

Teaching Light Scattering Spectroscopy: The Dimension and Shape of Tobacco Mosaic Virus

Nuno C. Santos and Miguel A. R. B. Castanho

Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade de Lisboa, Ed. C1-5°, Campo Grande, 1700 Lisboa, and Centro de Química-Física Molecular, Complexo I, Instituto Superior Técnico, 1096 Lisboa codex, Portugal

ABSTRACT The tobacco mosaic virus is used as a model molecular assembly to illustrate the basic potentialities of light scattering techniques (both static and dynamic) to undergraduates. The work has two objectives: a pedagogic one (introducing light scattering to undergraduate students) and a scientific one (stabilization of the virus molecular assembly structure by the nucleic acid). Students are first challenged to confirm the stabilization of the cylindrical shape of the virus by the nucleic acid, at pH and ionic strength conditions where the coat proteins alone do not self-assemble. The experimental intramolecular scattering factor is compared with the theoretical ones for several model geometries. The data clearly suggest that the geometry is, in fact, a rod. Comparing the experimental values of gyration radius and hydrodynamic radius with the theoretical expectations further confirms this conclusion. Moreover, the rod structure is maintained over a wider range of pH and ionic strength than that valid for the coat proteins alone. The experimental values of the diffusion coefficient and radius of gyration are compared with the theoretical expectations assuming the dimensions detected by electron microscopy techniques. In fact, both values are in agreement (length ≈ 300 nm, radius ≈ 20 nm).

INTRODUCTION

Light scattering techniques are very easy to perform and are very useful in studies about the structure of macromolecules and molecular assemblies. Although its application is becoming more common, suggestions for introductory laboratory work for undergraduate students are very scarce and deal only with static light scattering (e.g., Thompson et al., 1970, Matthews, 1984, Mougán et al., 1995). We present laboratory work for undergraduate students where light scattering techniques are used in multiple ways to characterize a molecular assembly.

The tobacco mosaic virus (TMV) was the chosen model molecular assembly because: 1) it has a very well-defined geometry; 2) it is not spherical (due to symmetry reasons, there is a tendency to emphasize spherical geometries too much); 3) it is very well characterized in terms of size and shape by means of independent techniques (for a review see e.g., Caspar, 1963); and 4) it is a fairly monodispersed system (both size and shape).

Moreover, the molecular assembly of TMV coat proteins is largely characterized in terms of structure changes with ionic strength (I) and pH (Fig. 1, e.g., Butler and Mayo, 1987). According to pH and I , TMV coat proteins can remain as separated monomers, self-assemble into disks, or self-assemble into disk stacks (rods). Such dramatic structure alterations are easily detected by means of light scattering spectroscopy techniques.

Can the nucleic acid stabilize the cylindrical shape of the virus in pH and I conditions where the coat proteins alone cannot? This is the first question students are challenged to answer.

THEORETICAL BACKGROUND

Light scattering intensity is monitored either in the microsecond or in the second time range domain. This is the basic difference between dynamic light scattering (DLS) and static light scattering (SLS), respectively. Fluctuations in the intensity of light scattered by a small volume of a solution in the microsecond time range are directly related to the Brownian motion of the solute. Averaging the intensity over the second time range interval will cause a loss of the solute dynamic properties information; that is why light scattering is named either static or dynamic.

The outlines of the theory related to light scattering techniques is described in biophysics (e.g., Brunner and Dransfeld, 1983; Marshall, 1978), chemistry (e.g., Oster, 1972), and polymer science textbooks (e.g., Munk, 1989). Introductory textbooks and review articles on light scattering applications in biochemistry are also available (e.g., Harding et al., 1992, Bloomfield, 1981). We shall only briefly describe some basic aspects.

The light scattering intensities are recorded according to the measurement angle (Fig. 2) and concentration.

Static light scattering

Light scattering intensity integrated over a period of time of seconds or more varies with the measurement angle (θ) and concentration according to (Zimm, 1948):

$$\frac{Kc}{R_\theta} = \frac{1}{MP_\theta} + 2A_2c \quad (1)$$

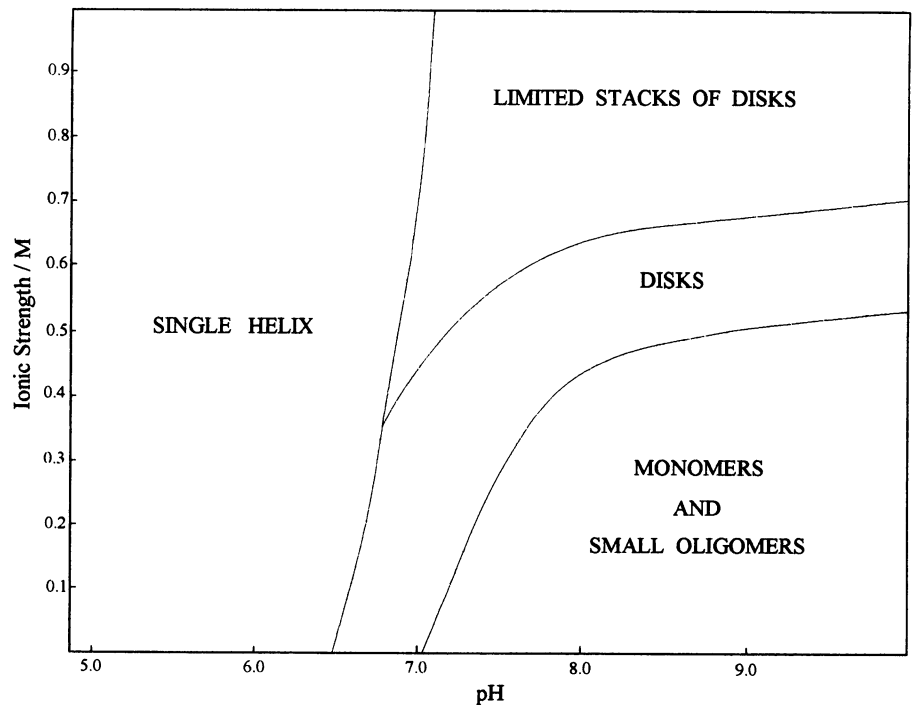
Received for publication 1 April 1996 and in final form 4 June 1996.

Address reprint requests to Dr. Miguel A. R. B. Castanho, Centro de Química-Física Molecular, Complexo I, Instituto Superior Técnico, 1096 Lisboa codex, Portugal. Tel.: 351-1-8419248; Fax: 351-1-3524372; E-mail: prieto@alfa.ist.utl.pt.

© 1996 by the Biophysical Society

0006-3495/96/09/1641/10 \$2.00

FIGURE 1 TMV coat proteins self-assembled structures according to pH and ionic strength (adapted from Butler and Mayo, 1987). In our study the whole TMV was studied at pH 5.0, 7.2, and 10.0 and $I = 0, 50, \text{ or } 200 \text{ mM}$.



where,

$$K = \frac{4\pi^2 n_o^2 (dn/dc)^2}{N_A \lambda^4} \quad (2)$$

$$R_\theta = \frac{d^2 I_s}{\sin^2 \theta_z I_o} \quad (3)$$

c represents concentration, I_o is the intensity of the incident light (vertical polarization), I_s is the scattered light intensity, θ_z is the measurement angle relative to the vertical axis, d is the sample - detector distance, n_o is the refractive index of the solvent, n is the refractive index of the solution, A_2 is the second virial coefficient (which accounts for interparticle interaction), M is the molecular weight, and,

$$P_\theta = \frac{I_{s,\theta}}{I_{s,\theta=0}} \quad (4)$$

is the intra-particle structure factor, which accounts for the interference of light scattered from different points in the same molecule or molecular assembly. P_θ can be evaluated by (e.g., Oster, 1972):

$$P_\theta \approx 1 - \frac{16\pi^2 n_o^2 R_g^2}{3\lambda^2} \sin^2\left(\frac{\theta}{2}\right) \quad (5)$$

where R_g is the radius of gyration.

Alternatively R_θ can be calculated by a calibration method using, for instance, benzene (e.g., Chu, 1991):

$$R_\theta = R_{\text{benzene},90^\circ} \frac{n_o^2}{n_{\text{benzene}}^2} \frac{I_s}{I_{s,\text{benzene}}} \quad (6)$$

($R_{\text{benzene},90^\circ} = 8.51 \times 10^{-6} \text{ cm}^{-1}$; Pike et al., 1975).

What happens if samples are polydisperse?

If the samples are polydisperse, than the values of M , R_g , and A_2 obtained by means of Eq. 1 are averaged. Textbooks and scientific papers often mention the averaged M ($\langle M \rangle$) as the weight average molecular weight (M_w) but a demonstration is hard to find. Moreover, what kind of average is obtained for R_g and A_2 is usually overlooked. In Appendix I it is demonstrated that:

$$\langle M \rangle = \frac{\sum_i w_i M_i}{\sum_i w_i} = M_w \quad (7)$$

$$\langle R_g^2 \rangle^{1/2} = \sqrt{\frac{\sum_i w_i R_{g,i}^2}{\sum_i w_i}} \quad (8)$$

$$\langle A_2 \rangle = \frac{\sum_i w_i^2 M_i A_{2,i}}{\left(\sum_i w_i \sum_i w_i M_i \right)} \quad (9)$$

where w_i , M_i , $R_{g,i}$, and $A_{2,i}$ represent the total mass, molecular weight, radius of gyration, and second virial coefficient, respectively, of kind i particles, in a polydisperse sample. The parameters between angle brackets represent the average value.

Dynamic light scattering

Light is scattered by the interaction of the electrons with the incident radiation (only the electric component will be considered here). The oscillating electric field causes a vibration on the electrons turning them into oscillating dipoles. These dipoles reemit radiation. As the electrons are moving sources (due to the Brownian motion) of radiation, the

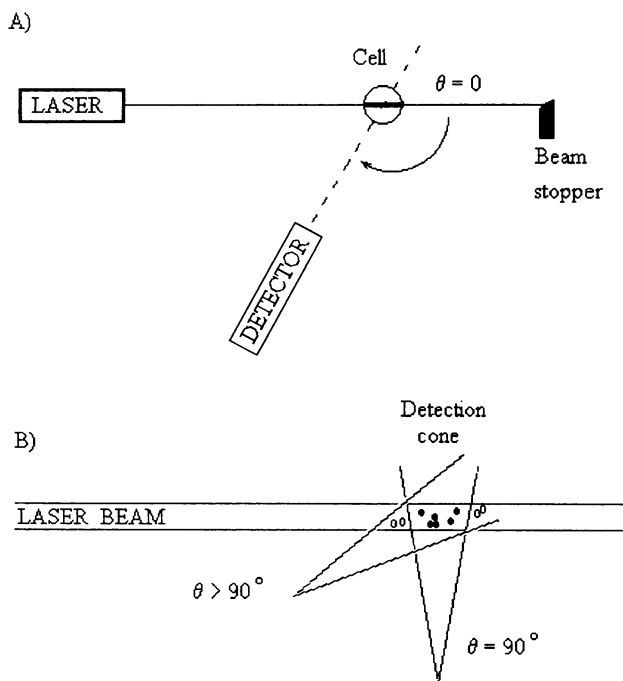


FIGURE 2 (A) Schematic representation of a light scattering apparatus. Light scattering intensities are recorded according to sample concentration and angle (by means of a goniometer). The detector might rotate, as depicted, or be fixed and connected to an optical fiber supported by the rotating arm. In any case, the detection device and the laser source must be aligned toward the geometrical center of the sample cell. The measurement angle (θ) origin is the way of the transmitted laser beam. (B) A geometrical effect has to be considered when light scattering intensities are recorded according to angle. As the detection angle deviates from 90° , the probed volume increases, including particles that are not detected at right angle (open circles). This geometric effect is corrected by a sinus function.

frequency of the radiation is shifted to higher or lower frequencies depending on its velocity and direction relative to the detector (Doppler effect). Molecules in solution move in all directions with equal probability and have a continuous speed distribution, thus a continuous broadening of the spectrum, relative to the incident frequency line (ν_0) (Fig. 3) is observed. The scattering of light is not exactly elastic, but quasi-elastic instead. This is why DLS is also named quasi-elastic light scattering (QELS). Because in SLS we are only concerned with the total intensity of the scattered light, ignoring the spectral distribution, SLS is also named intensity light scattering (ILS).

The power spectrum broadening is related to the Brownian motion of the particles in solution and hence to their diffusion coefficient, D , which in turn is related to the size and shape. However, the motion of large molecules is so slow that the broadening in the power spectrum is too small to be studied by interferometry. Therefore, instead of working in the frequency domain, we will work in the time domain (Fourier transform of the power spectrum) (Fig. 3). Our attention will be focused on how to obtain the time domain function to obtain the characteristic decay time of this function.

Light scattering intensity fluctuations detected in a small volume and in the microsecond time range (Fig. 4) are related to the Brownian motion of the particles due to density fluctuations, caused by incidental agglomeration of molecules and variation in the number of molecules in the scattering volume. The diffusion coefficient of the solute can be measured by means of an autocorrelation function ($g_2(t)$). (This methodology justifies the name of photon correlation spectroscopy sometimes used to name DLS. However, we do not think this is appropriate because several other spectroscopic methodologies can claim the same designation, e.g., fluorescence correlation spectroscopy is also a photon correlation spectroscopy.) Consider $I_{t'}$, the number of photons arriving at the detector at the time interval t' . The correlation function is built multiplying the number of photons from two successive time intervals and storing the result in the first instrumental channel. This calculation is repeated hundreds of thousands of times, averaged, and stored in channel 1. In the successive channels the average products of $I_{t'}I_{t'+t}$ are stored where t is the delay time:

$$g_2(t) = \langle I_{t'}I_{t'+t} \rangle. \quad (10)$$

At the limits:

$$\lim_{t \rightarrow 0} g_2(t) = \langle I_{t'}^2 \rangle \quad (11)$$

$$\lim_{t \rightarrow \infty} g_2(t) = \langle I_{t'} \rangle^2 \quad (12)$$

because correlation is maximal for close instants (most molecules have not collided, yet) and does not exist for very distant instants (Fig. 5). For small monodispersed particles and homogeneous spheres of any size the normalized scattered electric field autocorrelation function ($g_1(t)$) is:

$$g_1(t) = e^{-\Gamma t} \quad (13)$$

which is related to the intensity correlation function (Eq. 10) by the Siegert relation (Eq. 14); the quadratic dependence comes from the relation between the amplitude of the electric wave and the intensity, i.e., the rate of flow of radiation through unit area.

$$g_2(t) = \langle I_{t'}^2 \rangle \cdot b \cdot g_1^2(t) + \langle I_{t'} \rangle^2 \quad (14)$$

where b is an instrumental constant that reflects the deviations from ideal correlation (ideally $b = 1$) and

$$\Gamma = \frac{D}{q^2} \quad (15)$$

$$q = \frac{4n_0\pi}{\lambda} \sin\left(\frac{\theta}{2}\right) \quad (16)$$

where D is the diffusion coefficient, and Γ is the reciprocal of the characteristic decay time ($\tau = 1/\Gamma$).

For continuous polydispersed systems, the equation must be integrated over all the possible sizes (i.e., over all cor-

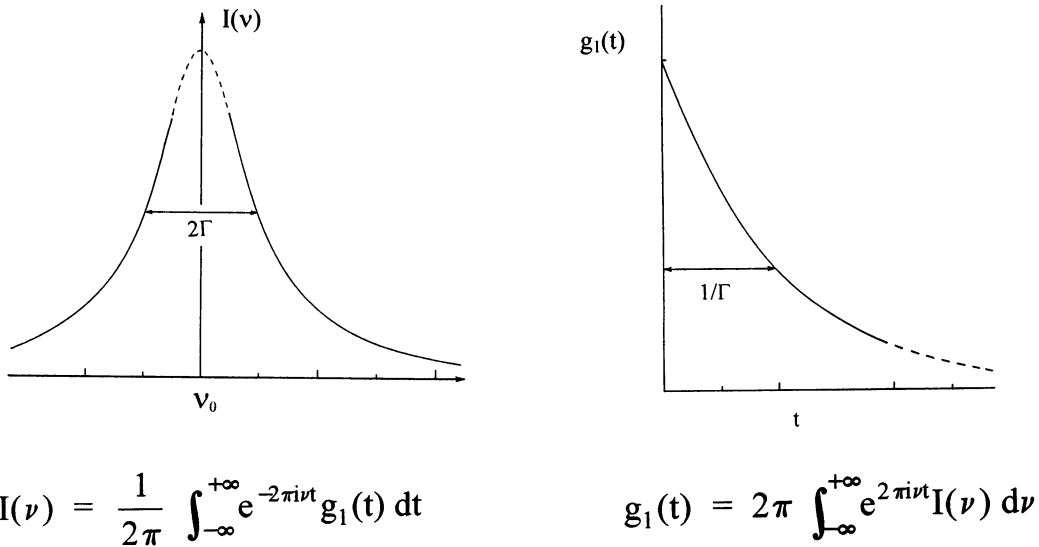


FIGURE 3 Power spectrum ($I(\nu)$) and corresponding Fourier transform ($g_1(t)$) with a characteristic decay time of $1/\Gamma$. The power spectrum frequency (ν) broadening is due to the Doppler effect on the Brownian motion of macromolecules.

responding Γ):

$$g_1(t) = \int_0^\infty G(\Gamma)e^{-\Gamma t}d\Gamma \quad (17)$$

$G(\Gamma)$ is the Γ distribution function and can be evaluated by inverse Laplace transform techniques, as $g_1(t)$ is the Laplace transform of $G(\Gamma)$. The most common routine used to perform the inverse Laplace Transform (ILT) is CONTIN (Provencher, 1982). CONTIN is a numerical method that starts from a preliminary unsmoothed solution in a frame of equally spaced $\log\Gamma$. After that, a regularization process is used according to statistical criteria. In the end, a “chosen solution” is selected. Such a chosen solution might be misleading; and very high signal-to-noise ratio data are required so that a unique solution can be admitted as the

chosen solution. Anyone using CONTIN or related methods should be aware of its limitations and artifacts that might be generated by this routine, including spurious peaks. These artifacts are related to integration processes (e.g., over-smooth and undersmooth), baseline definition (e.g., spurious peaks in skewed distributions), and edge effects (e.g., ripples in the distribution). All these aspects were very clearly discussed by Johnsen and Brown (1992). Undergraduates should give attention to all these aspects of CONTIN and related ILT methods because they are often too dazzled by the routine output to think about them.

If other variables related to Γ are used instead of Γ , a transformation of equation 16 has to be considered (see Appendix II).

The measured D value changes according to angle and concentration:

$$D = D_0(1 + k'_D R_g^2 q)(1 + k_D c) \quad (18)$$

where k'_D and k_D are constants and $D_0 = \lim_{q \rightarrow 0, c \rightarrow 0} D$. D_0 is related to the hydrodynamic radius (R_h) by the Stokes-Einstein relationship:

$$D_0 = \frac{kT}{6\pi\eta R_h} \quad (19)$$

where k is the Boltzman constant and η is the solvent viscosity.

For broad distributions and even for narrow distributions (i.e., quasi-monodispersed) the reference D value, commonly named effective diffusion coefficient (D_{eff}), used for calculation is:

$$D_{eff} = \frac{\langle \Gamma \rangle}{q^2} \quad (20)$$

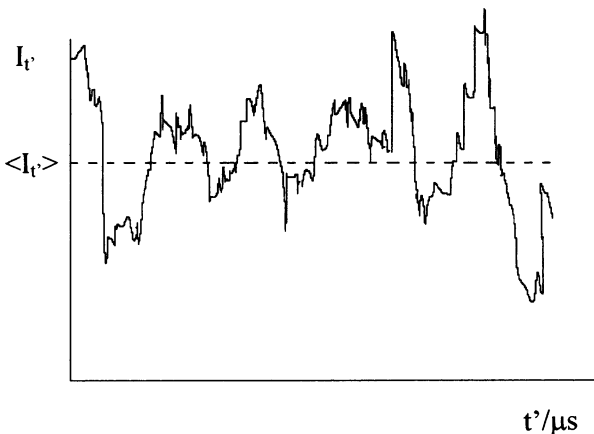


FIGURE 4 Schematic representation of light scattering intensity fluctuations from a small detection volume in the microsecond time range. $\langle I_t \rangle$ represents the average scattered intensity.

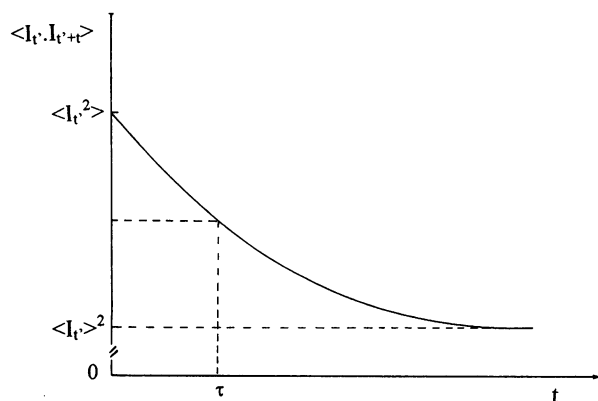


FIGURE 5 Schematic representation of an autocorrelation function. Several pairs of recordings of the scattered intensity (Fig. 4), occurring with a time t interval between them are multiplied, averaged, and represented against t . For small monodispersed particles and homogeneous spheres of any size, an exponential decay is obtained, with a characteristic decay time, $\tau = 1/\Gamma$.

where $\langle \Gamma \rangle$ is the average value of Γ in a peak. D_{eff} is the so-called z -average diffusion coefficient (D_z) (Eq. 20).

$$D_{\text{eff}} = \langle D \rangle = \sum_i w_i M_i D_i / \sum_i w_i M_i = D_z \quad (21)$$

(D_i is averaged by the scattered intensities since in Rayleigh scattering $I_i \propto n_i V_i^2 \propto w_i M_i$, where V_i is the volume of the particle i and n_i is the number of particles i in solution, per unit volume). Multimodal distributions (which are not the case in this work) are usually described by several $\langle \Gamma \rangle$, one for each peak of the distribution. If peaks are overlapped, it is impossible in practical terms to evaluate $\langle \Gamma \rangle$. In this situation it is a common procedure to consider the Γ value where the maximal value of peak occurs (Γ_{max}) instead. If the peaks are symmetrical, this approximation is always valid.

It should be stressed that some methods of data treatment lead to information on average diffusion coefficients over all the distribution function. This is the case, for instance, for the method proposed by Koppel (1972) (CUMULANTS) which results from the application of the statistical cumulants generating function. (The cumulants generating function of $G_x(t)$ is simply defined as $G_x(t) = \ln G_x(t)$.) If the constant baseline has already been subtracted from $g_2(t)$ the plot of $\ln(g_2(t))$ versus t should be strictly linear for a monodispersed system. Any deviation from this linear dependence is indicative of polydispersity. A series expansion yields:

$$\ln(g_1(t)) = K_0 - K_1 t + \frac{K_2 t^2}{2!} - \frac{K_3 t^3}{3!} + \dots \quad (22)$$

where K_0 is just an amplitude factor and

$$K_1 = \langle \Gamma \rangle \quad (23)$$

$$K_2 = \langle (\Gamma - \langle \Gamma \rangle)^2 \rangle \quad (24)$$

$$K_3 = \langle (\Gamma - \langle \Gamma \rangle)^3 \rangle \quad (25)$$

$$K_4 = \langle (\Gamma - \langle \Gamma \rangle)^4 \rangle - 3K_2^2 \quad (26)$$

$$K_n = \left[(-1)^n \frac{d^n}{dt^n} \ln(g_1(t)) \right]_{t=0} \quad (27)$$

K_n is the n th cumulant of $g_1(t)$. K_1 is the mean of Γ (z average) and K_2 is the variance of the Γ distribution. The ratio:

$$P = \frac{K_2}{K_1^2} \quad (28)$$

is the square of the relative SD (also known as the square of the coefficient of variation) and is called the polydispersity index. The bigger P is, the wider the distribution.

MATERIALS AND METHODS

Apparatus

The light scattering apparatus was from Brookhaven Instruments, Inc., model 2030AT, equipped with a He-Ne laser (632.8 nm) and a 128 channels autocorrelator, where the last six channels are used for baseline calculation. The UV-Vis absorbance spectrophotometer was from Jasco, model V800. A Sigma 2K15 centrifuge was also used.

Sample preparation

A TMV stock solution (American Type Culture Collection, MD, USA) 2 g/dm³, in 0.01 M phosphate buffer, pH 7.2, was diluted to a final concentration of 0.1 g/dm³ (checked by UV-Vis absorption using $\epsilon_{260\text{nm}} = 0.33 \text{ dm}^3 \text{ g}^{-1} \text{ cm}^{-1}$; Wilcoxon and Schurr, 1983). The solution pH (5.0, 7.2, and 10.0) was selected by the addition of HCl or NaOH. The ionic strength was controlled by the addition of NaCl (0 or 200 mM).

Undergraduates who had no previous practical contact with light scattering techniques might not have understand the concept of "dust" as a contaminant. Theoretically it is easy to explain the interference dust can cause by reminding students that in Rayleigh scattering the scattered intensity is proportional to concentration but is also proportional to the squared particle volume. Any dust particles in suspension, even at low concentration, might scatter significant light intensities due to their large volume (compared with the particles to be studied). In practical terms, students tend to underestimate the effort needed to obtain a dust-free sample, mainly in aqueous medium. Usually, the precautions taken to avoid air-borne dust contamination seem superfluous to students—the best strategy is to let them learn by themselves. After a few disappointments they will be more careful in sample preparation.

The samples were filtered through Millipore Millex GV 0.22 μm disposable filters, using a syringe, directly to cylindrical cells. Such cells were previously washed with chromo-sulfuric mixture and abundantly rinsed with dust-free water. Before measurement, the capped cells were centrifuged at $1300 \times g$ for 45 min to achieve the sedimentation of any remaining dust particle. Experiments were carried out at 22°C.

Sample concentrations should be chosen so that multiple scattering is as if absent and within the range that is valid for the use of Rayleigh-Gans-Debye light scattering formalisms. Criteria for these two conditions are not easy to settle. We have followed the suggestions of Glatter (1995) that states that when the turbidity of the sample leads to transmittance values that exceed 0.95, multiple scattering can be considered absent. If $2. \alpha \cdot (m - 1) \ll 1$, then the Rayleigh-Gans-Debye formalisms are valid ($\alpha = 2R/\lambda$; $\lambda = \lambda_0/n_{\text{solvent}}$; $m = n_{\text{solute}}/n_{\text{solvent}}$; n is the refractive index and R is the particle radius).

Data treatment

Although the software packages of most light scattering apparatus include an automatic SLS data treatment to calculate M_w , A_2 , and R_g , using the Zimm method, it is convenient that undergraduate students calculate these parameters step by step. This is the only way toward a complete understanding of the reasons beyond it. Otherwise, these parameters will be the result of a "black box." The concentration dependence in the range studied is negligible (Johnson and Brown, 1992) and so, even for the sake of simplicity and time management, we have admitted that $A_2 = 0$. Nevertheless, the methodology is very well demonstrated with R_g and M_w , only.

Equation 1 can be reformulated into:

$$\frac{1}{I_s} = \frac{A}{M_w} + \frac{A}{M_w} \frac{16\pi^2 n_0^2 R_g^2}{3\lambda^2} \sin^2\left(\frac{\theta}{2}\right) \quad (29)$$

where A is a constant. Plotting $1/I_s$ against $\sin^2(\theta/2)$ for low angles, a linear relationship is expected. Dividing the slope by the intercept, R_g can be calculated:

$$\frac{\text{Slope}}{\text{Intercept}} = \frac{16\pi^2 n_0^2 R_g^2}{3\lambda^2} \quad (30)$$

Once R_g is known, and assuming that dn/dc of TMV at 536 nm ($0.184 \text{ cm}^3 \text{ g}^{-1}$) (Huglin, 1989) is not significantly different from the one we would measure at 632.8 nm (an accurate measurement at 632.8 nm was prevented by the very low concentration of TMV used), then M_w can be obtained from Eq. 1 using the extrapolated data to $q = 0$ (zero angle conditions). The concentration dependence of D is also negligible in the conditions studied (Sano, 1987). This means $k_D = 0$ in Eq. 18 and thus $D_0 = \lim_{q \rightarrow 0} D$.

RESULTS

Dynamic light scattering

The size distribution function obtained at low angle are all fairly monodispersed and this characteristic is maintained regardless of pH and I . An example is depicted in Fig. 6.

The extrapolated D values to $q^2 = 0$ (D_0) are shown in Table 1, along with the corresponding R_h . The depicted values were obtained from $\langle \Gamma \rangle$ (CONTIN) but D_0 values obtained from the CUMULANTS method are similar.

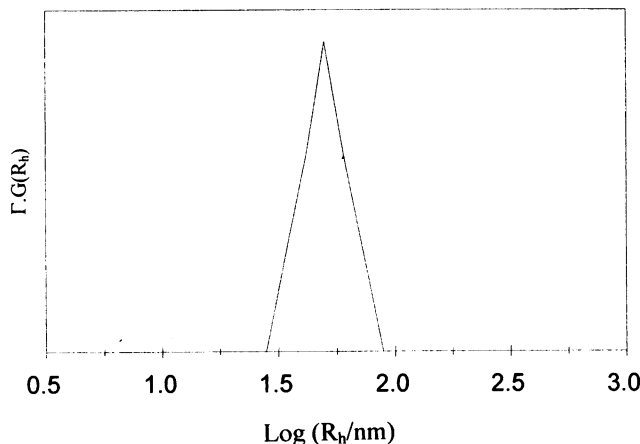


FIGURE 6 Example of CONTIN analysis output, obtained at $\theta = 50^\circ$, pH 7.2, and $I = 0$. This is a typical output, regardless of θ , pH, and I . All the results were monomodal and fairly monodispersed.

TABLE 1 Radius of gyration (R_g), hydrodynamic radius (R_h), diffusion coefficient at zero angle conditions (D_0), and weight average molecular weight of TMV at several pH and ionic strengths (I)

pH	I /mM	R_g /nm	R_h /nm	$D_0 \times 10^8/\text{cm}^2 \text{ s}^{-1}$	mol wt $10^{-6}/\text{g mol}^{-1}$
5.0	0	98.4	55	4.0	21.3
	50	108.8	51	4.3	18.3
	200	94.8	55	4.0	18.8
7.2	0	113.6	55	3.9	26.3
	50	103.1	51	4.2	23.5
	200	109.4	55	3.9	25.1
10.0	0	99.1	51	4.2	21.1
	50	92.6	51	4.3	22.9
	200	105.6	51	4.2	31.2

Static light scattering

The $1/I_s$ vs. $\sin^2(\theta/2)$ data were fitted by second degree polynomial functions. The residuals and χ^2 parameters were used to decide about the goodness of the fit. The fits were satisfactory in any case. Fig. 7 is an example (pH 7.2, $I = 0$ mM). R_g was calculated from the tangent to the second degree polynomial ($ax^2 + bx + c$) at zero angle ($bx + c$). R_g values are listed in Table 1 for the conditions studied. The corresponding values of M_w are also listed.

The experimental particle structure factor, P_q , is simply,

$$P_q = \frac{c}{ax^2 + bx + c} \quad (31)$$

which results directly from Eq. 4. Students tend to overlook the very simple definition of P_q (Eq. 4) and remember Eq. 5 only. It is noteworthy to stress that if we were dealing with concentration-dependent data, then Eq. 31 would be valid only for zero concentration (infinite dilution) extrapolated data.

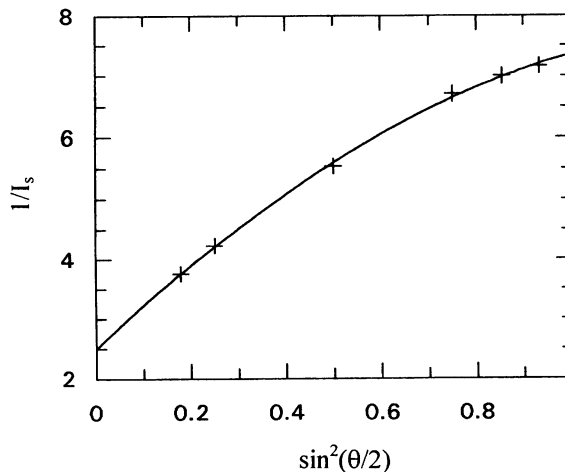


FIGURE 7 Static light scattering reciprocal intensity angular measurements. The experimental data (+; pH 7.2, $I = 0$) were fitted by a second degree polynomial (—). R_g and M_w are calculated from this data (see text). The results obtained at the other pH and I are similar.

DISCUSSION

None of the parameters studied seem to vary significantly with pH or I . The nucleic acid nucleic stabilizes the virus structure over a wider range of pH and I than those valid for the coat proteins alone (Fig. 1; for a review on coat protein assemblies alone see e.g., Buttler and Mayo, 1987). However, is this structure the cylinder that students are used to seeing on textbooks?

The first approach to solve this new problem can be to compare the experimental intraparticle static structure factor (named either P_q or P_θ) with the theoretical expectation for some of the more common geometry models.

Comparing experimental and theoretical P_q

The theoretical expectations for P_q according to each of several model geometries are quite complex. Eqs. 32–34 represent the theoretical expectations for a sphere (radius, R), infinitely thin rod (length, L), and Gaussian coil, respectively (e.g., Schmitz, 1990).

$$P_q(q.R) = \left\{ \frac{3}{(q.R)^3} [\sin(qR) - q.R.\cos(q.R)] \right\}^2 \quad (32)$$

$$P_q(q.L) = \frac{2}{q.L} \int_0^{q.L} \frac{\sin Z}{Z} dZ - \left[\frac{2}{q.L} \sin\left(\frac{q.L}{2}\right) \right]^2 \quad (33)$$

$$P_q(q.R_g) = \frac{2}{(q.R_g)^4} [\exp(-q^2.R_g^2) + (q.R_g)^2 - 1] \quad (34)$$

Some handbooks have long lists of P_q versus $q.R_g$ for several geometries (e.g., Casassa, 1989). Nevertheless, such handbooks are not common in most laboratories. That is why we present the following polynomial equations that result from fitting a fifth degree polynomial to the P_q versus $q.R_g$ data. Eqs. 35–37 correspond to the same three geometries referred above (sphere, thin rod, and random coil, respectively):

$$P_q(q.R_g) = 0.9977 + 0.0366qR_g - 0.4396(qR_g)^2 + 0.1072(qR_g)^3 + 0.0111(qR_g)^4 - 0.038(qR_g)^5 \quad (35)$$

$$P_q(q.R_g) = 0.9954 + 0.0837qR_g - 0.6068(qR_g)^2 + 0.3220(qR_g)^3 - 0.0656(qR_g)^4 + 0.0046(qR_g)^5 \quad (36)$$

$$P_q(q.R_g) = 0.9971 + 0.0579qR_g - 0.5388(qR_g)^2 + 0.2669(qR_g)^3 - 0.0520(qR_g)^4 + 0.036(qR_g)^5 \quad (37)$$

In these equations the parameters have no physical meaning because they result from a fit that is only intended to substitute the meaningful, but rather complex, original equations by simple, useful, and easy to handle equations. The deviations between fitted and “real” values are always

<1% for values of $qR_g < 2.3, 2.4,$ and $3.2,$ for a coil, sphere, and rod, respectively (data not shown).

In Fig. 8 the experimental P_q referring to pH 7.2 are plotted along with the theoretical P_q for different geometries. The results obtained for other values of pH are similar (data not shown). The data clearly suggest that the geometry of the virus is, in fact, a rod. However, the structure factor data are never absolutely convincing because even small degrees of polydispersity and/or excess scattering due to dust and/or some virus aggregation (Johnson and Brown, 1992) shift the experimental curves to higher values of P_q . Moreover, the rod theoretical expectation is calculated assuming $L \gg$ radius, which might not be totally true. An alternative approach is to compare R_g and R_h . R_g and R_h are not independent values because both concern the scattering particles dimension and geometry. The R_g/R_h ratio lies closer to the expectations for a cylinder (≈ 2.0 or more, depending on the axial ratio) than to any other model geometry. However, it is noteworthy that a random coil in a good solvent leads to $R_g/R_h = 1.78$ (Burchard, 1992). A rod-like structure is more reasonable to assume in the present work. R_g and R_h values are listed in Table 1.

The dimensions of the rod

At this time another question arises: *Because we are dealing with a rod, what are its dimensions?*

Cylinders have two basic parameters to describe their dimensions: length (L) and cross section radius (r). Each of the parameters on Table 1 (D_o , R_g , and M_w) depends on both. Therefore, it is not possible to find a unique solution for L and/or r from any of those parameters. An idea to overcome this limitation is to compare the experimental values of D_o and R_g with the theoretical expectations, assuming the dimensions detected by electron microscopy techniques. If the results are fairly coincident, it can be concluded that the nucleic acid stabilizes the virus with that specific length and radius, regardless of pH and I .

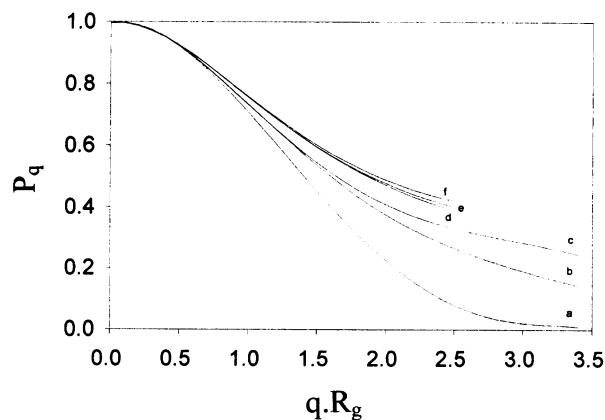


FIGURE 8 Theoretical (a, b, c) and experimental (from fitted data at pH 7.2; d, e, f) intraparticle static structure factor, P_q , according to $q.R_g$. a, Sphere; b, coil; c, rod; d, $I = 0$; e, $I = 200$ mM; and f, $I = 50$ mM. The results obtained at the other pH are similar.

The theoretical expectations are listed in Table 2 and the results are listed in Table 3. The experimental values of D_0 suggest we are dealing with cylinders having similar dimensions to the ones detected by electron microscopy techniques.

Not surprisingly R_g values are slightly higher than expected, probably due to the small but effective fraction of polydispersity and/or presence of trace amounts of dust and/or some end-to-end virus aggregation. The slightly overestimated R_g are a direct consequence from the "upward" shifts of experimental P_q relative to theoretical P_q . Nevertheless, our measurements are in agreement with a previously published result (Table 3).

The molecular weights evaluated by light scattering (Table 1) are compatible with the value of $(40 \pm 1) \times 10^6$ (Boedtker and Simmons, 1958; Weber et al., 1963), $39.4 \times 10^6 \pm 2\%$ referred by Caspar (1963), and 40.8×10^6 obtained from combining DLS and sedimentation data (Johnson and Brown, 1992). The translational diffusion coefficients calculated in this work are in close agreement with those available in the literature for TMV (see Table 3).

CONCLUSIONS

After data analysis and treatment, undergraduate students can easily conclude that: 1) The geometry of the TMV is invariant with pH (5.0–10.0) and I (0–200 mM); 2) the geometry of the TMV is a cylinder, and 3) the virus dimensions are $L \approx 300$ nm and $r \approx 20$ nm. The basic principles

TABLE 2 Equations used to calculate the theoretical expectations for radius of gyration (R_g), hydrodynamic radius (R_h), and diffusion coefficient (D) of TMV*

Equation	Reference/remark
$D = \frac{kT}{6\pi\eta b'} G(\rho)$	Broersma, 1960
$D = \frac{kT}{6\pi\eta b'} \ln(\rho)$	Doi and Edwards, 1986
$D = \frac{kT}{6\pi\eta b'} \left(\ln \rho + 0.3 + \frac{0.4738}{\rho} + \frac{0.4167}{\rho^2} - \frac{0.3394}{\rho^3} \right)$	Tirado and De la Torre, 1979
$D = \frac{kT}{6\pi\eta b} (1 - \frac{a^2}{b^2})^{-1/2} \ln \left(\frac{1 + (1 - (a^2/b^2))^{1/2}}{a/b} \right)$	Perrin, 1936 (cf., Berne and Pecora, 1990)
$R_g = \left(\frac{L^2}{12} + \frac{r^2}{2} \right)^{1/2}$	Thin rod
$R_g = \left(\frac{b^2}{3} + \frac{a^2}{4} \right)^{1/2}$	Ellipsoid

*Equations involving a and b are valid for an prolate ellipsoid of minor semiaxis a and major semiaxis b . The other equations are valid for rods. L , rod length; r , rod cross-sectional radius, $\sigma = \ln(L/r)$, $b' = L/2$, $\rho = L/(2r)$, $G(\rho) = \ln(2\rho) - 1/2(1.46 - 7.4((\ln^{-1}(2\rho) - 0.34)^2 - 4.2(\ln^{-1}(2\rho) - 0.39)^2))$.

and potentialities concerning size and shape determination using light scattering techniques (static and dynamic) are illustrated with this work.

CONCLUDING REMARKS

The interest and attention of undergraduates are often conquered by some curiosities about the techniques to be studied and the phenomena behind them. A white cloud, a dark blue sky, or an intense red sunset are good opportunities to talk about science with undergraduates, mentioning Mie and Rayleigh scattering. Some pedagogic bibliography can be provided on these natural spectroscopic phenomena (e.g., Young, 1982; Walker 1989). Students will find it interesting to know that one of the key questions that was in the minds of J. Strutt (later Lord Rayleigh) and J. Tyndall during the last century was *Why is the sky blue?* (the titles of their communications were very explicit toward this goal; e.g., Tyndall, 1869; Strutt, 1871).

The authors acknowledge the collaboration and helpful comments from the students graduating in Biochemistry (91/92-95/96) who attended the Molecular Biophysics Course (1995), Faculty of Sciences, University of Lisbon. N.C.S. acknowledges a grant from Junta Nacional de Investigaço Científica e Tecnolgica (J.N.I.C.T.) (Portugal). The authors acknowledge the collaboration of Dr. Manuel Prieto and the financial assistance of J.N.I.C.T. (Portugal)—Project PECS/C/SAU/144/95.

APPENDIX I

When the results obtained by static light scattering for a polydisperse sample are analyzed using the traditional Zimm equation, average parameters, total concentration, and total scattered intensities are considered:

$$K \sum_i c_i / K'' \sum_i I_i = \frac{1}{\langle M \rangle} \left(1 + K' \langle R_g^2 \rangle \sin^2 \left(\frac{\theta}{2} \right) \right) + 2 \langle A_2 \rangle \sum_i c_i \quad (\text{AI.1})$$

$$= \frac{1}{\langle M \rangle} + 2 \langle A_2 \rangle \sum_i c_i + \frac{K' \langle R_g^2 \rangle \sin^2(\theta/2)}{\langle M \rangle}$$

where K , K' , and K'' are constants (K has the usual meaning, $K' = q^2/(3 \cdot \sin^2(\theta/2))$ and $K'' = R_g/\sum I_i$), c_i and I_i are the concentration and light scattered by particles i in solution. Moreover,

$$K \sum_i c_i / K'' \sum_i I_i = \sum_i (K c_i / (K'' I_i) I_i / \sum_j I_j) \quad (\text{AI.2})$$

Having in mind that each of the components will follow the same general variation (eq. AI.3)

$$\frac{K c_i}{K'' I_i} = \frac{1}{M_i} \left(1 + K' R_{g,i}^2 \sin^2 \left(\frac{\theta}{2} \right) \right) + 2 A_{2,i} c_i \quad (\text{AI.3})$$

and that $I_i \propto n_i V_i^2 \propto n_i M_i^2$ (n_i , V_i , and M_i are the number, the volume and the molecular weight of particles i , respectively), we get to Eq. AI.4.

$$K \sum_i c_i / K'' \sum_i I_i = \sum_i \left(\left(\frac{1}{M_i} \left(1 + K' R_{g,i}^2 \sin^2 \left(\frac{\theta}{2} \right) \right) + 2 A_{2,i} c_i \right) n_i M_i^2 / \sum_j n_j M_j^2 \right) \quad (\text{AI.4})$$

TABLE 3 Theoretical values expected for the radius of gyration (R_g), hydrodynamic radius (R_h), and diffusion coefficient (D) when applying the dimension results (length, L and radius, r) of electronic microscopy (Caspar, 1963; Shepherd, 1975; Namba and Stubbs, 1986) in the equations of Table 2*

Parameter	Reference	References						This work	Other light scattering works
		Caspar		Shepherd		Namba and Stubbs			
		L/nm	r/nm	L/nm	r/nm	L/nm	r/nm		
		300	20	300	18	298	18		
$D \times 10^8/cm^2s^{-1}$	Broersma	2.8		3.0		3.0		3.9–4.3	2.8–4.5 [#]
	Doi and Edwards	2.9		3.0		3.0			
	Tirado and De la Torre	3.4		3.6		3.6			
	Perrin	4.9		5.0		5.0			
R_g/nm	Thin rod	87.7		87.5		87.0		93–113	92.4 [§]
	Ellipsoid	86.7		86.7		86.1			

*For the sake of comparison, the experimental measurements obtained in this and other light scattering works are also presented.

[#]Values of $2.8 < D < 10^8 \text{ cm}^2\text{s}^{-1} < 3.5$ were collected from several references by Shepherd (1975). Wilcoxon and Schurr (1983) calculated $D_0 = 4.35 \times 10^{-8} \text{ cm}^2\text{s}^{-1}$. Sano (1987) obtained $D_0 = 3.75 \times 10^{-8} \text{ cm}^2\text{s}^{-1}$ in phosphate buffer and $D_0 = 4.50 \times 10^{-8} \text{ cm}^2\text{s}^{-1}$ in Tris buffer. Johnson and Brown (1992) calculated $D_0 = 4.27 \times 10^{-8} \text{ cm}^2\text{s}^{-1}$.

[§]Brunner and Dransfeld, 1983.

For the sake of simplicity we will use,

$$B_i = n_i M_i^2 / \sum_j n_j M_j^2 \quad (\text{AI.5})$$

Combining Eqs. AI.4 and AI.5:

$$K \sum_i c_i / K' \sum_i I_i = \sum_i \frac{B_i}{M_i} + 2 \sum_i (A_{2,i} B_i c_i) + K' \sin^2 \left(\frac{\theta}{2} \right) \sum_i \frac{R_{g,i}^2 B_i}{M_i} \quad (\text{AI.6})$$

To keep the coherence between the equivalent Eqs. AI.1 and AI.6, the terms dependent on concentration, scattering angle and neither of these parameters, must be identical in the two equations:

$$\left\{ \begin{array}{l} \frac{1}{\langle M \rangle} = \sum_i \frac{B_i}{M_i} \\ \langle A_2 \rangle \sum_i c_i = \sum_i (A_{2,i} B_i c_i) \\ \langle R_g^2 \rangle / \langle M \rangle = \sum_i \frac{R_{g,i}^2 B_i}{M_i} \end{array} \right. \quad (\text{AI.7})$$

Using the equations of the system, keeping in mind the definition of B_i (Eq. AI.5) and also that of $n_i M_i \propto w_i$ and $c_i \propto w_i$, we finally get the kinds of averages of M_w , R_g , and A_2 :

$$\langle M \rangle = \sum_i w_i M_i / \sum_i w_i = M_w \quad (\text{AI.8})$$

$$\langle R_g^2 \rangle^{-1/2} = \sqrt{\sum_i w_i R_{g,i}^2 / \sum_i w_i} \quad (\text{AI.9})$$

$$\langle A_2 \rangle = \sum_i w_i^2 M_i A_{2,i} / \left(\sum_i w_i \sum_i w_i M_i \right) \quad (\text{AI.10})$$

APPENDIX II

When we are dealing with a sample having a continuous distribution of sizes, the normalized field correlation function is:

$$g_1(q, t) = \int_0^\infty G(\Gamma, q) \exp(-\Gamma t) d\Gamma \quad (\text{AII.1})$$

A plot of $G(\Gamma, q)$ versus Γ represents the normalized intensity distribution of the sample. The area of each peak is the fraction of light scattered by the population of particles responsible for that peak. However, for the sake of clarity, most of the time plots are not represented with the abscissa Γ but with a more suitable variable instead. In this work, we have used the abscissa $\log(R_h)$. Therefore, the following transformation has to be carried out in Eq. AII.1:

$$g_1(q, t) = \int_0^\infty G(\Gamma, q) \exp(-\Gamma t) \frac{d\Gamma}{df(\Gamma)} df(\Gamma) \quad (\text{AII.2})$$

$$f(\Gamma) = \log(R_h) = \log \left(\frac{kTq^2}{6\pi\eta\Gamma} \right) \quad (\text{AII.3})$$

Combining AII.2 and AII.3:

$$g_1(q, t) = \ln 10 \int_0^\infty \Gamma G(\Gamma, q) \exp(-\Gamma t) d\log(R_h) \quad (\text{AII.4})$$

$\ln 10$ is a normalization constant. The fraction of the area corresponding to a certain peak is equal to the fraction of the total light intensity scattered by the particles population described by such peak only if $\Gamma G(\Gamma)$ is plotted when $\log(R_h)$ is the abscissa, instead of Γ .

It can also be easily demonstrated that when the abscissa are $\log D$, τ , or $\log \tau$, the ordinates should be $\Gamma G(\Gamma)$, $A(\tau) = \Gamma^2 G(\Gamma)$ and $\tau A(\tau) = \Gamma G(\Gamma)$, respectively.

The same kind of problem should be kept in mind in other kinds of spectroscopies, such as fluorescence spectroscopy. When converting a

spectrum in wavelength (λ) to frequencies or wavenumbers ($\bar{\nu}$), for instance, it may take more than only converting the abscissa. If I_f represents the fluorescence intensity, then

$$\int_{\lambda_1}^{\lambda_2} I_f d\lambda = \int_{\nu_1}^{\nu_2} I_f (d\lambda d\bar{\nu}) d\bar{\nu} = \int_{\nu_1}^{\nu_2} I_f \bar{\nu}^{-2} d\bar{\nu}.$$

Such problem does not exist in an absorption spectrum, where relative measurements are carried out, canceling all the needed correction terms in the data analysis.

REFERENCES

- Berne, B. J. and R. Pecora. 1990. Dynamic Light Scattering with Applications to Chemistry, Biology and Physics. Robert E. Krieger Pub. Co., Malabar. 143–144.
- Bloomfield, V. A. 1981. Quasi-elastic light-scattering in biochemistry and biology. *Annu. Rev. Biophys. Bioeng.* 10:421–450.
- Boedtker, H. and N. S. Simmons. 1958. The preparation and characterization of essentially uniform tobacco mosaic virus particles. *J. Am. Chem. Soc.* 80:2550–2556.
- Broersma, S. 1960. Rotational diffusion constant of a cylindrical particle. *J. Chem. Phys.* 32:1626–1635.
- Brunner, H. and K. Dransfeld. 1983. Light scattering by macromolecules. In *Biophysics*, W. Hoppe, W. Lohmann, H. Markl, and H. Ziegler, editors. Springer-Verlag, New York. 93–100.
- Burchard, W. 1992. Static and dynamic light scattering approaches to structure determination of biopolymers. In *Laser Light Scattering in Biochemistry*. S. E. Harding, D. B. Sattelle, and V. A. Bloomfield, editors. Royal Society of Chemistry, Cambridge. 3–22.
- Butler, P. J. G. and M. A. Mayo. 1987. Molecular architecture and assembly of tobacco mosaic virus particles. In *The molecular biology of the positive strand RNA viruses*. Academic Press Inc., London. 237–257.
- Casassa, E. F. 1989. Particle scattering factors in Rayleigh scattering. In *Polymer Handbook*. J. Brandrup and E. H. Immergut, editors. John Wiley & Sons, New York. VII/485–491.
- Caspar, D. L. D. 1963. Assembly and stability of the tobacco mosaic virus particle. *Adv. Protein Chem.* 18:37–121.
- Chu, B. 1991. *Laser Light Scattering. Basic Principles and Practice*. Academic Press, New York. 13–20.
- Doi, M. and S. F. Edwards. 1986. *The Theory of Polymer Dynamics*. Oxford University Press, Oxford. 289–323.
- Glatter, O. 1995. Modern methods of data analysis in small-angle scattering and light scattering. In *Modern Aspects of Small-Angle Scattering*. H. Brumberger, editor. Kluwer Academic Publishers, Dordrecht. 107–180.
- Harding, S. E., D. B. Sattelle, and V. A. Bloomfield. 1992. *Laser Light Scattering in Biochemistry*. Royal Society of Chemistry, Cambridge.
- Huglin, M. B. 1989. Specific refractive index increments of polymers in dilute solutions. In *Polymer Handbook*. J. Brandrup and E. H. Immergut, editors. John Wiley & Sons, New York. VII/409–484.
- Johnsen, R. and W. Brown. 1992. An overview of current methods of analyzing QLS data. In *Laser Light Scattering in Biochemistry*. S. E. Harding, D. B. Sattelle, and V. A. Bloomfield, editors. Royal Society of Chemistry, Cambridge. 77–91.
- Johnson, P. and W. Brown. 1992. An investigation of rigid rod-like particles in dilute solution. In *Laser Light Scattering in Biochemistry*. S. E. Harding, D. B. Sattelle, and V. A. Bloomfield, editors. Royal Society of Chemistry, Cambridge. 161–183.
- Koppel, D. E. 1972. Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants. *J. Chem. Phys.* 57:4814–4820.
- Marshall, A. G. 1978. *Biophysical Chemistry: Principles, Techniques and Applications*. John Wiley & Sons, New York. 463–503.
- Matthews, G. P. 1984. Light Scattering by Polymers. Two Experiments for Advanced Undergraduates. *J. Chem. Educ.* 61:552–554.
- Mougan, M. A., A. Coello, F. Meijide, and J. V. Tato. 1995. Spectrofluorimeters as light scattering apparatus. Application to polymers molecular weight determination. *J. Chem. Educ.* 72:284–286.
- Munk, P. 1989. *Introduction to Macromolecular Science*. John Wiley & Sons, New York. 375–400.
- Namba, K. and G. Stubbs. 1986. Structure of tobacco mosaic virus at 3.6 Å resolution: implications for assembly. *Science*. 231:1401–1406.
- Oster, G. 1972. Light Scattering. In *Techniques of Chemistry*, Vol. I, Part IIIA. A. Weissberger and B. W. Rossiter, editors. Wiley-Interscience, New York. 75–117.
- Perrin, F. 1936. Brownian movement of an ellipsoid. II. Free rotation and depolarization of fluorescence. Translation and diffusion of ellipsoidal molecules. *J. Phys. Radium*. 7:1–11.
- Pike, E. R., W. R. M. Pomeroy, and J. M. Vaughan. 1975. Measurement of Rayleigh ratio for several pure liquids using a laser and monitored photon counting. *J. Chem. Phys.* 62:3188–3192.
- Provencher, S. W. 1982. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comput. Phys. Commun.* 27:213–227.
- Sano, Y. 1987. Translational diffusion coefficient of tobacco mosaic virus particles. *J. Gen. Virol.* 68:2439–2442.
- Schmitz, K. S. 1990. *An Introduction to Dynamic Light Scattering by Macromolecules*. Academic Press, San Diego. 22.
- Shepherd, I. W. 1975. Inelastic laser light scattering from synthetic and biological polymers. *Rep. Prog. Phys.* 38:565–620.
- Strutt, J. W. 1871. On the light from the sky, its polarization and color. *Phil. Mag.* 41:107–120, 274–279.
- Thompson, A. C., K. G. Kozimer, and D. Stockwell. 1970. A light scattering experiment for physical chemistry. *J. Chem. Educ.* 47:828–831.
- Tirado, M. M. and J. G. De la Torre. 1979. Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders. *J. Chem. Phys.* 71:2581–2587.
- Tyndall, J. 1869. On the blue color of the sky, the polarization of skylight and the polarization of light by cloudy matter generally. *Phil. Mag.* 37:384–394.
- Walker, J. 1989. The colors seen in the sky offer lessons in optical spectroscopy. *Sci. Am.* 260:84–87.
- Weber, F. N., Jr., R. M. Elton, H. G. Kim, R. D. Rose, R. L. Steere, and D. W. Kupke. 1963. Equilibrium sedimentation of uniform rods of tobacco mosaic virus. *Science*. 140:1090–1092.
- Wilcoxon, J. and J. M. Schurr. 1983. Dynamic light scattering from thin rigid rods: anisotropy of translational diffusion of tobacco mosaic virus. *Biopolymers*. 22:849–867.
- Young, A. T. 1982. Rayleigh scattering. *Phys. Today*. 35:42–48.
- Zimm, B. H. 1948. Apparatus and methods for measurement and interpretation of the angular variation of light scattering; preliminary results on polystyrene solutions. *J. Chem. Phys.* 16:1099–1116.