molecular weight measurements: principle and theory

Static Light Scattering is a well-known characterization technique used in chemistry and macromolecules research which principles has been described many decades ago[2,3]. It is based on the fact that a portion of a laser beam focused into a sample is scattered in all directions by the macromolecules or particles dispersed in the liquid (see figure 1). Yet, this scattered light intensity depends on one hand, the sample properties, such as the molecular weight of solubilized species, their refractive index contrast, and their concentration for instance, and on the other hand, the optical set-up characteristic, such as the observation angle, the distance between the detector and the sample cell, the quantum efficiency of the detector, the scattering volume, *etc*.

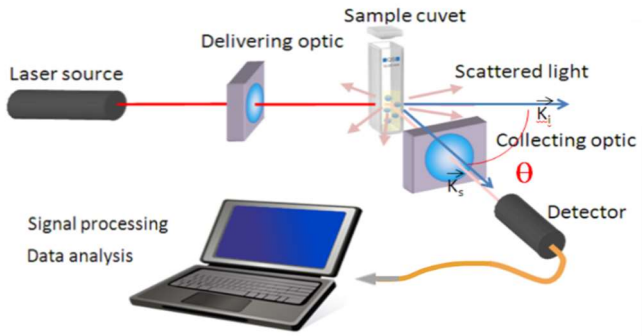

Figure 1: General set up of a light scattering experiment.

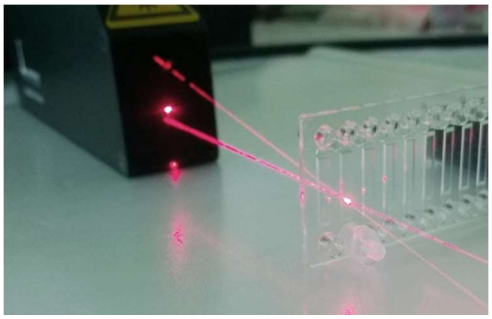

Figure 2: in-situ optical head of "VASCO Kin".

[1] *C. Tanford (1961), "Physical Chemistry of Macromolecules". John Willey & Sons-ISBN 0 471 84447 0*
[2] *P. Debeye (1944), "light scattering in solutions". J Appl Phys. 15 (4) :338*
[3] *B.H. Zimm (1945). "Molecular Theory of the scattering of lights in fluids". J. Chem . Phys 13 (4) :141*

Then, it has been shown that the Rayleigh ratio **R(θ)** of a sample, related to the ratio of the scattered intensity to the incident intensity for a given scattering angle θ, can be approximated by the Zimm/Rayleigh equation:

$$\frac{KC}{R(\theta)} = \frac{1}{M_W\, P(\theta)} + \, 2\, A_2 C + \cdots \tag{1}$$

Where $\mathbf{M_w}$ is the (weight averaged) molecular weight and **C** the macromolecules concentration.
K is called the optical contrast constant and is defined as:

$$K = \frac{4\pi^2\, n_0^2\, (\frac{dn}{dC})^2}{N_A\, \lambda^4} \tag{2}$$

With $\boldsymbol{n_0}$ the refractive index of the solvent, $\boldsymbol{\lambda}$ the laser wavelength, $\boldsymbol{\frac{dn}{dC}}$ the differential refractive index increment of the solution upon the macromolecules concentration and $\boldsymbol{N_A}$ is the Avogadro number (6.023 x 10^{23}).
P(θ) is the corrective form factor that we have to consider with large particles/molecules with respect to the laser wavelength (> $\lambda/20$). When characterizing smaller objects like proteins or macromolecules, this factor can be approximated by 1 (Rayleigh domain). Note that it is precisely in this operating regime that we should use the "VASCO kin". Actually, addressing the Mw estimation of larger objects implies to determine P(θ) through the light intensity measurement at many angles (typically 6 or more) or to calculate it theoretically knowing the structure of the objects (Gaussian coil, rigid, sphere, rod..). Then, in the next part of this note, we assume the form factor equal to 1.
Finally, $\mathbf{A_2}$ is the second virial coefficient, a corrective factor for non-ideal solution related to the macromolecules-solvent interactions. The value and the sign of this coefficient is directly proportional to the affinity of the macromolecules with the solvent[4].

The Rayleigh ratio **R(θ)** has to be determined experimentally. In principle, it can be calculated by measuring the ratio of scattered light intensity on the incident light intensity at given angle and distance r (and normalized by unit of scattering volume and unit of scattering solid angle). In practical, these intensities of light are very difficult to measure due to multiple instrumental factors. In order to circumvent their tedious determination, one measures intensity ratios with respect to a reference sample, typically toluene for example. Knowing the theoretical Rayleigh ratio of toluene $\mathbf{R_{tol}}$ we can then write the following equation:

$$R(\theta) = \frac{I_A(\theta) - I_0(\theta)}{I_{tol}(\theta)} R_{tol} \left(\frac{n}{n_{tol}}\right)^2 N(\theta, n) \tag{3}$$

Where $\boldsymbol{I_A(\theta)}$ is the intensity scattered by the sample at the angle θ, $\boldsymbol{I_0(\theta)}$ and $\boldsymbol{I_{tol}(\theta)}$ the intensity scattered by the solvent only and the toluene in the same conditions (respectively the background and calibration measurement). $\boldsymbol{n}$ is r the refractive index of the sample and $\boldsymbol{n_{tol}}$ the one of toluene. The $\left(\frac{n}{n_{tol}}\right)^2$ term is involved in this equation to correct the solid angle normalization of scattered light between toluene and the sample.
$\boldsymbol{N(\theta, n)}$ is another correction factor related to the change of scattering volume between toluene and the sample and highly depends on the optical set-up geometrical configuration (square or round cuvette, *etc.*).

---

[4] For A_2 >0, Macromolecules/particles have a good affinity for the solvent and the colloid should be rather stable (and polymers form large/spread random coils). For A_2 <0, the macromolecules have a better affinity for themselves than for the solvent and then tends to aggregate (and polymers form dense random coils). When A_2 =0, the strength of macromolecules-solvent interaction is the same magnitude as that of macromolecules - macromolecules interaction (Theta solvent).

Debye plot: analysis and interpretation

Equation (1) shows that the Rayleigh ratio has, at first approximation, a linear dependence with concentration C. Then, if one plots KC/R(θ) vs C at a constant scattering angle θ, we should in principle obtain a line having a slope equal to 2.$\mathbf{A_2}$ and an intercept at C= 0 equal to $\frac{1}{M_W}$ (assuming P(θ)= 1) as shown on figure 3. This representation is called the Debye plot.

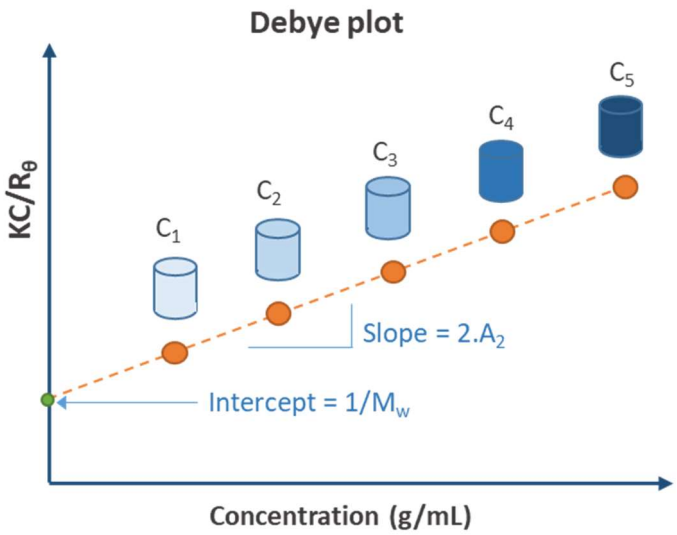

Figure 3: Debye plot. Values of KC/R(θ) are plotted as a function of sample concentration. The slope of the corresponding linear regression is equal to 2xA_2 while the intercept at C=0 is equal to 1/M_w.

Basically, the Mw measurement principle consists in measuring the average sample scattered intensity at different concentrations. Such a light intensity measurement has also to be performed for the solvent alone (background measurement) and for toluene or another reference material (calibration measurement). Then, for each concentration, the corresponding ratio KC/R(θ) is calculated using (2) and (3) and plots on a graph versus the concentration C. A linear regression fit allows to determine the intercept and the slope of the straight line from which Mw and A2 are deduced.

Examples of Mw measurement using the VASCO Kin

Bovine serum Albumin (BSA)

BSA is a small protein which is commonly used as a standard. According to literature, ideal monomeric albumin has a molecular weight of 66.5 kDa determined by mass spectrometry[5] but higher values, up to 69kDa, have also been reported probably due to the co-presence of dimers or higher oligomers in solution[6]. BSA shows a natural tendency to dimerize depending on the buffer properties (ionic strength, pH…)[7].

The BSA used in this study is provided by Sigma-Aldrich under lyophilized powder form (≥ 96%, ref: A6003). This powder has been dissolved at different concentrations between 1 and 10 mg/mL in a 0.2µm-filtrated 0.1M KCL buffer. Measurements of both hydrodynamic diameters (DLS) and average scattered intensity (SLS) have been performed using "Vasco Kin" with its *in-situ* optical head in optical glass cuvette at 25°C (see figure 2).
Alongside with the samples measurements, the characterization of the solvent background, and filtered toluene were carried out in order to calibrate the instrument.

---

[5] *Hirayama, K., et al., Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS, BBRC, 173(2), 639 (1990).*
[6] *Scatchard, G., et al., Preparation and properties of serum and plasma proteins; osmotic equilibria in solutions of serum albumin and sodium chloride J. Amer. Chem. Soc. 68, 26.0 (1946).*
[7] *D.C. Carter, J.X. Ho, Structure of serum albumin, Adv. Protein Chem.45 (1994) 153–203.*

Molecular weight calculation processes the average intensity of light scattered by the studied species in solution. However, some samples may contain several populations of macromolecules/particles which all participate to the overall amount of light scattered by the sample and, then, affect significantly the result. This effect is well illustrated for the case of BSA sample as shown below.
Actually, DLS size characterization using the Vasco Kin shows that at least two populations are detected in these BSA solutions (see figure 4): a population having a hydrodynamic diameter close to 7nm, likely to contain mainly BSA monomers, and a larger population around 20nm, either oligomers, aggregates or impurities, present in low amount.

In order to improve the calculation accuracy of the monomers molecular weight, the instrument software allows to select the population at 7nm to focus the processing to its sole scattering intensity.

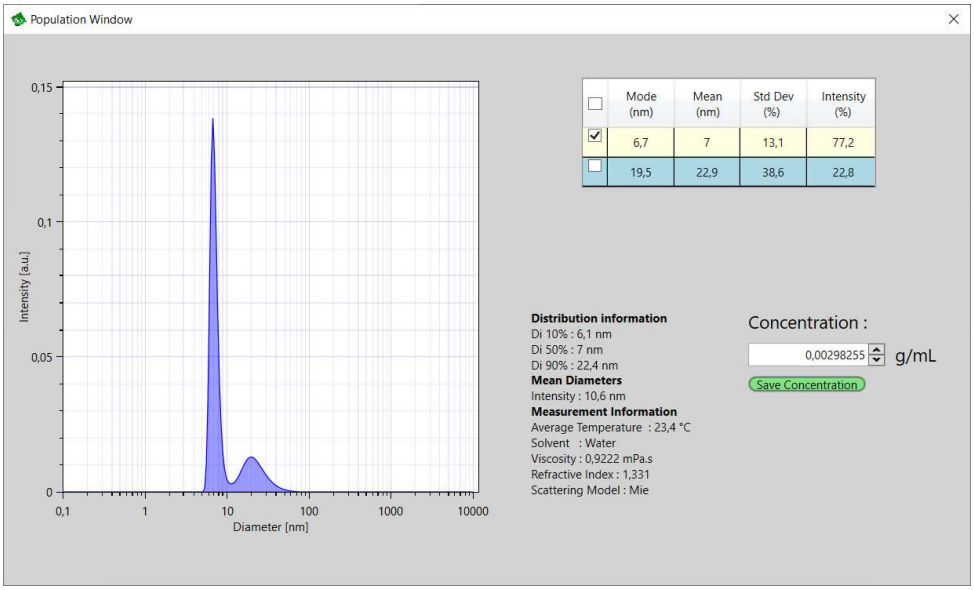

Figure 4: "NanoKin" software panel showing the DLS sizes measurement of BSA in 0.1M KCl buffer. This window also allows to select on which population is processed the Mw calculation.

By using this functionality, a dn/dc of 0.190 mL/g[8] and a $R_{tol}(\theta)$ toluene of 1.402E-5 cm-1[9], the software draws the corresponding Debye plot (see figure 5) and calculates the Mw value at **65.7kDA** with a virial coefficient of **3.0E-4 mol/g**. As seen above, this result is well in the range of expected values for such a protein.

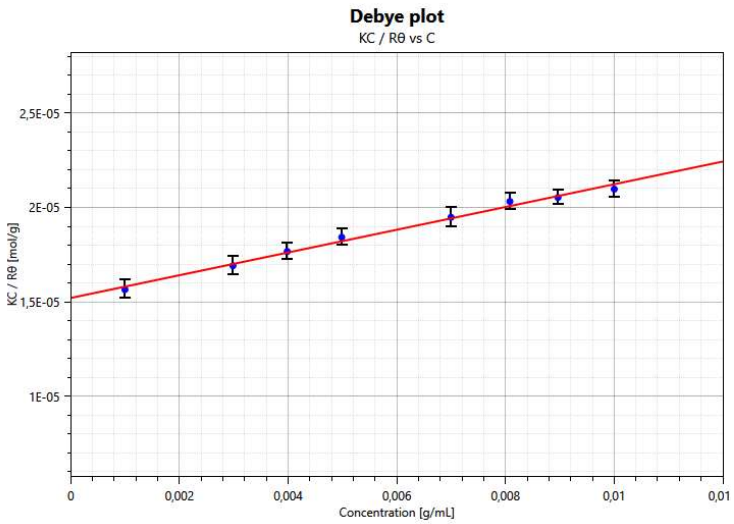

Figure 5: Debye plot of BSA in 0.1M KCl showing Values of KC/R(θ) are plotted as a function of sample concentration (blue dots). The red line show the corresponding linear regression.

[8] *A. Theisen and al, Refractive Increment Index Data-Book, Nottingham University Press.*

[9] *W. Kaye and al. Low-angle laser light scattering: Rayleigh factors and depolarization ratios. Appl. Opt., 13:1934–1937, 1974. 27, 96.*

For each sample, the standard deviation of the average scattered light intensity measurement is displayed on the graph by an error bar (between 2 and 3% in this case). The overall error estimations is determined at 2.6kDa for the molecular weight and 5E-5 mol/g for the virial coefficient. Note that this calculation is performed using the method of maximum slops (as define in ISO17025) and mainly involves the results obtained at minimum and maximum concentrations. Then, the accuracy of these two measurements is critical for such an overall error estimation.

Others species in water

The procedure described above for the BSA molecular weight measurement have been performed for others species soluble in water: Lysozyme proteins from chicken egg white, having a theoretical molecular weight of 14.3 kDa (ref 2970, from Sigma-Aldrich) and three different Polyethylene Glycol polymers (PEG) having nominal molecular weights of 6kDa, 20kDa, and 40kDa (respectively, reference 81293, 81298 and 96699 from Sigma-Aldrich).
In each case, the samples have been prepared at different concentrations by dispersing the raw powder in filtered 0.1M KCl Buffer for Lysozymes and in pure water for PEGs.
The results have been calculated using a dn/dc of 0.183 mL/g[10] for the Lysozyme and 0.134 mL/g for the PEGs[11]. The corresponding Debye plots are shown on figure 6. Resulting molecular weights and virial coefficients are reported in the table 1 below.

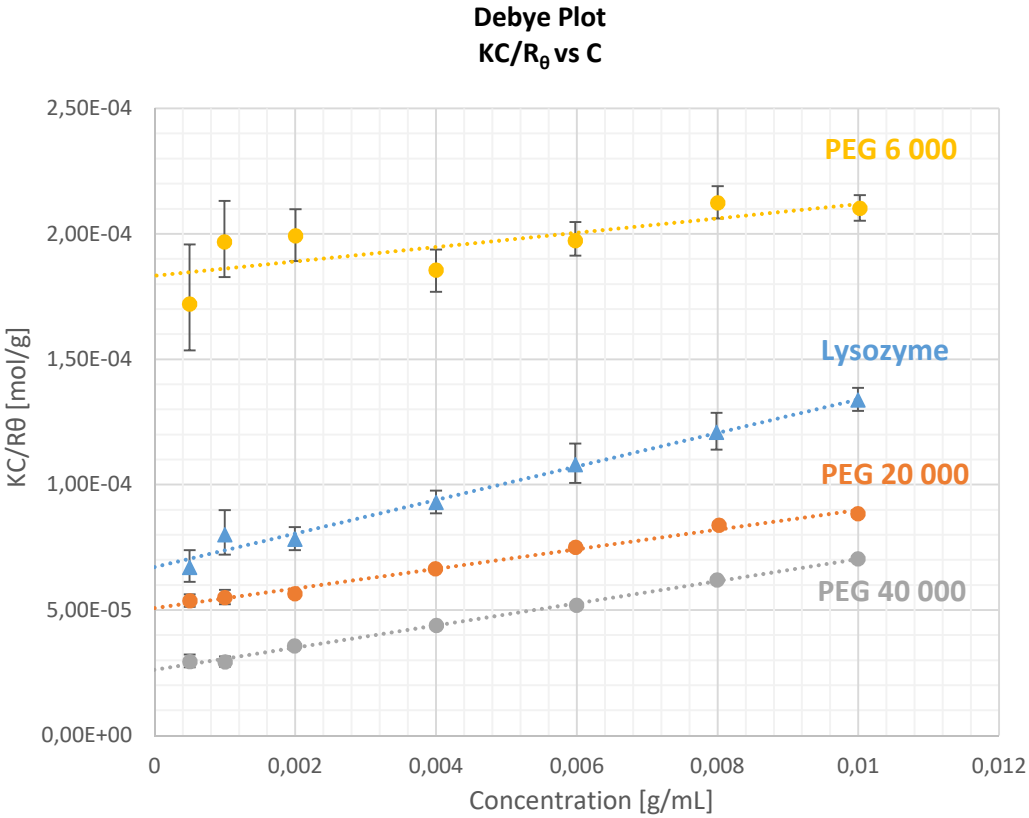

Figure 6: Debye plot showing values of KC/R(θ) are plotted as a function of sample concentration for PEG 6 000 in water (yellow dots), Lysozyme in 0.1M KCl (blue triangles), PEG 20 000 in water (orange dots), and PEG 40 000 in water (grey dots). The dash lines show the corresponding linear regressions.

[10] *Narayanan, J., & Liu, X. Y. (2003). Protein Interactions in Undersaturated and Supersaturated Solutions: A Study Using Light and X-Ray Scattering. Biophysical Journal, 84(1), 523–532.*
[11] *Hasse H., et al. Osmotic Virial Coefficients of Aqueous Poly(ethylene glycol) from Laser-Light Scattering and Isopiestic Measurements, Macromolecules 1995, 28, 10, 3540–3552.*

	Mw (kDa)	ΔMw (kDa)	A_2 (mol/g)	ΔA_2 (mol/g)
Lysozyme	**14.9**	1.4	**3.34E-03**	5.75E-04
PEG 6 000	**5.6**	0.7	**1.76E-03**	1.38E-03
PEG 20 000	**19.7**	1.0	**1.96E-03**	2.06E-04
PEG 40 000	**38.1**	1.2	**2.20E-03**	1.14E-04

Table 1: Molecular weights (M_w) and virial coefficients (A_2) calculated from the measurements shown in figure 6. ΔMw and ΔA_2 are the overall measurement errors estimated by the method of maximum slops.

These results are in good agreement with what is expected for their molecular weights. Note that despite a very low intensity scattered by the diluted PEG 6 000 (which results in the larger error bars observed on the corresponding Debye plot), the sensitivity of the Vasco Kin still allows a reliable measurement of its molecular weight. On the other hand, the high ΔA_2 value estimated for the PEG 6 000 calls for more caution on its second virial coefficient calculation.

Conclusion

The results shown in this note show that the molecular weight of various macromolecules have been successfully measured using the Vasco kin and its dedicated software tool. These measurements are in good agreement with the expected values and were achievable within the range of 10% maximum error estimation for most of tested proteins and polymers. Only the PEG 6000 shows a higher estimated error of 12.5% due to the very low light scattering properties of such a specie in water. Despite of that, the molecular weight calculated for this polymer (5.6kda) remains quite close to its nominal value (6kDa).
We also observe with the BSA measurement that a multi-populated sample can be correctly characterized by using the corresponding DLS size distribution to focus the algorithm processing on the population of interest (the population of monomers in the case of BSA). This functionality featured in the" Nanokin" software appears to be critical in such an optical measurement technic which can be so sensitive to the presence of large aggregates or inhomogeneities in the studied samples.