Contribution of Slow Charge Fluctuations to Light Scattering from a Monodisperse Solution of Macromolecules[1]

George D. J. Phillies

Department of Chemistry, University of California, Los Angeles, Los Angeles, California 90024. *Received November 14, 1975*

ABSTRACT: For a monodisperse solution of charged macromolecules, a model calculation shows that the second cumulant (or "variance") of the spectrum of the quasielastically scattered light includes a term which is proportional to the root-mean-square range of charges on the macromolecules. This term is also proportional to the strength of the electrostatic interactions between the macromolecules (and is thus inversely proportional to the ionic strength of the solution). Increases in the variance of the spectrum attendant to a reduction in the ionic strength of the solution therefore need not be interpreted as arising from an increase in the polydispersity of the system at low ionic strength. Comparison of the model calculation with the experimental data of Doherty and Benedek (*J. Chem. Phys.*, **61**, 5426 (1974)) indicates that bovine serum albumin of average charge $\bar{Z} = -10$ has root-mean-square charge fluctuations of 3–4.5, which is in agreement with previous measurements of this quantity.

I. Introduction

Quasielastic light-scattering spectroscopy has been widely used to study the diffusion of macromolecules in aqueous solution. In early experiments on dilute solutions, in which interactions between solute molecules were taken to be negligible, it was usually sufficient to assume that a single solute species was present and that the measured mutual diffusion coefficient D of the solute was related to its hydrodynamic drag coefficient f by the Stokes–Einstein equation

$$D = K_B T/f \tag{I.1}$$

where K_B is Boltzman's constant and T is the absolute temperature.[2a] With these assumptions, the intensity–intensity correlation function of the scattered light is characterized by a single exponential with decay time Dk^2, k being the magnitude of the scattering vector **k** in the experiment.

Advances in experiment technique and studies of more complicated experimental systems have shown that these assumptions are not always adequate to interpret experimental results. When examined with sufficient care, the experimentally measured intensity–intensity correlation function is generally found to contain a distribution of exponentials. This distribution is conveniently parameterized by its cumulants.[2b] The first cumulant is proportional to the light-scattering intensity weighted average diffusion coefficient, while the second cumulant, or "variance", is proportional to the root-mean-square range of exponentials present in solution. A nonzero value for the second cumulant is usually taken to arise from the polydispersity of the solution. That is, rather than containing a single diffusing species, it is frequently found that solutions of biological macromolecules contain appreciable numbers of solute dimers, trimers, and higher oligomers. Each solute species contributes a separate exponential to the spectrum. In practise, only averages over all exponentials present can be measured. (Hocker et al.[3] have discussed how the first two cumulants may, in specific circumstances, be combined to estimate the monomer diffusion coefficient.)

In many solutions, interactions between solute molecules make an appreciable contribution to diffusion.[4] In a system containing a single solute species, whose interactions have a range short by comparison with a light wavelength, the mutual diffusion coefficient is found to be[5]

$$D = \frac{(\partial\pi/\partial c)_{P,T}}{f}(1 - \phi)$$

$\partial\pi/\partial c$ being the osmotic compressibility of the solute at constant temperature and pressure and ϕ being the volume fraction of the solute molecules. (Long-range solute–solute interactions have been predicted[6] and, independently, found[7] to cause the mutual diffusion coefficient measured by light scattering to depend on the scattering vector **k**.)

Most discussions of light scattering from concentrated macromolecule solutions have confined themselves to analyzing the contribution of solute–solute interactions to the average diffusion coefficient.[6,8] In this paper, certain aspects of the contribution of solute–solute interactions to the range of diffusion coefficients seen in a multicomponent system will be developed. For simplicity, the discussion will be confined to effects arising from the presence in an elsewise monodisperse protein solution of molecules bearing different charges. For this simple model system, the second cumulant of the diffusion coefficient distribution is predicted to be proportional to the mean-square charge fluctuation δ of the macromolecules. This simple result will be compared with the recently published experimental data of Doherty and Benedek.[4]

II. Calculation. Comparison with Experimental Data

Quasielastic light-scattering spectroscopy measures the mutual diffusion coefficient by studying the decay of spontaneous fluctuations in the solute concentration. Solute–solute interactions affect the spectrum in two distinct ways:

(1) The intensity of the light scattered by each species in solution is proportional to the average size of the concentration fluctuations of that species and to the correlations between concentration fluctuations of the different solute species. The contribution of such solute–solute interactions was previously considered by Kirkwood and Goldberg,[9] who calculated the intensity of the light scattered from a multicomponent nonideal solution.

(2) In the absence of solute–solute interactions, a concentration fluctuation can only relax through the Brownian motion of the individual solute molecules. If a concentration gradient is present, interactions between solute molecules may augment or reduce the solute current due to Brownian motion, thereby changing the rate at which a concentration fluctuation decays. Such effects, with especial attention to light-scattering spectroscopic applications, have been described, for example, by this author.[8]

Here we consider the contribution of charge heterogeneity to the spectrum of light scattered from a monodisperse protein solution. A real system would contain protein molecules with a substantial number of different charges. A realistic calculation would treat each charge state of the protein as a distinct chemical component of the solution, which would lead to extremely complicated results. In this paper we shall consider

Table I
Chemical Reactions Involved in Proton Binding in Bovine Serum Albumin[14]

Reaction	k_r,[a] M^{-1} s^{-1}	k_D,[a] s^{-1}
(i) $OH^- + C_3N_2H_5^+ \rightleftharpoons C_3N_2H_4 + H_2O$ (imidazole)	2.5×10^{10}	2.5×10^3
(ii) $H^+ + HCO_2^- \rightleftharpoons HCO_2H$	4×10^{10}	5×10^5
(iii) $OH^- + {}^+H_3NCH_2CO_2^- \rightleftharpoons H_2N\,CH_2CO_2^- + H_2O$	1.4×10^{10}	8.4×10^5

[a] k_r and k_D are the rate constants for the front and back reactions, respectively.

only a simple model for the problem. The model system is assumed to contain two solute species of charge $q_1 = Z + \delta$ and $q_2 = Z - \delta$, respectively. The species are taken to be present at equal number density n, so that the average charge and root-mean-square charge fluctuation on a solute molecule will be Z and δ, respectively. The solute concentration is assumed to be small enough that reference frame corrections, which arise if the solute occupies an appreciable fraction of the volume of the solution, are not important.

The importance of the charge fluctuations will depend in part on their temporal behavior. If a protein molecule keeps the same charge for many decay times of a concentration fluctuation, molecules of different charges will diffuse essentially as separate species. On the other hand, if the molecules in a concentration fluctuation each change their charges many times before that fluctuation can decay, the molecules will move as though they permanently carried their average charge. This latter effect plays a prominent part in light-scattering electrophoretic measurements.[10]

Table I presents rate constants[11] for the important chemical reactions which contribute to proton binding by bovine serum albumin, the protein studied by Doherty and Benedek. These rate constants should be compared with the decay time τ of a concentration fluctuation; for serum albumin ($D \approx 6 \times 10^{-7}$ cm^2/s) illuminated by 5000-Å light at a scattering angle of $\theta = 90°$, τ is ~ 30 μs. At pH 7.0 (for which[12] $Z \approx -10$), reaction (i) has a time constant $\sim 10^{-3}$ s, which is much longer than τ. When studied by light scattering, protein molecules whose charge differs by virtue of reaction (i) will appear as distinct species. Reactions (ii) and (iii) have $k_D \approx 5$–8×10^5 s^{-1}, which is on the order of τ. However, at pH 7, the equilibrium ratios of the dissociated and bound states of (ii) and (iii) are very different from unity. Consequently, near pH 7.0 reactions (ii) and (iii) will only affect the charge of a small number of protein molecules, so the contribution of these reactions to the light-scattering spectrum will be very small.

At a different pH or scattering angle, the conclusions of the above paragraph would need to be modified. At a higher or lower pH, reactions (ii) or (iii) could contribute substantially to charge fluctuations; by reducing the scattering angle, one can attain $\tau \sim 1$ ms, which is comparable to the charge-changing time of reaction (i). In either of these cases, the spectrum of the scattered light will be affected by the dynamics of the chemical reaction.

Protein molecules gain and lose charge by binding and releasing protons from and into the solution. Such a small change will, in general, have little effect on the polarizability of the molecule. The polarizability of the two protein species in our calculation will, therefore, be assumed to be equal. This means that the amount of light scattered by a volume of solution will depend on the total number of protein molecules in that volume, but not on the relative number of molecules in either of the charge states. If the polarizability of the molecules did depend on the number of bound protons, the chemical reaction would contribute directly to the spectrum of the scattered light. Despite many attempts, no repeatable observation of a reactive contribution to light scattering has apparently been made in any system.

The interactions between the prótein molecules are described by the derivatives of their chemical potentials, which are here written

$$\left(\frac{\partial \mu_i}{\partial n_j}\right)_{P,T} = \frac{KT}{n_j}\,\delta_{ij} + Iq_iq_j \qquad \text{(II.1)}$$

where δ_{ij} is the Kronecker δ and I is a measure of the strength of the electrostatic solute–solute interactions. By using this form, we take the interaction between the two charge states of the protein to be symmetric. Since the calculation is limited to monodisperse systems, all protein molecules will have the same drag coefficient f; differences in diffusion coefficient arise entirely from the derivatives (eq II.1) of the chemical potential. The diffusion coefficients are related to the chemical potential derivatives by

$$D_{ij} = n_i \frac{(\partial \mu_i/\partial n_j)_{P,T}}{f} \qquad \text{(II.2)}$$

On combining the above and the results of ref 8 and 9, the time correlation function of the scattered field is predicted to be

$$\langle E(0)E(t)\rangle = A_1 + A_2(D_w{}^+ e^{-D^+k^2t} + D_w{}^- e^{-D^-k^2t}) \qquad \text{(II.3)}$$

where A_1 and A_2 are constants and

$$\begin{aligned} D_w{}^+ &= 2Z^2(K_BT + 2nI(Z^2 - \delta^2)) \\ D_w{}^- &= 2\delta^2(K_BT + 4nIZ^2) \\ D^+ &= [K_BT + 2nI(Z^2 + \delta^2)]/f \\ D^- &= K_BT/f \end{aligned} \qquad \text{(II.4)}$$

The spectrum of the scattered light is seen to contain two exponentials of time constant D^+ and D^-. If the two protein species did not interact with each other, D^+ and D^- would be the diffusion coefficients of the two protein species in the solution. However, since the two species do interact, D^+ and D^- reflect the coupled motion of the protein molecules, which can be described in terms of two overdamped normal modes. The first of the modes, with time constant D^-, relaxes only by the random walk of the individual solute molecules. This mode corresponds to coupled concentration fluctuations in species 1 and 2 which are so arranged that the total density of protein-bound charge remains everywhere the same. The second mode, with time constant D^+, is necessarily orthogonal to the first. It corresponds to the relaxation of a gradient in the density of protein-bound charge.

The spectrum of the light is thus seen to be a sum of decaying exponentials. If the spectrum could be measured with sufficient accuracy, each exponential could be separately identified, thereby permitting a complete characterization of the ensemble-average temporal behavior of the scattered light. This is generally impossible.

An alternate way to characterize the spectrum is the method of cumulants.[2] One may in general write

$$\langle E(0)E(t)\rangle = \int_0^\infty A(\Gamma)e^{-\Gamma t}\,d\Gamma \qquad \text{(II.5)}$$

where $A(\Gamma)$ is the distribution of decay times. The method of cumulants expands this quantity as a power series

$$\int_0^\infty A(\Gamma)e^{-\Gamma t}\,d\Gamma = \exp\left(-\sum_{i=0}^{\infty} \mu_i t^i\right) \qquad \text{(II.6)}$$

the μ_i being the cumulants of the expansion. Denoting $\Gamma^n = \int_\infty^0 A(\Gamma)\Gamma^n\,d\Gamma$, $\mu_1 = \bar{\Gamma}$ is simply the decay time, while $\mu_2 = \overline{\Gamma^2} - \bar{\Gamma}^2$ is a measure of the width of $A(\Gamma)$. It is convenient to

describe the spread of decay times in the terms of the dimensionless parameter $\sigma = \sqrt{\mu_2/\mu_1}$, which is often referred to as the "variance". If $A(\Gamma)$ were a Gaussian, σ would be its half-width at half-height. With modern light-scattering equipment, the average diffusion coefficient and the variance can readily be measured. In special cases, reproducible values for the third and higher cumulants have been obtained.[13]

For the model system being discussed, one may readily obtain analytical expressions for the average diffusion coefficient and the variance:

$$\bar{D} = \frac{K_B T + 2nI(Z^2 - \delta^2)}{f} + \frac{2nI\delta^2 K_B T}{(K_B T + 2nIZ^2)f} \qquad \text{(II.7)}$$

$$\sigma = \frac{2nI\delta Z}{K_B T}\{[(1+a)^4 + (a+b-2)(1+a)^3 + (1+ab-b^2-2a)(1+a)^2 + (1+a)(a-b-ab^2+2b^2-b^3) + ab(1-b)^2] \times [a^2(1+a)^4 + 2ab(1+a)^3 + 2a^2b(1-b)(1+a)^2 + b^2(1+a)^2 + 2ab^2(1-b)(1+a) + a^2b^2(1-b)^2]^{-1}\}^{1/2}$$

where $a = 2nIZ^2/K_B T$ and $b = 2nI\delta^2/K_B T$.

The effect of the interactions between the macromolecules is to increase the average diffusion coefficient from the free diffusion value $D = K_B T/f$ to that given in eq II.7. The effect of spread in charges δ, however, may be seen to reduce $\bar{D}$ by an amount $2nI\delta^2[K_B T/(K_B T_1 + 2nIZ^2) - 1]$. This reduction occurs because the molecules will preferentially arrange themselves so that molecules of charge $Z + \delta$ tend to have as neighbors molecules of charge $Z - \delta$. Such a rearrangement from a random molecular configuration reduces the average intermolecular electrostatic interaction energy.

In order to compare these results with experiment, it is convenient to re-express eq II.7 in terms of measurable parameters. While the average charge $\bar{Z}$ can be determined from a protein titration curve, values for the interaction strength I can only be determined indirectly. We will, therefore, describe our results, not in terms of I, but in terms of the parameter α

$$\alpha = (\bar{D} - D_0)/D_0 \qquad \text{(II.8)}$$

where D_0 is the mutual diffusion coefficient in the limit of high salt concentration, i.e., the limit in which $I \to 0$. α was previously used by Doherty and Benedek[4] to describe their measurements on the diffusion of serum albumin in solutions of low salt concentration.

Figure 1 presents (dashed lines) values for the variance, as a function of α and δ, evaluated at an average charge $\bar{Z} = -10$. The points indicate the measurements of Doherty and Benedek on serum albumin with the same average charge. It is seen that σ initially increases sharply with increasing α and then tends to saturate for $\alpha \gtrsim 2$. Figure 1 suggests that by measuring the dependence of the variance on the salt concentration, it should be possible to estimate the magnitude of the charge heterogeneity, at least in some monodisperse systems. A refinement of the above calculation, so as to use a better approximation for the macromolecule charge distribution, should give better values for δ.

Figure 1 also indicates the major limitation of this technique; there are other sources for the variance. Doherty and Benedek found that the variance of their bovine serum albumin solutions was never less than 20%, even in 1.0 M NaCl. In many systems, reducing the salt concentration will change the degree of aggregation in the system, thereby changing the contribution of polydispersity to the variance. While further experimental study is necessary, aggregation introduced by changes in the salt concentration will probably, in many systems, prevent one from using the dependence of the variance on salt concentration to estimate the charge heterogeneity.

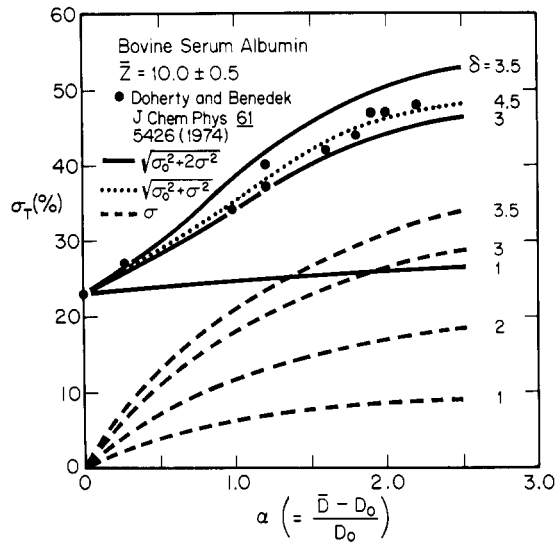

Figure 1. Variance σ $(= \sqrt{\bar{D}^2 - \bar{D}^2}/\bar{D})$ in the spectrum of the light scattered from a solution of charged macromolecules ($\bar{Z} = -10$), displayed as a function of effective intermolecular interaction strength α and RMS macromolecular charge fluctuation δ. Dashed lines indicate the predicted variance for a monodisperse system. Solid and dotted lines show the predicted variance for an initially polydisperse system on different assumptions as to the total variance expected for such a system.

However, as pointed out by Timasheff et al.,[14] in isoionic solutions light-scattering intensity measurements can be used to identify aggregation and charge fluctuation effects. The former's contribution to the turbidity is linear in the protein concentration, while the latter contributes to the turbidity only as the square root of the protein concentration. Timasheff et al. found that isoionic serum albumin does not aggregate appreciably when the ionic strength of the solution is reduced. Thus, in comparing the experimental data of Doherty and Benedek with the model calculation, it is not unreasonable to assume that the polydispersity of their system is not appreciably dependent on the salt concentration.

Because the oligomers interact with each other and with the solute monomers, reducing the salt concentration can change the contribution of the other species to the variance. The solid and dotted lines of Figure1 are estimates of the total expected variance σ_T of the system. σ_0 is the background variance seen at high salt concentration, while σ is the electrostatic contribution to the variance. The dotted and solid lines correspond to $\sigma_T = (\sigma_0^2 + \sigma^2)^{1/2}$ and $\sigma_T = (\sigma_0^2 + 2\sigma^2)^{1/2}$. Physically, one expects that the increase in I increases both the spread in the sets of exponentials contributed by monomer and oligomer and the spread between these two sets of exponentials. Additional information about the source of the observed variance σ_0 would permit a better estimate of the total variance σ_T and its dependence on α.

If $3.0 \leq \delta \leq 4.5$, either of the above assumptions gives a good estimate of the total variance. This value for δ compares favorably with the results of Timasheff et al., who used light-scattering intensity measurements to estimate that for isoionic bovine serum albumin $\delta \approx 3.6$–3.7.

While the values of δ obtained by the two light-scattering techniques are in agreement, they are consistently higher than the value expected from the titration curve of serum albumin[12] and the Linderstrom–Lang equation,[15]

$$\langle Z^2 \rangle - \langle Z \rangle^2 = \frac{1}{2.302}\frac{\partial Z}{\partial \mathrm{pH}} \qquad \text{(II.9)}$$

which together predict $\delta \approx 1.8$ for this pH range. At pH 7.0,

charge fluctuations of serum albumin will probably arise primarily from its 16 imidazole groups, whose pK is near 7.0;[12] as indicated in Table I, fluctuations in the charge on these groups are sufficiently slow that light scattering will measure the different imidazole charge states as distinct chemical species. The discrepancy between the value of δ predicted by the titration curve and the value of δ found by light scattering probably arises from microheterogeneities in the serum albumin sample, leading to the presence of subspecies of serum albumin with different isoionic pH's.[16] As long as each subspecies has the same titration curve, use of the Linderstrom–Lang equation will predict the part of the total charge fluctuation due to members of the same subspecies having different charges, but not the part of the charge fluctuation caused by the presence of several subspecies of serum albumin with different average charges.

III. Discussion

The contribution of intermolecular interactions to the variance in the light-scattering spectrum might reasonably be expected on simple physical grounds. The random forces on each solute molecule cause it to diffuse through the solution. If solute–solute interactions be present, they can contribute to diffusion, their contribution being a particular average over the force between solute molecules. If several species of solute molecules are present, the averages over each of the species–species interactions will give rise to a range of diffusion coefficients. The range of exponentials in the spectrum, as measured by the variance, may then be expected to be proportional to the range of intermolecular interaction strengths in the system, which is precisely the result found above in a simple model.

When the ionic strength of a protein solution is reduced, it is not uncommon for the variance of its spectrum to increase. It is easy to interpret such spectral changes as arising from an increased polydispersity of the sample at low ionic strength. However, as is shown in the above, one may not unambiguously associate a high variance with a substantial degree of polydispersity. Intermolecular interactions may also act to increase the variance. Timasheff et al.[14] have noted how a study of concentration dependence can, in a special case, resolve between aggregation and charge fluctuation effects; however, this convenient differentiation is probably only useful for nearly isoionic systems, for which (since $Z \approx 0$) the variance contributed by electrostatic effects (cf. eq II.7) will be small.

References and Notes

(1) (a) Contribution No. 3602. (b) This research was supported in part by grants from the National Science Foundation (Grant No. NSF MPS 7424612A-03) and the National Heart and Lung Institute (HL 14322-03, R. W. Mann, Principal Investigator).
(2) (a) S. B. Dubin, J. H. Lunacek, and G. B. Benedek, *Proc. Natl. Acad. Sci. U.S.A.*, **57**, 1164 (1967). (b) D. Koppel, *J. Chem. Phys.*, **57**, 4814 (1972); N. M. Mazer, S. B. Thesis, MIT (1972).
(3) L. Hocker, J. Krupp, G. Benedek, and J. Vournackis, *Biopolymers*, **12**, 1677 (1973).
(4) P. Doherty and G. B. Benedek, *J. Chem. Phys.*, **61**, 5426 (1974), and references cited therein.
(5) G. D. J. Phillies, N. M. Mazer, and G. B. Benedek, unpublished; G. D. J. Phillies, *J. Chem. Phys.*, **60**, 976 (1974).
(6) M. M. Stephen, *J. Chem. Phys.*, **55**, 3878 (1971).
(7) P. N. Pusey, D. W. Schaeffer, D. E. Koppel, R. D. Camerini-Otero, and R. M. Franklin, *J. Phys. (Paris), Colloq.*, **33**, 163 (1972); D. W. Schaeffer and B. J. Berne, *Phys. Rev. Lett.*, **32**, 1110 (1974); J. C. Brown, P. N. Pusey, J. W. Goodwin, and R. M. Ottewill, *J. Phys. A: Gen. Phys.*, **8**, 644 (1975).
(8) G. D. J. Phillies, *J. Chem. Phys.*, **60**, 983 (1974).
(9) J. G. Kirkwood and R. J. Goldberg, *J. Chem. Phys.*, **18**, 54 (1950).
(10) B. R. Ware and W. M. Flygare, *Chem. Phys. Lett.*, **12**, 81 (1971); B. Berne and R. Giniger, *Biopolymers*, **12**, 1161 (1973).
(11) M. Eigen and L. DeMaeyer, "Technique of Organic Chemistry", Vol. VIII, Part II, A. Weissberger, Ed., Interscience, New York, N.Y. 1963, p 1031.
(12) C. Tanford, S. A. Swanson, and W. S. Shore, *J. Am. Chem. Soc.*, **77**, 6414 (1955).
(13) N. M. Mazer, private communication.
(14) S. N. Timasheff, N. M. Dintzis, J. G. Kirkwood, and B. D. Coleman, *Proc. Natl. Acad. Sci. U.S.A.*, **41**, 710 (1955); S. N. Timasheff, H. M. Dintzis, J. G. Kirkwood, and B. D. Coleman, *J. Am. Chem. Soc.*, **79**, 782 (1957).
(15) J. T. Edsall and J. Wyman, "Biophysical Chemistry", Vol. 1, Academic Press, New York, N.Y., 1958, p 511.
(16) L. S. Baskin (*J. Phys. Chem.*, **72**, 2958 (1968)) has previously suggested this possibility.