# Modeling Biological Macromolecules in Solution. II. The General Tri-axial Ellipsoid

STEPHEN E. HARDING, *Department of Biochemistry, University of Bristol, Bristol BS8 1TD, England;* and ARTHUR J. ROWE, *Department of Biochemistry, University of Leicester, Leicester, LE1 7RH, England*

## Synopsis

The use of the general (tri-axial) ellipsoid as a model for the conformation of biological macromolecules in solution is discussed. The recent derivation of an expression for the viscosity increment of dispersions of tri-axial ellipsoids [S. E. Harding et al. (1979) *IRCS Med. Sci.* **7**, 33; and (1981) *J. Colloid Interface Sci.* **79**, 7–13] makes the fitting of such a model possible by the derivation of appropriate volume-independent functions. We now derive these volume-independent functions for tri-axial ellipsoids and investigate by exhaustive computer simulation the possibility of deriving the two axial ratios $(a/b, b/c)$ from data of various types, in every case containing plausible experimental error. It is shown that transport properties alone cannot be used to yield estimates for the axial ratios, given current experimental precision. However, a combination of transport and rotational diffusion properties is more promising, and an algorithm is developed and tested that will reliably yield estimates from simulated data obtained by the methods of sedimentation, viscosity, and electric birefringence.

## INTRODUCTION

The most general and almost ubiquitously applied shape for modeling biological macromolecules in solution has been the ellipsoid of revolution model (i.e., an ellipsoid with two equal axes).[1–5] For globular proteins, a much better model, in principle at least, might be obtained if the restriction of the two equal axes on the ellipsoid were removed to allow use of the more general tri-axial ellipsoid. The recent derivation of the viscosity increment $\nu$ for tri-axial ellipsoids[6,7] represents a significant step forward, although a given value of $\nu$ does not uniquely fix a value for the two independent axial ratios. We now show by combining this result with information from sedimentation velocity and electric birefringence decay that it is possible, with the currently available experimental precision, to uniquely determine the axial dimensions of a macromolecule in solution in terms of a general tri-axial ellipsoid model. This model is independent of any assumptions concerning the swelling of the macromolecule due to solvent association other than that it is the same in the three types of experiment.

# THEORY

## Viscosity Increment $\nu$ for Tri-axial Ellipsoids

The fact that a dissolved macromolecular solute increases the viscosity of a fluid has been well described, and Simha[8] and Saito[9] related this increase to axial ratio for ellipsoids of revolution in terms of the viscosity increment $\nu$. An explicit expression in terms of axial ratio $(a/b, b/c)$ for the viscosity increment of a dilute dispersion of general ellipsoids $(a \geqslant b \geqslant c)$ has recently been given[6,7] assuming that the particles rotate on average with the local undisturbed angular velocity of the fluid:

$$\nu \equiv \frac{[\eta]}{v_s} = \frac{1}{abc} \left\{ \frac{4(\alpha_0'' + \beta_0'' + \gamma_0'')}{15(\beta_0''\gamma_0'' + \gamma_0''\alpha_0'' + \alpha_0''\beta_0'')} + \frac{1}{5} \left[ \frac{\beta_0 + \gamma_0}{\alpha_0'(b^2\beta_0 + c^2\gamma_0)} \right. \right.$$
$$\left. \left. + \frac{\gamma_0 + \alpha_0}{\beta_0'(c^2\gamma_0 + a^2\alpha_0)} + \frac{\alpha_0 + \beta_0}{\gamma_0'(a^2\alpha_0 + b^2\beta_0)} \right] \right\} \quad (1)$$

where $[\eta]$ is the intrinsic viscosity (mL/g), $v_s$ the swollen specific volume (mL/g), $M_r$ the molecular weight, $V_e$ the volume of a macromolecule (mL), $N_A$ Avogadro's number, and the $\alpha_0'$, etc., are elliptic integrals defined by Jeffrey.[10] This assumption had only been rigorously proved for ellipsoids of revolution, but it can be shown to be extremely accurate for globular particles $(a/b: 1.0 \rightarrow 3.0, b/c: 1.0 \rightarrow 3.0)$ and for tapes $(a \gg b \gg c)$. For certain particles of intermediate asymmetries, the value of a small extra term (Ref. 11; and J. M. Rallison and H. Brenner, personal communications), representing the departure of particles from rotating with the fluid is generally no more than 1% of the RHS of Eq. (1):

$$- \frac{1}{5abc} \left\{ \left[ \frac{a^2 - b^2}{a^2\alpha_0 + b^2\beta_0} + \frac{b^2 - c^2}{b^2\beta_0 + c^2\gamma_0} + \frac{c^2 - a^2}{c^2\gamma_0 + a^2\alpha_0} \right]^2 \middle/ \right.$$
$$\left. \left[ \frac{a^2 + b^2}{a^2\alpha_0 + b^2\gamma_0} + \frac{b^2 + c^2}{b^2\beta_0 + c^2\gamma_0} + \frac{c^2 + a^2}{c^2\gamma_0 + a^2\alpha_0} \right] \right\} \quad (1')$$

The elliptic integrals in Eqs. (1) and (1') can be solved numerically using standard computational routines,[12] and hence $\nu$ can be given for any value of $(a/b, b/c)$. Although a given value of $(a/b, b/c)$ uniquely fixes a value of $\nu$, the converse is not true, as is clear from inspecting a table of values of $\nu$ as a function of $(a/b, b/c)$[6,7]; rather, a given value of $\nu$ has a *line solution* of possible values of $(a/b, b/c)$. In order to determine a unique solution for $(a/b, b/c)$ and hence the axial dimensions of a macromolecule in solution, other hydrodynamic information must be used; we must therefore consider the translational and rotational frictional properties. Scattering of light or x-rays can also be used to give useful information on, for example, molecular weight, but any interpretation of the data in terms of shape suffers from the serious drawback that the macromolecule is assumed to have uniform electron density.

## Translational Frictional Ratio: The $\beta$ and $R$ Functions

Although Perrin[13,14] had, as long ago as 1936, provided an explicit expression for the translational frictional ratio ($f/f_0$ or $P$) for a tri-axial ellipsoid:

$$\frac{f}{f_0} \equiv P = \frac{M_r(1 - \bar{v}\rho_0)}{N_A 6\pi\eta_0 s}\left(\frac{4\pi}{3V_e}\right)^{1/3}$$

$$= \frac{2}{(abc)^{1/3}} \int_0^\infty \frac{d\lambda}{[(a^2 + \lambda)(b^2 + \lambda)(c^2 + \lambda)]^{1/2}} \quad (2)$$

(where $f$ is the frictional coefficient of the macromolecule, $f_0$ the corresponding coefficient for a hydrated sphere of the same volume and molecular weight, and $\eta_0, \rho_0$ the viscosity and density, respectively, of the solvent), the elliptic integral in Eq. (2) could only be solved for the special case of ellipsoids of revolution. This restriction is, however, no longer necessary with the availability of high-speed computers and can be solved using standard numerical routines.[12] As with $\nu$, a given value of $P$ has a line solution of possible values of $(a/b, b/c)$. However, in principle at least, by combining graphically both line solutions, a unique solution can be found from their intersection. [This is analogous to the Oncley[15] treatment for determining solvation and axial ratio for ellipsoids of revolution. There is no possibility of an analytic simultaneous solution to the problem corresponding to the Scheraga and Mandelkern[14] treatment, since expressions for $\nu$ and $P$ in terms of individual axial ratios ($a/b$ or $b/c$) are not separable from Eqs. (1) or (2).]

In Fig. 1 we have assumed a particle of axial ratios $(a/b, b/c) = (1.5, 1.5)$, computed the corresponding values of the parameters $\nu$ and $P$ (accurate to four significant figures), and then plotted the line solutions (Ref. 16). Unfortunately, Fig. 1 reveals that the intersection between the two line solutions is very shallow and allowing for $\pm 1\%$ experimental error in each, there is no intersection at all in the "globular macromolecule" range of Fig. 2. There is also the additional problem that in order to determine both $\nu$ and $P$, we need to know the (swollen) volume of the macromolecules in solution.

However, since both the Scheraga and Mandelkern $\beta$ and Rowe $R$ volume-independent shape parameters[3,5,14] are functions only of $\nu$ and $P$, they too can be computed for tri-axial ellipsoids:

$$\beta = \frac{N_A^{1/3}}{(16,200\pi^2)^{1/3}} \frac{\nu^{1/3}}{f/f_0} \equiv \frac{N_A \, s[\eta]^{1/3}\eta_0}{M_r^{2/3}(1 - \bar{v}\rho_0)100^{1/3}}$$

$$R = \frac{2}{\nu}\left(1 + \left(\frac{f}{f_0}\right)^3\right) \equiv \frac{k_s}{[\eta]}$$

where $k_s$ is the sedimentation concentration regression coefficient.[3]

As can be seen from Fig. 2, there is no reasonable intersection between the four line solutions—the sensitivity of the $\beta$ function to error is such that
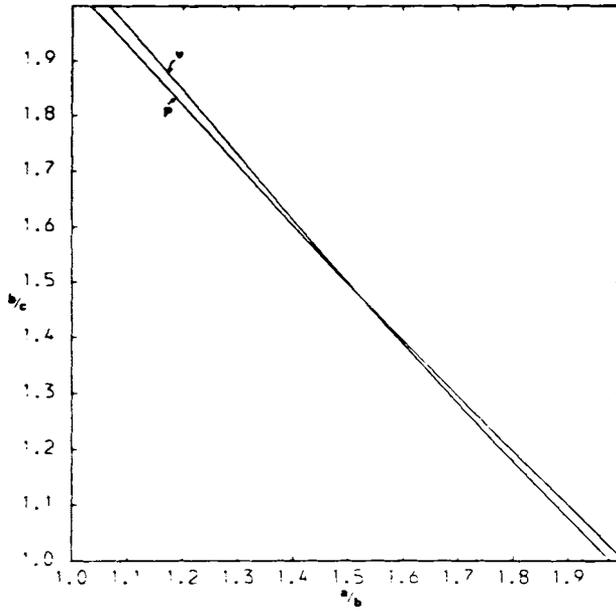
Fig. 1. Plots of the contours of constant value for the functions $\nu$ and $P$ in the $(a/b, b/c)$ plane, computed for the values $a/b = 1.5$, $b/c = 1.5$ for the two axial ratios of a tri-axial ellipsoid.
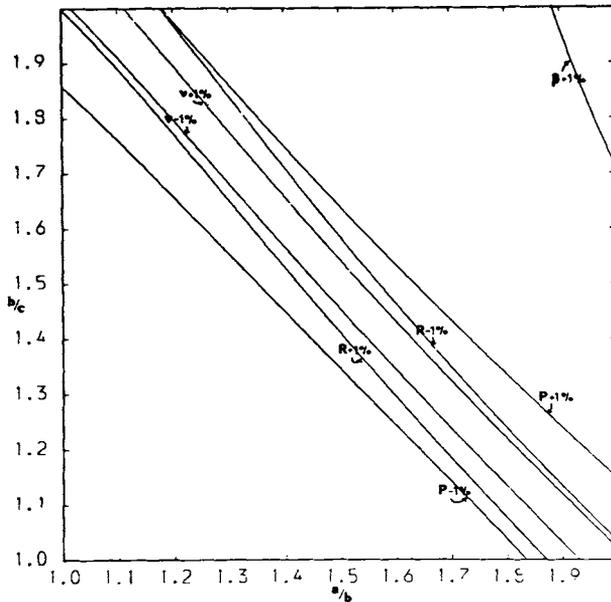


Fig. 2. Plots of the contours of constant value for the functions $\nu$, $P$, $\beta$, and $R$ at $\pm 1\%$ levels to simulate error in measured values.  Other details as in Fig. 1.

the $\beta$-1% line is off the area mapped. An important conclusion to be drawn from this is that no combination of measurements of translational parameters can suffice to determine the two axial ratios of a tri-axial ellipsoid.

An obvious alternative approach to the problem is to combine the most precisely determinable volume-independent translational function $(R)$ with one or more rotational frictional or rotational relaxation functions. Again, no analytical solution will be possible, and we consider the use of graphical solutions applied to plausible data with simulated error.

## Rotational Relaxation Line Solutions

Solutions for the three rotational frictional ratios $\zeta_i/\zeta_0$ $(i = a, b, c)$ and the three dielectric dispersion relaxation time ratios are now available[5] but are impossible to resolve experimentally. The same is true for the five fluorescence anisotropy relaxation time ratios.[17] It therefore follows that since all these functions are of no apparent use at the moment, the same must be true of compound volume-independent functions involving them.

The tri-axial $\Psi$ function is also now available but is extremely insensitive to axial ratio. The same is unfortunately true for the tri-axial $\Lambda$ function in the globular particle range.[5] For volume-independent functions involving electric birefringence decay parameters, the situation is more encouraging.

## Electric Birefringence Decay: The $\delta_+$ and $\delta_-$ Functions

The decay of electric birefringence for a homogeneous dispersion of asymmetric macromolecules consists of two exponential terms[18]:

$$\Delta n = \frac{N}{2n_l}\{A_+ e^{-6\theta_+ t} + A_- e^{-6\theta_- t}\} \tag{3}$$

where $\Delta n$ is the birefringence, $N$ the number density of particles in suspension, and $n_1$ the refractive index of the suspending medium. $A_+$ and $A_-$ are complicated functions depending on the initial orientation of the particles and their dielectric and diffusion properties. We may rewrite $NA_\pm/2n_l$ as the "pre-exponential factors" $A'_\pm$, and Eq. (3) becomes

$$\Delta n = A'_+ e^{-6\theta_+ t} + A'_- e^{-6\theta_- t}$$

$\Theta_+$ and $\Theta_-$ are related to the rotational diffusion coefficients $\Theta_i$ (and hence the rotational frictional coefficients, since $\zeta_i = kT/\Theta_i$) by[18]

$$\theta_\pm = \tfrac{1}{3}\sum\theta_i \pm \left\{\left(\tfrac{1}{3}\sum\theta_i\right)^2 - \tfrac{1}{3}\sum\sum_{i>j}\theta_i\theta_j\right\}^{1/2}$$

$$= \frac{kT}{3}\left\{\sum_i\frac{1}{\zeta_i} \pm \left(\sum_i\frac{1}{\zeta_i^2} - \sum\sum_{i>j}\frac{1}{\zeta_i\zeta_j}\right)^{1/2}\right\}$$

The dimensions of this equation are of energy/(volume × viscosity); we therefore "reduce" it to a function of shape alone[18]:

$$\theta_\pm^{\text{red}} \equiv \left(\frac{\eta_0}{kT}\right) V_e \theta_\pm = \frac{abc}{12}\left\{\left(\frac{1}{\zeta_a''}+\frac{1}{\zeta_b''}+\frac{1}{\zeta_c''}\right)\right.$$
$$\left.\pm\left[\left(\frac{1}{\zeta_a''^2}+\frac{1}{\zeta_b''^2}+\frac{1}{\zeta_c''^2}\right)-\left(\frac{1}{\zeta_a''\zeta_b''}+\frac{1}{\zeta_b''\zeta_c''}+\frac{1}{\zeta_c''\zeta_a''}\right)\right]^{1/2}\right\}$$

where

$$\zeta_a'' = \frac{b^2+c^2}{b^2\beta_0+c^2\gamma_0}\ ,\qquad \zeta_b'' = \frac{c^2+a^2}{c^2\gamma_0+a^2\alpha_0}\ ,\qquad \zeta_c'' = \frac{a^2+b^2}{a^2\alpha_0+b^2\beta_0}$$

and the elliptic integrals $\alpha_0$, etc., are defined by Jeffrey.[10]

Related volume-independent functions can be obtained by combining $\theta_\pm^{\text{red}}$ either with the viscosity increment or the translational frictional ratio:

$$\delta_\pm = 6\theta_\pm^{\text{red}}\nu \equiv \frac{6}{N_Ak}\left(\frac{\eta_0\theta_\pm}{T}\right)[\eta]M_r \tag{4}$$

$$\gamma_\pm = 6\theta_\pm^{\text{red}}\left(\frac{f}{f_0}\right)^3 \equiv \frac{M_r^3(1-\bar{v}\rho_0)^3\theta_\pm}{27N_AkT\pi^2\eta_0^2s^3} \tag{5}$$

The $\delta_\pm$ and $\gamma_\pm$ functions are new and $\delta_\pm$ plotted, with the $R$ function, in Fig. 3 (the $\gamma_\pm$ vs $R$ plot is very similar to the $\delta_\pm$ vs $R$ plot of Fig. 3). It is
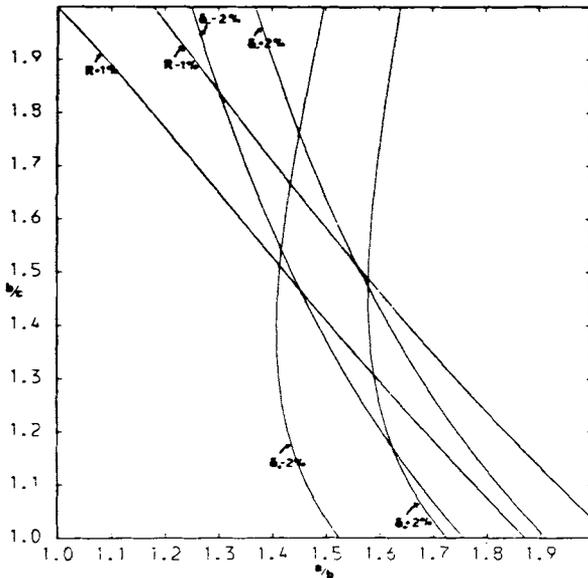


Fig. 3. Plots of the contours of constant value for the functions $R$, $\delta_+$, and $\delta_-$ at $\pm 1\%$ levels. Other details as in Fig. 1.

seen that both give very reasonable intersections with $R$ and are sensitve to axial ratio. The $\delta_\pm$ functions are preferred over the $\gamma_\pm$, since they require fewer experimental measurements and do not involve squared or cubed terms, and hence, in principle, can be measured more accurately. There remains, however, the problem of measuring $\Theta_\pm$, i.e., resolving the exponential decay into its two component decay constants. To date, reliable resolution has not been possible except for very accurate synthetic data—far beyond current experimental accuracy. We now investigate aspects of this problem and show that with a new "constrained" least-squares algorithm using intersection with the $R$ curve as the constraint, this is now possible with currently available experimental precision.

## NUMERICAL METHODS

Two detailed reviews of methods for resolving multiexponential decays have recently been given by Jost and O'Konski[16] and O'Connor, Ware, and Andre.[19] It is clearly evident from their work that the nonlinear least-squares iterative procedure is the best available method for resolving a two-term exponential birefringence decay. In this method, the weighted sum of the squares of the residuals, $\chi^2$ is calculated between a set of experimental data points and the function to be fitted. If $x_j$ represents the value of the $j$th experimental point and $\xi_j(X_m)$ the corresponding computer point for a given estimate for the $X$ (the independent variables), then we define our "goodness of fit" parameter $\chi^2$ by

$$\chi^2 = \sum_{j=1}^{n} \left( \frac{x_j - \xi_j}{\sigma_j} \right)^2$$

where $\sigma_j$ is the standard error on the $j$th experimental point. The best values of the $X_m$ are such that $\partial\chi^2/\partial X_m = 0$, for all the $X_m$. For the particular case of electric birefringence, $\sigma_j$ is approximately constant for all the $x_j$,[16] and the minimization condition becomes

$$\frac{\partial F}{\partial X_m} = 0$$

where

$$F = \sum_{j=1}^{n} \{x_j - \xi_j\}^2$$

In the case of a two-term birefringence decay, the minimization condition is said to be "nonlinear" in that the data are to be fitted to a function that is the sum of terms each consisting of the product of an adjustable parameter (i.e., a pre-exponential factor) with the function of another adjustable parameter (i.e., a decay constant or relaxation time). In order to evaluate $\partial F/\partial X_m$ for a current estimate of the parameters $X_m$, the solution either has to be linearized using a Taylor expansion as outlined by Jost and O'Konski,[16] or alternatively a quadratic or quasi-Newtonian procedure can

be employed: Gill and Murray's[20] quadratic algorithm is particularly attractive in that upper and lower limits for adjustable parameters can be specified and included as external constraints.

## Computer Simulation

The requirement on the accuracy of the experimental data was tested by computer simulation: Three hypothetical ellipsoidal proteins of known (tri-axial) dimensions and hence axial ratio ($a/b$, $b/c$) were assumed to have a partial specific volume $\bar{v}$ of 0.73 and a swelling ratio ($v_s/\bar{v}$), (where $v_s$ is the swollen specific volume) of 1.3 (typical for globular proteins). From these values the molecular weight $M_r$, viscosity increment $v$, $R$ function, $\delta_\pm$ functions, intrinsic viscosity $[\eta]$, and hence decay constants $\Theta_\pm$ could be predicted for each protein (Table I). Pre-exponential factors $A'_+$ and $A'_-$ of 0.07 and 0.05 rad were also assumed [taken from an initial birefringence ($\equiv A'_+ + A'_-$) of 0.12 rad], and hence the unperturbed decay curve for each simulated protein could be given.

TABLE I
Assumed and Derived Characteristics of Three Hypothetical Globular Proteins

| Parameter | Protein 1 | Protein 2 | Protein 3 |
|---|---|---|---|
| | Assumed Values | | |
| $a, b, c$ (Å) | 45, 30, 20 | 42.5, 25, 20 | 42.5, 34, 20 |
| $\bar{v}$ (mL/g) | 0.730 | 0.730 | 0.730 |
| $\bar{v}_s/\bar{v}$ | 1.3 | 1.3 | 1.3 |
| | Derived Values | | |
| $a/b, b/c$ | 1.50, 1.50 | 1.70, 1.25 | 1.25, 1.70 |
| $\bar{v}_s$ (mL/g) | 0.949 | 0.949 | 0.949 |
| Swollen molecular volume, $V_e = \frac{4}{3}\pi abc$ (cm³) | $1.131 \times 10^{-19}$ | $0.890 \times 10^{-19}$ | $1.211 \times 10^{-19}$ |
| Anhydrous molecular volume, $V [= (\bar{v}/\bar{v}_s)V_e]$ (cm³) | $0.870 \times 10^{-19}$ | $0.685 \times 10^{-19}$ | $0.931 \times 10^{-19}$ |
| Molecular weight, $M_r [= (N_A/\bar{v})V]$ | 71,744 | 56,510 | 76,853 |
| $v$ | 2.892 | 2.870 | 2.840 |
| $[\eta]$ ($= N_A V_e v/M_r$) (mL/g) | 2.75 | 2.72 | 2.695 |
| $R$ | 1.479 | 1.482 | 1.496 |
| $\theta_+^{red}, \theta_-^{red}$ | 0.163, 0.116 | 0.171, 0.115 | 0.155, 0.125 |
| $\delta_+, \delta_-$ | 2.821, 2.016 | 2.943, 1.982 | 2.645, 2.125 |
| Decay constants,[a] | | | |
| $\theta_\pm = \dfrac{N_A kT}{6\eta_0[\eta]M_r}\delta_\pm$ (s$^{-1}$) | $5.815 \times 10^6$, $4.156 \times 10^6$ | $7.766 \times 10^6$, $5.229 \times 10^6$ | $5.187 \times 10^6$, $4.167 \times 10^6$ |
| Relaxation times, $\tau_\pm = 1/6\theta_\pm$ (ns) | 28.660, 40.010 | 21.461, 31.873 | 32.130, 39.992 |

[a] $T = 293$ K, $\eta_0 = 0.01$ g cm$^{-1}$ s$^{-1}$

One then places simulated experimental error on each of 100 data points for each decay curve using a computer normal ("Gaussian") pseudo-random number generator,[12] and first of all assuming no errors in the molecular weight or intrinsic viscosity, investigate how much error in the decay data points is tolerable before the algorithm becomes unstable, i.e., fails to give back the correct decay constants and hence axial ratios within reasonable limits. If successful, the algorithm would then be tested for the effect of errors in the intrinsic viscosity and molecular weight (as well as the error in the decay curve) and for various initially assumed values of $A'_+$ and $A'_-$ (which in the analyses are, of course, regarded, with the $\Theta_\pm$, as unknown variables).

## Nonlinear Least-Squares Method

The quasi-Newtonian quadratic method for minimizing any function (i.e., in this case the sum of the squares of the residuals, $F$) given by Gill and Murray[20] and incorporated in the UK,NAG mk VI[12] subroutine EØ4JAF was used. In this algorithm the user, besides supplying the subroutine for calculating the value of $F$ at any point $X_m$, has also to supply fixed upper and lower bounds on the independent variables $X_1 \cdots X_m \cdots X_N$. This routine was incorporated in a FORTRAN IV program that generated its own hypothetical decay curve with normal pseudo-random error generated on each data point (using NAG routine GØ5ADF), the amount specifiable by the user. The program attempted to retrieve the decay constants, hence the $\delta_\pm$ functions (from $[\eta]$, $M_r$) and hence, from a plot (Fig. 3), the axial ratios $(a/b, b/c)$. Owing to the problem of the danger of the routine falling into subsidiary minima,[16] it was necessary to repeat the method for a large number (30) of initial guesses.

Unfortunately, even synthetic birefringence data as accurate as $0.001°$ standard error (SE) on each data point [about two orders of magnitude greater than the current experimental precision (B. Jennings and V. Morris, personal communication)] failed to give back the correct $(a/b, b/c)$ within reasonable limits, and for data of machine accuracy (14 significant figures), the exact value was not retrieved, as Fig. 4 illustrates.

## A New $R$-Constrained Nonlinear Least-Squares Algorithm

A significant improvement in the least-squares algorithm can be gained if information from the $R$ function line solution, which can be found relatively accurately and easily,[3] is included as a constraint. The problem is effectively reduced from one of four independent variables $(\Theta_+, \Theta_-, A'_+, A'_-)$ to one of three $(a/b, A'_+, A'_-)$. The solution is constrained to lie on the $R$ curve; thus, a given estimate for $a/b$ will necessarily give a "constrained" value for $b/c$; the computer program can then calculate the values for $\delta_+$
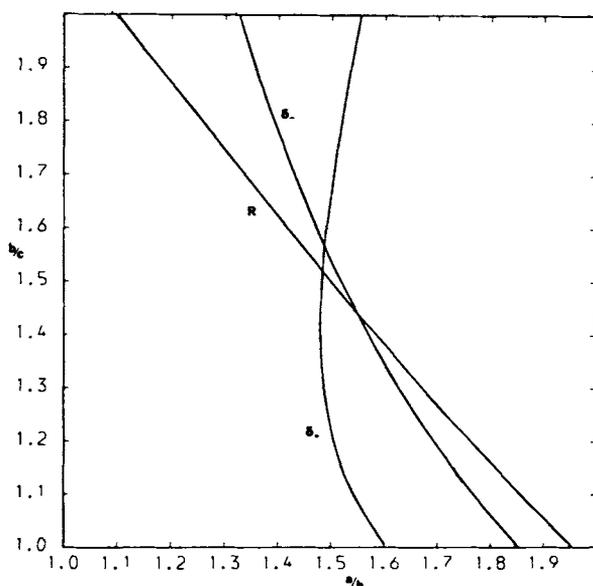
Fig. 4. Plots of the contours of constant value for the functions $R$, $\delta_+$, and $\delta_-$ (as in Fig. 3) assuming no error greater than computer rounding error in supplied data. The nonlinear least-squares method (see text) was used.

and $\delta_-$ corresponding to this estimate and, hence, the decay constants $\Theta_\pm$ [using also the values for $[\eta]$ and $M_r$—Eq. (4)], the corresponding decay curve, and finally the sum of the squares of the residuals (SSR) between the computer points and the "experimental" curve. By iterating along the $R$ curve for $a/b$ and the two pre-exponential factors $A'_\pm$, the best estimate for $(a/b, b/c)$ can be found from the minimum value of the SSR. The constraint of the $R$ curve was included in the algorithm by use of a cubic spline routine: the coordinates of knots in the curve are specified in the program (or alternatively, the whole curve digitized) and the routine interpolates between these points using a cubic polynomial spline fit. Numerical integration of the elliptic integrals for determining values of $\delta_\pm$ was accomplished using the NAG routine DØ1GAF.[12]

It was found in pilot runs of the complete program that the danger of the algorithm falling into subsidiary minima[16] was no longer significant. The number of initial guesses was thus reduced from 30 to 3 to save on computer time: the final estimates were generally the same for all three initial guesses. The values of the $(a/b, b/c)$ retrieved did, however, depend on the cutoff time specified for the decay curve. If there were no error in the decay data points, then very long cutoff times would be desirable, since this region is dominated by the longer relaxation time (or shorter decay constant, $\Theta_-$). However, the effect of a given absolute error is more pronounced the lower the birefringence signal.

The optimum cutoff time, and hence the best value for $(a/b, b/c)$ was

TABLE II
Determination of the Optimum Cutoff Time for Protein 1:   True $(a/b, b/c) = (1.5, 1.5)$[a]

| | a/b at Cutoff Time (ns) | | | | | |
| | 80 | 100 | 110 | 115 | 120 | 140 |
|---|---|---|---|---|---|---|
| Stream 1 | 1.580 | 1.534 | 1.513 | 1.503 | 1.493 | 1.454 |
| Stream 2 | 1.946 | 1.785 | 1.692 | 1.654 | 1.619 | 1.497 |
| Stream 3 | 1.591 | 1.512 | 1.483 | 1.468 | 1.452 | 1.392 |
| Stream 4 | 1.644 | 1.487 | 1.425 | 1.396 | 1.367 | 1.249 |
| Stream 5 | 1.623 | 1.480 | 1.426 | 1.401 | 1.377 | 1.287 |
| Stream 6 | 1.186 | 1.275 | 1.303 | 1.315 | 1.326 | 1.364 |
| Stream 7 | 1.573 | 1.645 | 1.678 | 1.694 | 1.710 | 1.772 |
| Stream 8 | 1.716 | 1.623 | 1.590 | 1.575 | 1.562 | 1.514 |
| Mean | 1.607 | 1.543 | 1.514 | 1.509 | 1.488 | 1.441 |
| $\sigma_D$[b] | 0.210 | 0.149 | 0.134 | 0.132 | 0.134 | 0.163 |
| $\sigma_E$[c] | 0.0741 | 0.0527 | 0.0473 | 0.0468 | 0.0475 | 0.0578 |

Optimum cutoff time = 115 ns
Best estimate for $a/b$ = 1.501 ($\pm$0.047)
Corresponding estimate for $b/c$ = 1.498

[a] No assumed error in $R$.
[b] $\sigma_D$, standard deviation.
[c] $\sigma_E$, standard error.

found by repeating for eight streams of normal random data, specified by the NAG routine[12] GØ5BAF(0.$N$), where $N$ represents the stream number of the random data:   the optimum cutoff time for each decay curve was then determined by finding the best standard deviation ($\sigma_D$) or standard error ($\sigma_E$) of the $a/b$'s from the eight streams for increments of 5 ns in the cutoff times.   The corresponding best mean value for $a/b$ (and hence $b/c$), together with the corresponding standard error for the eight streams of data, could then be found (Table II).   This procedure was then repeated allowing for $\pm$1% experimental error in the $R$ curves.   If the points on the three curves corresponding to $\langle(a/b, c/d)\rangle + \sigma_E$ (where $\langle$ $\rangle$ denotes the mean value) are joined and then similarly those of $\langle(a/b, b/c)\rangle - \sigma_E$, regions of allowed values for $(a/b, b/c)$ permitted by the data would then be found. Figures 5–8 illustrate the results for birefringence decay data of 0.1° SE on each data point (about the current experimental precision):   the mean values are seen to agree very closely with the actual values (Table III).

TABLE III
Mean Values for the Retrieved Axial Ratios Compared with Real Values

| Protein | Retrieved $(a/b, b/c)$ | Real $(a/b, b/c)$ |
|---|---|---|
| 1 | (1.501, 1.498) | (1.50, 1.50) |
| 2 | (1.652, 1.305) | (1.70, 1.25) |
| 3 | (1.284, 1.695) | (1.25, 1.70) |
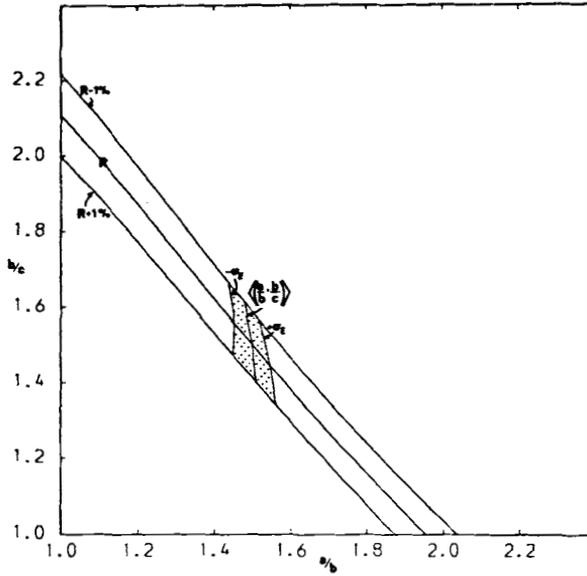
Fig. 5. Contour plots for a hypothetical Protein 1 (see text) showing the area (dotted) within which estimates for the axial ratios are found to lie by use of the $R$-constrained least-squares algorithm (see text). Simulated experimental error of $\pm 1°$ on each data point in the electric birefringence decay curve was assumed. Error in $R$ of $\pm 1\%$ was assumed. [True $(a/b, b/c)$ = (1.5, 1.5).]
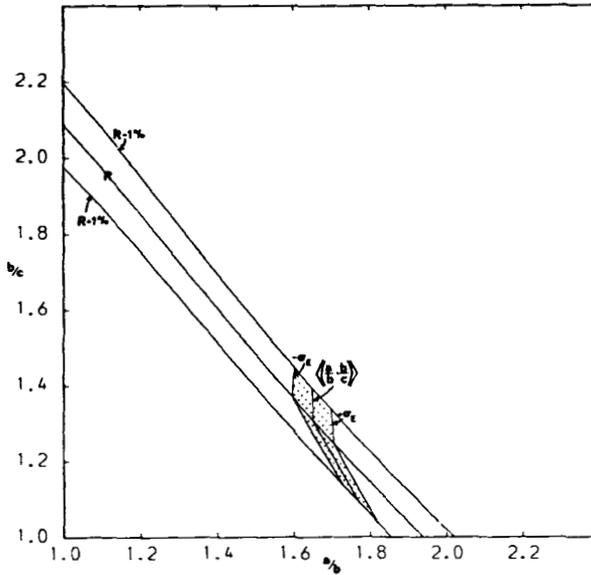


Fig. 6. Contour plots for a hypothetical Protein 2 (as in Fig. 5). [True $(a/b, b/c)$ = (1.7, 1.25).]
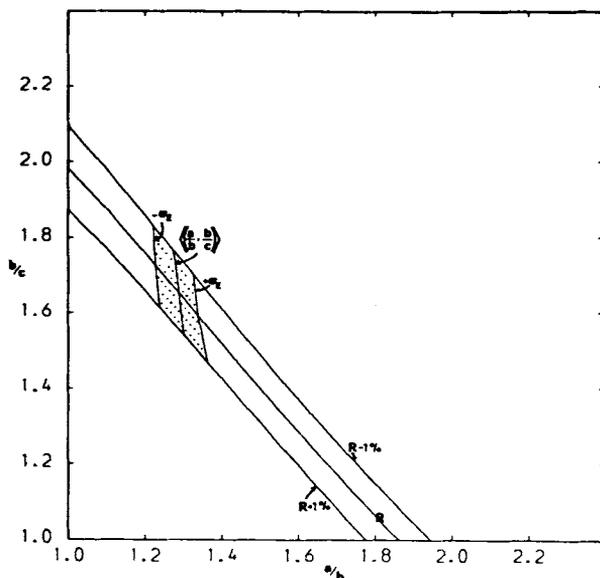
Fig. 7. Contour plots for a hypothetical Protein 3 (as in Fig. 5). [True $(a/b, b/c)$ = (1.25, 1.7).]

The algorithm was then tested for the effect of experimental errors in the intrinsic viscosity (±1%) and molecular weight (±1.4%) [Eq. (4)]. These were found to be not significant (Table IV). In any case, molecular weights of many macromolecules can now be found precisely from the results of

TABLE IV
Effect of Experimental Errors in the Intrinsic Viscosity and Molecular Weight[a]

| Stream No. | $a/b$ | | |
| of Random Data | −1.7% | No Error | +1.7% |
| --- | --- | --- | --- |
| 1 | 1.493 | 1.503 | 1.520 |
| 2 | 1.638 | 1.654 | 1.679 |
| 3 | 1.455 | 1.468 | 1.487 |
| 4 | 1.374 | 1.396 | 1.424 |
| 5 | 1.383 | 1.401 | 1.425 |
| 6 | 1.305 | 1.315 | 1.333 |
| 7 | 1.695 | 1.694 | 1.704 |
| 8 | 1.566 | 1.575 | 1.593 |
| Mean $a/b$ | 1.489 | 1.509 | 1.521 |
| $\sigma_D{}^b$ | 0.136 | 0.132 | 0.130 |
| $\sigma_E{}^c$ | 0.0482 | 0.0468 | 0.0467 |

[a] Assumed error in $[\eta]$ = ±1%; in $M_r$, ±1.4%. Total error in product $[\eta] M_r \simeq$ ±1.7% (calculated from the formula of Ref. 21). Results are for Protein 1: cutoff time, 115 ns, ±0.1° SE on each of the 100 data points.

[b] $\sigma_D$, standard deviation.

[c] $\sigma_E$, standard error.

sequence analyses. Finally, the algorithm was tested for the effect of taking different initially assumed values for the pre-exponential factors $A'_+$ and $A'_-$ (Table V). Again, these were found to have no significant effect on the results: even for pre-exponential factors differing by two orders of magnitude, although the retrieved $A'_-$ was poor, the retrieved $a/b$ was in close agreement with the other values.

Once the value for the axial ratio $(a/b, b/c)$ has been found for a particular protein, it can then be combined with the swollen volume of the protein to determine the axial dimensions. "Model-dependent" estimates for the swollen volumes, $V_e$, of each protein have been found by back-substitution of the mean value of $(a/b, b/c)$ determined above into Eq. (1) for the viscosity increment. The semi-axial dimensions $a, b, c$ for the three proteins are then found to be (Å):

Protein 1:  45.00, 29.98, 20.01   (45.0, 30.0, 20.0)
Protein 2:  42.28, 25.59, 19.61   (42.5, 25.0, 20.0)
Protein 3:  43.11, 33.58, 19.81   (42.5, 34.0, 20.0)

These are in excellent agreement with the actual values.


## DISCUSSION

In applying this algorithm to real macromolecular solutions, several important factors must be taken into consideration.

1. Two or more decay constants can also arise if the system is polydisperse. It is therefore essential that the solution be rendered homogeneous by, for example, gel filtration or other appropriate techniques.

2. It has now been well established[22] that the single-exponential decay constant previously resolvable from monodisperse decays shows a concentration dependence. One must therefore naturally assume this to be true for the two decay constants now resolvable, and hence they should be

TABLE V
Effect of Different Initial Values for the Pre-exponential Factors $A'_\pm$[a]

| Assumed | | Retrieved | | |
| --- | --- | --- | --- | --- |
| $A'_+$ | $A'_-$ | $a/b$ | $A'_+$ | $A'_-$ |
| 0.06 | 0.06 | 1.683 | 0.057 | 0.064 |
| 0.07 | 0.05 | 1.674 | 0.065 | 0.055 |
| 0.09 | 0.03 | 1.660 | 0.083 | 0.038 |
| 0.11 | 0.01 | 1.664 | 0.099 | 0.021 |
| 0.119 | 0.001 | 1.644 | 0.109 | 0.012 |

[a] Protein 1, cutoff time = 100 ns, 0.1 SE, on each of the 100 data points. The data for this table were obtained after the UK NAG Mk VI routines had been updated to Mk VII; the new random-number routines corresponding to G05ADF and G05BAF in Mk VI are G05CAF and G05CBF.

extrapolated to infinite dilution. On the other hand, because of the constraint in our algorithm that they must correspond to $\delta_+$ line solutions that intersect with the $R$ curve, the values for the decay constants are such that they are not the "true" decay constants for each particular concentration but closer to their infinite-dilution values. Since the extrapolation procedure must be empirical, the best estimates for $a/b$ at particular solute concentrations rather than these "damped" decay constants may be extrapolated to infinite dilution. Once the extrapolated value for $a/b$ has been found, the corresponding value for $b/c$ can thus be found from the $R$ curve.

3. The requirement on the precision of the electric birefringence apparatus is not only in producing transient decays to a precision of 0.1° SE on each data point, but also the availability of response times (i.e., the finite time it takes for the orienting pulse to be switched off) of about an order of magnitude less than that of the faster relaxation time. Adequate response times (Ref. 23; and B. Jennings and V. Morris, private communication), are now available, however, with apparatus that use a laser light source, cable discharge generator, and a memory oscilloscope, giving a response time of $\sim$5 ns.

4. In the above analysis it has been shown that greater accuracies in obtaining the axial ratios can be had if the optimum cutoff time for the decay is found; this corresponds in practice to recording several decays of the same preparation. Different samples of the same preparation should be used because of the risk of denaturing the protein by pulsing through high electric fields.

5. It has also been assumed that the $R$ function can be measured to a precision of $\sim \pm 1\%$. Sedimentation coefficient values in a $s_c$ versus concentration plot can be determined to within $\sim \pm 1\%$. The intrinsic viscosity $[\eta]$ can also be measured to within $\sim \pm 1\%$, the limiting factor here being the accuracy to which flow times can be measured. The error in $R$ will thus be of the order of 1% after taking into consideration that any systematic errors in measuring absolute solute concentrations will cancel in the ratio $k_s/[\eta]$.[3]

6. Finally, it should be pointed out that because of polarization effects on the electrodes and also the danger of high electric fields mentioned above, solutions of low ionic strength ($<0.01M$) generally have to be used. [This could account for the nonideality observed in Pt. (2).] On the other hand, an interesting new method is being developed in which an ultrasonic field rather than an electric field is used for initial orientation of the macromolecules before the decay is observed.[24] This "acoustic birefringence" method does not suffer from the problems of electrode polarization and denaturation associated with higher ionic strengths, allowing the possibility for the investigation of less soluble materials.

The experimental application of our newly developed algorithm, in which data from viscosity, sedimentation, and electric birefringence are combined to determine unique values for $(a/b, b/c)$, the two axial ratios of the tri-axial

ellipsoid, remains to be accomplished.   Detailed simulation shows that our
algorithm is stable to known experimental errors.   Problems as yet un-
suspected may arise, but equally, experimental technique may well advance,
especially in the area of electric birefringence.[25]

Squire and Himmel[26] have recently reemphasized the role that those
methods often loosely called "hydrodynamic" can play in investigating the
conformation of macromolecules in solution, including the presence of
associated solvent.   Indeed, Squire and Himmel's data suggest that the
ellipsoid of revolution remains a very adequate model for many proteins,
and from this we would predict that our algorithm would yield a value of
either $a/b$ or $b/c$ close to unity in such cases.   Equally, there exist well-
known proteins that cannot be modeled by any sort of ellipsoid: sometimes
for very obvious reasons of total shape discrepancy (e.g., myosin) and
sometimes for reasons that are as yet uncertain but that may become ap-
parent when the three-dimensional structure in solution is known or that
may arise from very particular types of solute–solvent interaction (e.g.,
serum albumin[5,27]).   Nonetheless, the number of macromolecular solutes
capable of being modeled by an ellipsoid, at least to a level where experi-
mental precision rather than the goodness of the model is limiting, is
probably large.   As probes of conformation and conformational changes
in such solutes "hydrodynamic" methods are nondestructive, use well-
defined methodology, and yield results that apply to the solute particles
in solution and that can illuminate the increasingly interesting area of so-
lute–solvent interaction.   There is, however, little point in continuing the
all-too-prevalent tradition of reporting "axial ratios" and "hydration val-
ues" where it is often painfully obvious that the (usually unspecified) real
imprecision in the estimate probably exceeds the actual numerical value
of the estimate itself.   Our present study shows many approaches that
might appear mathematically plausible to be hopelessly unstable from the
numerical standpoint.   Numerically stable procedures can, however, be
defined for ellipsoids of revolution and now for tri-axial ellipsoids.   Ex-
perimental results suggest that the simpler model can yield valid and in-
teresting results (Squire and Himmel[26]; and A. J. Rowe, data to be pub-
lished on the use of the $R$ function).   Further experimental data will be
required to test the real usefulness of our presently defined approach.

# References

1.  Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York.
2.  Jeffrey, P. D., Nichol, L. W., Turner, D. R. & Winzor, D. J. (1977) *J. Phys. Chem.* **81**, 776–778.
3.  Rowe, A. J. (1977) *Biopolymers*, **16**, 2595–2611.
4.  Squire, P. G. (1978) *Electro. Op. Ser.* **2**, 565–600.
5.  Harding, S. E. & Rowe, A. J. (1982) *Int. J. Biol. Macromol.* **4**, 160–164.
6.  Harding, S. E., Dampier, M. & Rowe, A. J. (1979) *IRCS Med. Sci.* **7**, 33.
7.  Harding, S. E., Dampier, M. & Rowe, A. J. (1981) *J. Colloid Interface Sci.* **79**, 7–13.
8.  Simha, R. (1940) *J. Phys. Chem.* **44**, 25–34.

9. Saito, N. (1951) *J. Phys. Radium* **7**, 1–11.

10. Jeffrey, G. B. (1922) *Proc. R. Soc. London, Ser. A* **102**, 161–179.

11. Harding, S. E. (1980) Ph.D. thesis, University of Leicester.

12. 'NAG' FORTRAN Library Manual (1978) Numerical Algorithms Group, Oxford.

13. Perrin, E. (1936) *J. Phys. Radium* **7**, 1–11.

14. Scheraga, H. A. & Mandelkern, L. (1953) *J. Am. Chem. Soc.* **79**, 179–184.

15. Oncley, J. L. (1941) *Ann. NY Acad. Sci.* **41**, 121–150.

16. Jost, J. W. & O'Konski, C. T. (1978) *Electro. Opt. Ser.* **2**, 529–564.

17. Small, E. W. & Isenberg, I. (1977) *Biopolymers* **16**, 1907–1928.

18. Ridgeway, D. (1968) *J. Am. Chem. Soc.* **90**, 18–22.

19. O'Connor, D. V., Ware, W. R. & Andre, J. C. (1979) *J. Phys. Chem.* **83**, 1333.

20. Gill, P. E. & Murray, W. (1976) Natl. Phys. Lab. Rep. NAC 72.

21. Paradine, C. G. & Rivett, B. H. P. (1960) *Statistical Methods for Technologists*, English Universities Press, London.

22. Riddiford, C. L. & Jennings, B. (1967) *Biopolymers* **5**, 757–771.

23. Williams, R. C., Ham, W. T. & Wright, A. K. (1976) *Anal. Biochem.* **73**, 52–64.

24. Ballinger, K. W. A. & Jennings, B. (1979) *Nature* **282**, 699–701.

25. Krause, S., Ed. (1981) *Electro. Op. Ser.* **3**.

26. Squire, P. G. & Himmel, M. (1979) *Arch. Biochem. Biophys.* **196**, 165–177.

27. McCammon, J. A., Deutch, J. M. & Bloomfield, V. A. (1975) *Biopolymers* **14**, 2479–2484.