

The ELLIPS Suite of Whole-Body Protein Conformation Algorithms for Microsoft WINDOWS

STEPHEN E. HARDING, HELMUT CÖLFEN AND ZAHID AZIZ

1 Introduction: ELLIPS1, 2, 3, 4

ELLIPS1, 2, 3 and 4 are simple to use algorithms for the representation of the overall hydrodynamic shape of proteins in solution in terms of tri-axial ellipsoids (three semi-axes $a > b > c$ and shape characterised by two axial ratios a/b , b/c) and bi-axial ellipsoids or “ellipsoids of revolution” where two of the semi-axes are approximated as equal: a prolate ellipsoid has semi-axes a , b , b an oblate ellipsoid has a , a , b , with $a > b$ in both cases and the asymmetry defined by the axial ratio a/b . These algorithms have previously been available only from mainframe or MSDOS platforms.¹ This short chapter briefly reviews the purpose of these FORTRAN/QUICK-BASIC algorithms and describes their fresh implementation onto a Microsoft WINDOWS platform.

Hydrodynamic methods provide a useful approach to the study of macromolecular conformation in solution. In the study of the conformation and flexibility of linear types of macromolecules – for example synthetic polymers and polysaccharides, consideration of how hydrodynamic parameters such as the intrinsic viscosity, the sedimentation coefficient or the radius of gyration vary with molecular weight of a homologous polymer series has provided the means of estimating particle dimensions and flexibility (via the persistence length) in solution.

For the representation of quasi-rigid types of macromolecule – many proteins for example (except at very short timescales) can be thought of as such – hydrodynamic approaches have also proved useful.

There are two approaches to representing the conformation of fairly rigid proteins in solution. The first approach is the bead modelling whereby a macromolecule or macromolecular assembly is approximated as an array of spherical beads. Using computer programs that are currently available (based on how these spheres interact) such as HYDRO, SOLPRO, HYDROPRO, HYDROSUB,²⁻⁴ it is possible for a given Bead

model to predict its hydrodynamic properties. One can model quite sophisticated structures by this approach, and recent efforts have focussed on its application to flexible structures.⁵ There are uniqueness problems, for example, one can predict the sedimentation coefficient for a particular complicated model, but there will be many other equally complicated models which give the same sedimentation coefficient. This type of modelling is therefore best for choosing between plausible models for a structure, or for refining a close starting estimate for a structure from, say, X-ray crystallography.

A complementary approach to bead modelling is to make no assumptions concerning starting estimates and to calculate the shape directly from hydrodynamic measurements. This is called the "ellipsoid" or "whole body" approach,⁶ so called because the investigator instead of approximating the macromolecule as an array of spheres approximates the macromolecule instead as a smooth whole regular structure – an ellipsoid, or "three-dimensional ellipse" characterised by three perpendicular semi-axes $a \geq b \geq c$. Of course, only simple representations are possible but by combining shape parameters together there are no hydration or uniqueness problems. This approach is best for giving a relatively quick idea of the overall dimensions or shape of a macromolecule in solution. There are two types of ellipsoid approach: the *ellipsoid of revolution* and the *general triaxial ellipsoid*. The simplest of these (which has been used in one form or another for over half a century⁷) is the ellipsoid of revolution in which two of the three semi-axes are equal ($c=b$). Ellipsoids of revolution are so-called because they are the shapes formed by rotating an ellipse of semi-axes, a, b either about the major (a) axis to give a prolate ellipsoid (semi-axes a, b, b) or about the minor (b) axis to give an oblate ellipsoid (a, a, b), both defined by the axial ratio (a/b) (where $a \geq b$). One hydrodynamic measurement can uniquely define (a/b), after assuming a value for the (time averaged) hydration of the molecule; two hydrodynamic measurements are normally sufficient to define (a/b) without assumptions concerning hydration; a third is occasionally necessary to distinguish whether an oblate ellipsoid or prolate ellipsoid is the more appropriate (usually the latter for proteins). In the extremes $a \gg b$ the prolate \rightarrow rod and the oblate \rightarrow disc and the other extreme of $a=b$ is of course a sphere.

The most sophisticated of the whole body approaches is the general tri-axial ellipsoid where the restriction of two equal axes $b=c$ is removed. This allows a much greater variety of conformations ranging from rods ($a \gg b=c$), discs ($a=b \gg c$) and tapes ($a \gg b \gg c$) as well as the prolate ($a > b=c$), oblate ($a=b > c$) ellipsoids of revolution and the sphere ($a=b=c$). All the necessary theoretical developments for applying either of these "ellipsoid" strategies are in place. What had been lacking is a coherent set of easy-to-use algorithms – available on PC as opposed to computer mainframes – which the general user has access to. The launch of a suite of four ELLIPS algorithms in 1997¹ for an MSDOS platform helped addressing this: Table 1 gives a summary of what these do. ELLIPS1 is written in QUICKBASIC, the others are in FORTRAN. Obviously for some classes of molecule – antibodies are a good example – this type of whole-body modelling is not applicable and bead- approaches need to be employed. Even here, however, ellipsoidal representations of the major domains (Fab, Fc) have helped in the bead modelling of the intact assembly.⁸⁻¹³

Table 1 *The ELLIPS routines*

<i>Routine</i>	<i>Language</i>	<i>Model</i>	<i>Purpose</i>
ELLIPS1	QUICKBASIC	Ellipsoid of revolution	Prediction of axial ratio (a/b) (equivalent prolate or oblate ellipsoid of revolution) from user-specified shape function
ELLIPS2	FORTRAN	General triaxial ellipsoid	Evaluates the values of all the hydrodynamic shape functions from user-specified (a, b, c) or ($a/b, b/c$) ^a
ELLIPS3	FORTRAN	General triaxial ellipsoid	Evaluates ($a/b, b/c$) from combinations of hydration-independent shape functions
ELLIPS4	FORTRAN	General triaxial ellipsoid	Evaluates ($a/b, b/c$) from electro-optic decay combined with other hydrodynamic data

^aEquivalent to SOLPRO³ for bead models.

2 Universal Shape Functions: Hydration-Dependent and Hydration-Independent

In common with the bead modelling program SOLPRO³ the ELLIPS algorithms all use *Universal shape functions*. By this we mean each is specifically a function of shape alone (and not volume). It makes no odds what the size is: a Universal shape function will have the same value, and will only depend on the shape. All these universal shape functions have been worked out in terms of the axial ratio (a/b) for ellipsoids of revolution and now the two axial ratios ($a/b, b/c$) for general ellipsoids. The relations of all these to (a/b) or ($a/b, b/c$) are given in Harding¹⁴ and will not be repeated here: all of these *exact* formulae are inbuilt into the ELLIPS routines.

To measure these Universal shape functions experimentally, many require knowledge of the hydration δ (mass in g of H₂O bound per g of dry macromolecule) or hydrated volume V (mL) of the particle, the others do not. Hydration is a dynamic process, and so δ and V represent time-averaged values. The particle volume V is often presented in two equivalent forms:

$$V = v_s M/N_A \quad (1)$$

where M is the molecular weight or molar mass (g mol⁻¹), N_A is Avogadro's number (6.02205×10^{23} mol⁻¹), and v_s is the specific volume (mL g⁻¹) of the hydrated macromolecule (volume occupied by the hydrated macromolecule per unit mass of dry macromolecule) or

$$V = (\bar{v} + \delta/\rho_o)M/N_A \quad (2)$$

where \bar{v} is the partial specific volume (mL g⁻¹).

3 Hydration-Dependent Universal Shape Functions

Harding *et al.*¹ give a complete list of those Universal shape functions requiring knowledge of δ or V for their experimental measurement. We give here only the most useful ones:

- *Viscosity increment*:^{15,16}

$$\nu = [\eta]M/(N_A V) \quad (3)$$

in which $\nu = 2.5$ for a sphere.^{17,18}

- *Perrin*¹⁹ *function*:

$$P = (f/f_0)\{1 + \delta/(\bar{\nu}\rho_0)\}^{-1/3} \quad (4)$$

where (f/f_0) , the frictional ratio,⁷ is related to the sedimentation coefficient $s_{20,w}^0$ by

$$(f/f_0) = M(1 - \bar{\nu}\rho_0)/(N_A 6\pi\eta_0 s_{20,w}^0)(4\pi N_A/3\bar{\nu}M)^{1/3} \quad (5)$$

or the translational diffusion coefficient $D_{20,w}^0$ by

$$(f/f_0) = \frac{k_B T}{6\pi\eta_0} \left(\frac{4\pi N_A}{3\bar{\nu}M} \right)^{1/3} \frac{1}{D_{20,w}^0} \quad (6)$$

where $T=293.15$ K, η_0 is the viscosity of water at 293.15 K (0.010 P), ρ_0 is the density of water at 293.15 K (0.99823 g mL⁻¹) and k_B is the Boltzmann's constant (1.3807×10^{-16} erg K⁻¹). $P=1$ for a sphere.¹⁹

- *Reduced excluded volume*:²⁰

$$u_{\text{red}} = u/V = \{2BM^2 - Z^2/2I\}/(N_A V) \quad (7)$$

where u is the excluded volume (mL), B the second thermodynamic (or "osmotic pressure") virial coefficient (mL mol g⁻²) from osmotic pressure, light scattering or sedimentation equilibrium measurements, Z is the valency of the macromolecule, measurable by titration²¹ and I is the ionic strength of electrolyte in the solvent (mol/mL). At sufficient ionic strengths, the $Z^2/2I$ term becomes negligible compared with $2BM^2$. Of course, for uncharged macromolecules and proteins at the isoelectric point $Z=0$. $u_{\text{red}}=8$ for a sphere.⁷

- *Harmonic mean rotation relaxation time ratio*:

$$\tau_h/\tau_0 = \{k_B T/\eta_0 V\} \tau_h \quad (8)$$

where τ_h (s) is the harmonic mean rotational relaxation time, traditionally measured using steady-state fluorescence depolarisation methods,^{22,23} and τ_0 the corresponding value for a spherical particle of the same volume:

$$\tau_0 = \eta_0 V/k_B T \quad (9)$$

In earlier representations a factor of 3 was introduced because the rotational relaxation time was referred to on a dielectric dispersion basis (compensated

for in the equations for steady-state anisotropy depolarisation) although this is no longer necessary – compare ref. 22 with ref. 23. This is further discussed in Garcia de la Torre *et al.*³ $\tau_h/\tau_o = 1$ for a sphere.²⁴

- *Reduced electro-optic decay constants:*

$$\theta_i^{\text{red}} = (\eta_o V/k_B T)\theta_i \quad (10)$$

where θ^i are the electric birefringence or electric dichroism decay constants. For ellipsoids of revolution that are homogeneous, *i.e.* where the geometric axis of symmetry coincides with the electrical axis, $i=1$. For general ellipsoids that are homogeneous, *i.e.* where the geometric axes coincide with the electrical axes, $i=2$, termed “+” and “–” (refs. 25 and 26); for general particles $i=1-5$ (ref. 27). For a sphere, $\theta_i^{\text{red}}=0.66667$.

To assist with the calculation of the salient Universal parameters P , ν and u_{red} from the sedimentation coefficient, intrinsic viscosity and second virial coefficient, respectively, a spreadsheet algorithm has been set up called ELLIPSPRIME, which can be downloaded along with the other algorithms described here from the NCMH web site <http://www.nottingham.ac.uk/ncmh>.

4 Hydration-Independent Universal Shape Functions

Harding *et al.*¹ also give a complete list of those Universal shape functions *NOT* requiring knowledge of δ or V for their experimental measurement. Again we give here only the most popular or useful ones:

- *The Scheraga–Mandelkern*²⁸ *parameter*

$$\beta = \{[\eta]^{1/3}\eta_o\}/\{M^{2/3}(1 - \bar{\nu}\rho_o)100^{1/3}\} = \{N_A^{1/3}/(16200\pi^2)^{1/3}\}\{\bar{\nu}^{1/3}/P\} \quad (11)$$

The β parameter is unfortunately very insensitive to shape, and Equation (11) is used more as an equation of consistency, or for measuring M from sedimentation velocity and viscosity measurements. $\beta=2.1115\times 10^6$ for a sphere.

- *The Pi function*²⁹

$$\Pi = \{2BM/[\eta]\} - \{Z^2/2IM[\eta]\} = u_{\text{red}}/\nu \quad (12)$$

with the second term in the parantheses (an approximation of the charge contribution for polyelectrolytes) $\rightarrow 0$ at sufficient values of I , and of course $=0$ for uncharged macromolecules or proteins at the isoelectric point ($Z=0$). $\Pi=3.2$ for a sphere

- *The Wales–van Holde*^{30,31} *parameter*

$$R = k_s/[\eta] = 2(1 + P^3)/\nu \quad (13)$$

where k_s (mL g^{-1}) is the concentration dependence parameter of the sedimentation coefficient in the limiting relation

$$s_{20,w} = s_{20,w}^o(1 - k_s c)$$

or

$$1/s_{20,w} = \{1/s_{20,w}^0\}(1 + k_s c)$$

Although the theory behind Equation (13) is less rigorous than that for Π (because of the greater complexity of “hydrodynamic” as opposed to “thermodynamic equilibrium”-based non-ideality), it does have a strong experimental basis.³¹⁻³³ To apply k_s in this way it is important that charge contributions to k_s are absent or if the macromolecule is a polyelectrolyte, charge contributions are suppressed by working with a solvent of sufficient ionic strength. $R=1.6$ for a sphere.

- The reduced radius of gyration function G (ref. 34)

$$G = R_g^2 \{4\pi N_A / (3vM)\}^{2/3} \quad (14)$$

where R_g is the radius of gyration (cm), determined from light scattering, X-ray scattering or neutron scattering measurements. v is the specific volume of the macromolecule. Originally it was thought that this was closest to the specific volume of the anhydrous macromolecule, *i.e.* there is no difference in scattering density of the surface-bound solvent compared with free solvent although current thinking seems to be in favour of something between the anhydrous and (time-averaged) hydrated macromolecule: strictly speaking G should be regarded as a hydration-dependent parameter. $G=0.6$ for a sphere.

- The Lambda function³⁵

$$\Lambda = (\eta_o[\eta]M) / (N_A k_B T \tau_h) = v / (\tau_h / \tau_o) \quad (15)$$

For spheres, $\Lambda=2.5$.

- Electro-optic delta functions³⁶

$$\delta_i = (6\eta_o / N_A k_B T) [\eta] M \theta_i = 6\theta_i^{\text{red}} v \quad (16)$$

(for homogeneous ellipsoids of revolution $i=1$ and for homogeneous triaxial ellipsoids, $i=+$ and $-$). For spheres $\delta_i=2.5$.

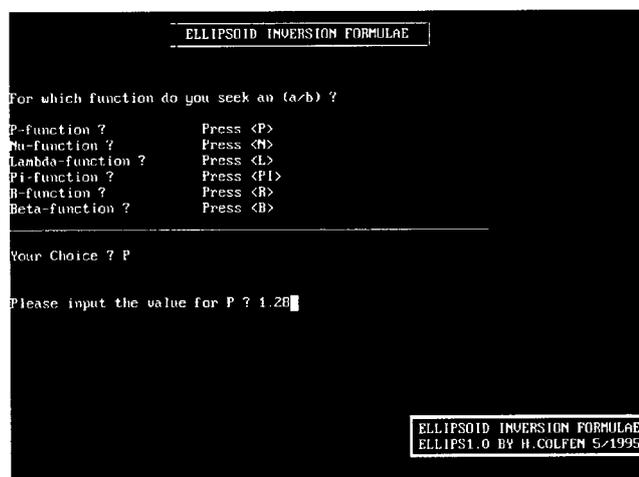
5 ELLIPS1

Aim: Prediction of axial ratio (a/b) (equivalent prolate or oblate ellipsoid of revolution) from a user specified value for a shape function.

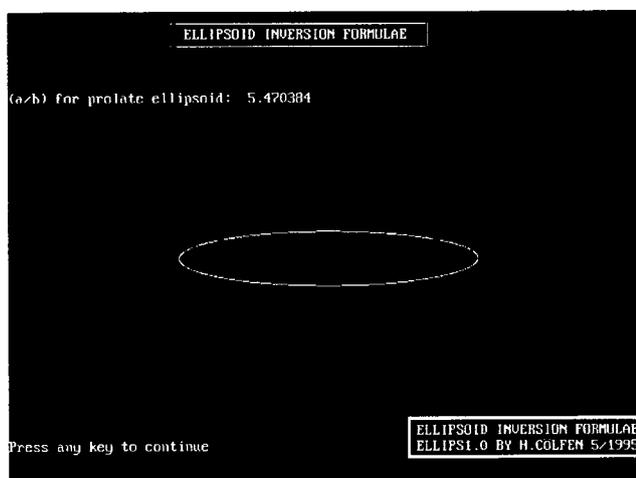
Description: ELLIPS1 is based on simple ellipsoid of revolution models (where two of the three axes of the ellipsoid are fixed equal to each other); if the user types in a value for a shape function from sedimentation or other types of hydrodynamic measurement, it will return a value for the axial ratio of the ellipsoid. The question an experimenter wishes to address usually is not “what is the shape function for a specified value of the axial ratio a/b ?” but rather “what is the axial ratio a/b for my macromolecule specified by my (Universal) shape function which I have experimentally measured?”. Although there are exact analytical formulae linking each shape function with a/b (ref. 37), the reverse is not true: inversion is analytically impossible. The QUICKBASIC algorithm ELLIPS1 uses the polynomial-based

inversion procedure of Harding and Cölfen³⁷ to give a/b vs. the various Universal shape functions to an acceptable degree of accuracy (*i.e.* to better than the precision of the measurement, which is normally no better than a few per cent).

Example: Figure 1 and Table 2 show an example for part of the complement receptor CR1 (modules 16/17). In a study by Kirkitadze and co-workers³⁸ they showed that the axial ratio of this protein from ELLIPS1 was ~ 5 from P values obtained from



(a)



(b)

Figure 1 ELLIPS1 screens for the determination of the axial ratio (a/b) for complement receptor CR1 (domains 16/17) using the Universal shape function P obtained from the sedimentation coefficient, $s_{20,w}^0 = (1.48 \pm 0.04)$ S, molecular weight, $M = 14513$ g mol⁻¹ and partial specific volume $\bar{v} = 0.725$ mL g⁻¹. (a) Input screen. (b) Output screen (result for prolate ellipsoid only shown). (Data from ref. 38)

Table 2 ELLIPS1 estimates of the axial ratio a/b for the complement receptor CR1 (domains 16/17) using the Universal shape function P for three different values of the time-averaged hydration parameter δ . (Adapted from ref. 38).

δ	P	a/b
0.30	1.28	5.5
0.35	1.26	5.2
0.4	1.23	4.6

measurement of $s_{20,w}^0$ and various possible values of the hydration parameter δ . Since the individual domains (16 and 17) each had an axial ratio ~ 3 they could conclude the domains were arranged approximately in an end-to-end fashion.

6 ELLIPS2

Aim: Evaluates the values of *all* the Universal hydrodynamic shape functions from user-specified axial dimensions (a, b, c) or axial ratios ($a/b, b/c$) for the macromolecule as modelled by a general triaxial ellipsoid.

Description: ELLIPS2 is essentially analogous to SOLPRO³ in that from a given structure (as represented by an array of beads in SOLPRO or as a general triaxial ellipsoid in ELLIPS2) the *complete set* of Universal shape functions is returned. Most of the universal shape functions involve one or more of 10 different elliptic integrals (called alpha 1...alpha 10 – see Harding¹⁴). These are solved by quadrature using the NAG³⁹ routine D01AMF for a one-dimensional integration with an infinite upper limit and routine D01DAF for a two-dimensional integration with finite limits.

Example: Figure 2 gives the output data for an $(a/b, b/c) = (1.23, 1.52)$, based on the crystallographic axial dimensions of $43 \times 35 \times 23 \text{ \AA}$ for Fab' of a chimaeric antibody B72.3.^{40,41} These were then used as the basis for the construction of a surface shell bead model. This procedure was repeated for Fc and a model for the intact immunologically active antibody was then constructed.^{9,10}

7 ELLIPS3

Aim: Performs the reverse of ELLIPS2 by evaluating the tri-axial shape of a macromolecule ($a/b, b/c$) using two possible combinations of universal shape functions.

Description: Whereas an $(a/b, b/c)$ specifies unique values for all the hydrodynamic shape functions, the reverse is unfortunately not true: measurement of P, v, R, Λ, \dots does not uniquely fix $(a/b, b/c)$ but rather gives a line solution of possible values. A graphical combination of the line solutions for two of these shape functions will in principle provide a unique solution for $(a/b, b/c)$. The main criteria for selection are (i) their ease of measurement, (ii) their sensitivity to shape and insensitivity to experimental error, (iii) the two give an intersection as orthogonal as possible and (iv) the lack of requirement of an estimate for the hydration for their experimental measurement. ELLIPS3 currently offers two such combinations: the

```

*****
                      ELLIPS2 Output Data
*****

(a/b, b/c) =      1.6000    1.4200
Viscosity increment, nu =      2.9120
Perrin function, P =      1.0458
Corresponding hydration independent functions:
Wales-van Holde, R =      1.4724
Scheraga-Mandelkern, 10**-6xbeta =      2.1245
Functions based on the 2nd virial coefficient:
Reduced excluded volume, u_red      9.1442
Corresponding hydration independent function:
Pi function =      3.1401
G function (from radius of gyration) =      0.7491
Reduced electro-optic decay constants:
  Theta+ =      0.1654
  Theta- =      0.1134
Corresponding hydration independent functions:
Delta+ =      2.8907
Delta- =      1.9806
Gamma+ =      1.6965
Gamma- =      1.1624
Harmonic mean rotational relaxational time ratio:
tau_h/tau_0 =      1.1956
Corresponding hydration-independent functions:
PSI (Squire-Himmel) function =      0.9854
LAMBDA =      2.4356
Fluorescence anisotropy relaxation time ratios:
tau_1/tau_0 =      1.0075
tau_2/tau_0 =      1.3083
tau_3/tau_0 =      1.3292
tau_4/tau_0 =      1.4703
tau_5/tau_0 =      1.0074
Corresponding hydration independent functions:
lambda_1 =      2.8903
lambda_2 =      2.2259
lambda_3 =      2.1908
lambda_4 =      1.9806
lambda_5 =      2.8907
psi_1 =      1.0432
psi_2 =      0.9562
psi_3 =      0.9512
psi_4 =      0.9197
psi_5 =      1.0433
*****

```

Figure 2 ELLIPS2 output for B72.3c Fab. The axial ratios of (a/b, b/c)=1.60, 1.42) used in this example were obtained from the crystal structure of Brady et al.⁴⁰ using the algorithm of Taylor et al.⁴¹

A function combined with the R function and the Π function combined with the G function. The former combination satisfies all four criteria: A requires the experimental measurement of the harmonic mean rotational relaxation time (from e.g. steady-state fluorescence measurements: no complicated resolution of exponentials is required), with the intrinsic viscosity. With regard to the latter parameter, the

traditional U-tube viscometers require relatively large quantities of material – the new-generation pressure imbalance methods now make this attractive for those materials available in low quantities.^{13,42,43} The Wales–van Holde parameter R is measured from the ratio of the concentration-dependent sedimentation term k_s to the intrinsic viscosity measurements. The latter combination also involves a hydration-independent function, namely Π (from measurement of the intrinsic viscosity and the second-thermodynamic virial coefficient B). G can also be measured without experimental measurement of hydration provided that the radius of gyration R_g is measured using neutron scattering. If X-ray or light scattering is used, evidence now suggests the v in Equation (14) is the hydrated volume. ELLIPS3 uses as its basis the function calculation routine of ELLIPS2 except that a whole array of such values are evaluated in the $(a/b, b/c)$ plane (a matrix of 40×40 values). A Contour plotting routine (RGCNTS from the Simpleplot Library) interpolates between these matrix points and can plot the Π , G , A and R functions (or any other of the universal shape functions if the programmer so decides) in the $(a/b, b/c)$ plane.

Example: Figure 3 shows an example of the determination of the triaxial shape of neurophysin monomers and dimers in solution based on data of Nicolas *et al.*⁴⁴ and Harding and Rowe.⁴⁵

8 ELLIPS4

Aim: Evaluates the tri-axial shape of a macromolecule $(a/b, b/c)$ from electro-optic decay based Universal shape functions combined with other hydrodynamic data.

Description: Rotational hydrodynamic shape functions, based on rotational diffusion measurements, are attractive for determining the shapes of macromolecules in solution since they are generally more sensitive functions of shape compared to other shape functions. This sensitivity comes however at a price because they are generally more difficult to measure, due to problems surrounding the resolution of multiexponential decay functions. Electro-optic measurements are more attractive than time-resolved fluorescence depolarisation anisotropy measurements in the sense that for homogeneous triaxial ellipsoids at least, there are only two exponential terms to resolve (the decay constants or reciprocal relaxation times θ_+ and θ_-) as opposed to five (τ_1 – τ_5):

$$\Delta n = A'_+ \exp(-6\theta_+ t) + A'_- \exp(-6\theta_- t) \quad (17)$$

(see refs. 26 and 36) where Δn is the birefringence or dichroism (often expressed as “optical retardation” in degrees) at time t after the aligning electric field has been switched off. A practical problem with electro-optic decay methods is the potential local heating effects from the high electric fields used, especially if the experiments are conducted in solutions of high ionic strength: the investigator is advised to consult an article by Pörschke and Obst⁴⁶ describing how these effects can be minimised. After eliminating hydration (via *e.g.* combination with $[\eta]$) to give the Universal hydration-independent shape functions δ_+ and δ_- and graphical combination with another Universal hydration-independent shape functions such as R^{36} or Π ,⁴⁷ the triaxial shape as represented by the two axial ratios $(a/b, b/c)$ can be evaluated.

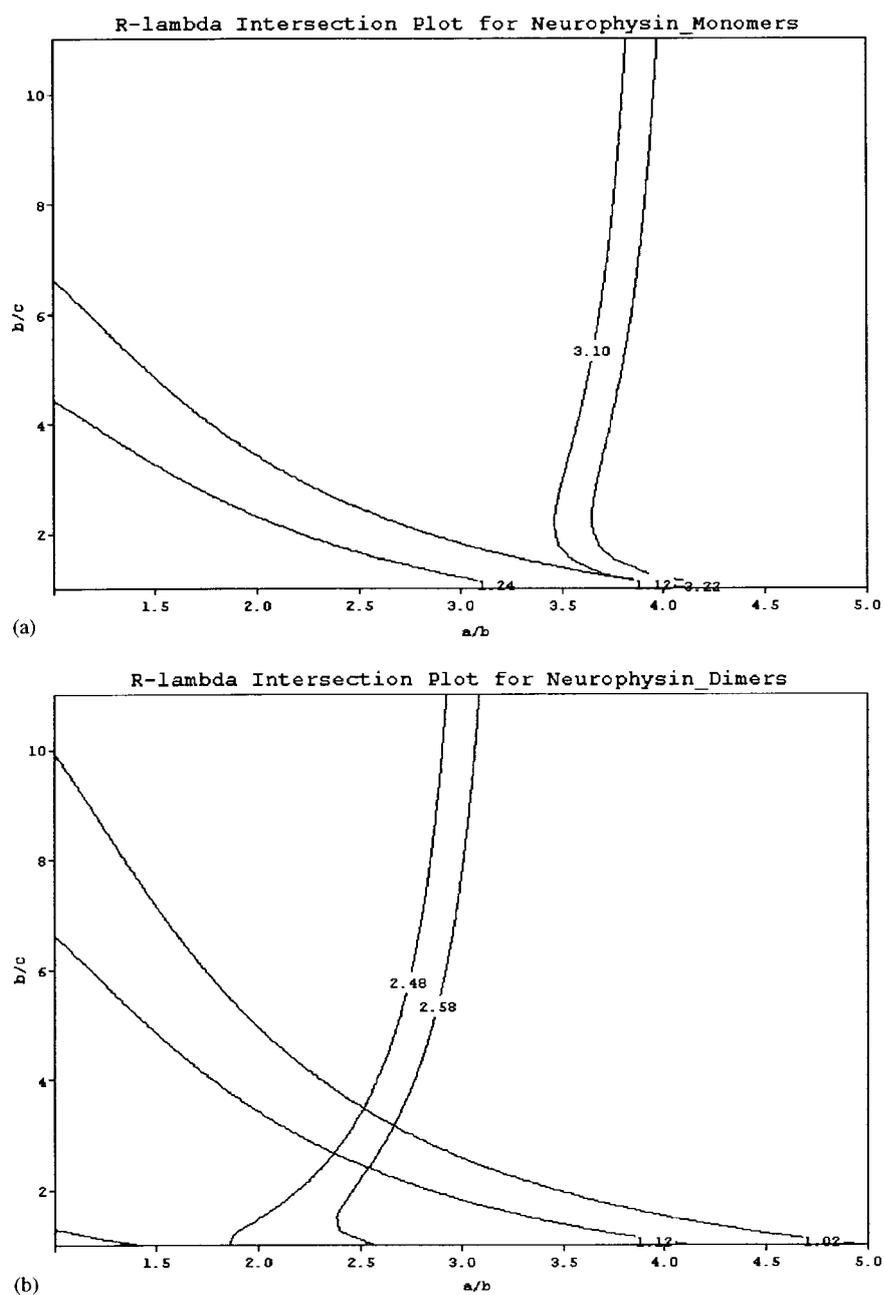


Figure 3 ELLIPS3 output. Λ - R plot applied to neurophysin (a) monomers, $\Lambda=3.16$, $R=1.18$ (b) dimers, $\Lambda=2.53$, $R=1.07$. The lines allow for experimental error of $\pm 2\%$ in Λ and $\pm 5\%$ in R . The intersections indicate an $(a/b, b/c)=(4.0, 1.0)$ for neurophysin monomer (i.e. a prolate ellipsoid of axial ratio 4.0) and an $(a/b, b/c)=(2.5, 2.9)$ for neurophysin dimers, suggesting the dimerisation is a side by side as opposed to end-to-end process

Resolution however of even two exponential terms is not easy, particularly for globular macromolecules where θ_+ and θ_- are similar,⁴⁸ irrespective of the form of mathematical deconvolution applied, whether it be non-linear least squares or more refined types of analysis;⁴⁹⁻⁵¹ see Johnsen and Brown⁵² for the analogous problem in dynamic light scattering analysis of polydisperse systems. In our hands^{36,49} we have found a more reliable method of extraction is to use another hydrodynamic function as a constraining parameter in the analysis of the electro-optic decay data: in this way the problem is reduced from one of four variables (A'_+ , θ_+ , A'_- , θ_-) to one of three (A'_+ , A'_- , a/b) since a/b will specify, by the constraining function a unique value for b/c (and hence θ_+ , θ_-). ELLIPS4 had been written to facilitate this procedure for PC¹ based on an earlier non-interactive version of the program written for mainframe computer^{49,53} and is now in WINDOWS. Its use is best illustrated by application to synthetic data (with error) generated for a macromolecule "Bonington," which includes the following characteristics: $(a/b, b/c)=(1.5, 1.5)$; $M=71744$ Da; $[\eta]=2.74$ mL g⁻¹, and the following electro-optic decay parameters: $A'_+=0.07$, $A'_-=0.05$, $\theta_+=5.81538 \times 10^6$ s⁻¹, $\theta_-=4.15646 \times 10^6$ s⁻¹, $T=293.15$ K, $\eta_0=0.01$ P. Figure 4(a) shows the electro-optic decay for this based on expected error (standard deviation) of $\pm 0.1^\circ$ (optical retardation) or ± 0.0017 rad random normal error on the decay data. With ELLIPS4 the user puts his electro-optic decay data (Δn vs. t) into a data file which is read in. The user also has to specify values for $[\eta]$ (mL g⁻¹), the molecular weight M (Da), the solvent viscosity (Poise) and temperature (K) at which the electro-optic measurements were made. The user also needs to specify the coordinates of a line of $(a/b, b/c)$ values (based on measurement of R , Π or some other hydration-independent Universal shape function) in a second data file: Figure 4(b) shows such a constraining line of allowed $(a/b, b/c)$ values for "Bonington" which has an R function value of 1.479. This constrains each iteration of $(a/b, b/c)$ and hence θ_+ , θ_- , to work along the line specified by the constraining function, since each value of $(a/b, b/c)$ specifies a value for δ_+ and δ_- (worked out using the the NAG routine D01AJF) which, combined with the user entered values for T , η_0 , $[\eta]$ and M gives the θ_+ , θ_- for each iteration. This reduces the risk of the fitting routine falling into subsidiary minima. When the minimum of the least-squares procedure has been formed ELLIPS4 successfully returns a/b , its corresponding value of b/c and the preexponential factors A'_+ and A'_- . The program runs automatically four times using successively the four different values of a/b entered from the constraining function (excluding the first and last data points) as starting estimates for a/b : this provides a further check against the dangers of subsidiary minima. For the starting estimates for A'_+ and A'_- the routine automatically takes these as $\Delta n_{\max}/2$. The routine has various inbuilt error warnings concerning the reliability of each estimation. If no error warning is returned the result for the evaluation from a particular starting point should be reliable.

Example: Figure 4(c) shows the output for a run on the data of Figure 4(a) for Bonington, which returns a value for $(a/b, b/c) \sim (1.62, 1.34)$ - *i.e.* to within two-tenths of an axial ratio unit of the true axial ratios. In practical terms, however, the user is advised to (i) repeat the whole operation several times with various cut-off times for the decay data (at longer times the signal/noise data gets progressively worse; on the other hand more information concerning the slower relaxation time, or larger decay constant, θ_+ is contained in this region) to be certain of no subsidiary-minima

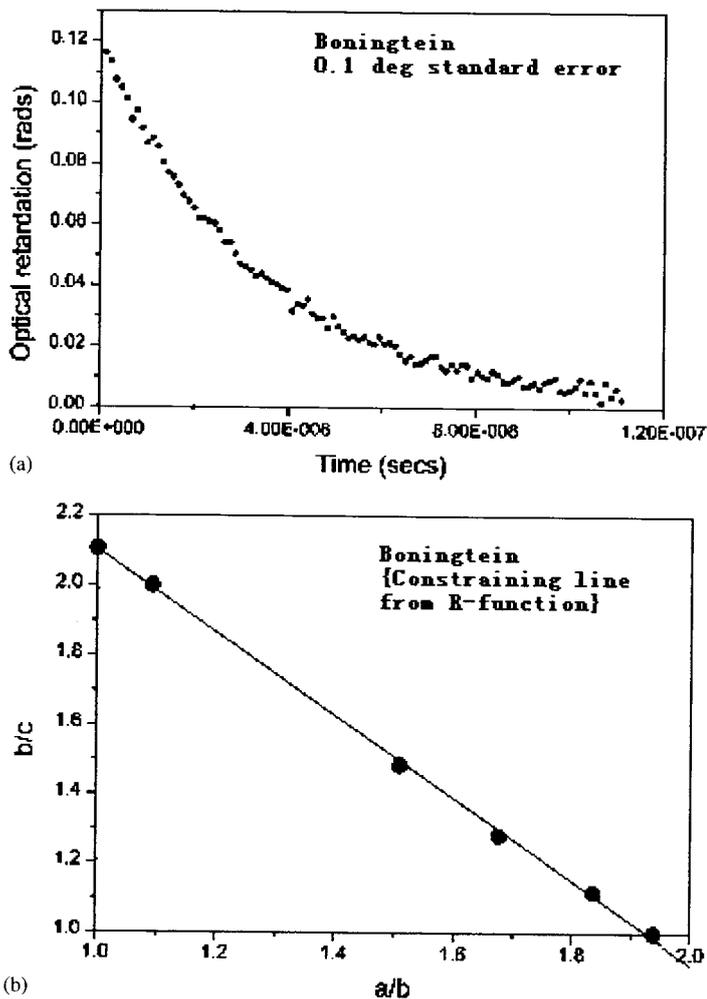


Figure 4 ELLIPS4. (a) Electro-optic decay (expressed as the decay of optical retardation with time, t (s)). Synthetic data shown corresponding to a protein of true $(a/b, b/c) = (1.5, 1.5)$ ("Boningtein"). The electro-optic data are fed in as a data-file into ELLIPS4. (b) R-constraining data for assisting with the resolution of the exponential terms. The user takes his experimental value of R or other suitable constraining function ($\Pi, G, \Lambda \dots$) and plots the line of corresponding values of $(a/b, b/c)$ using ELLIPS3 to do this for him. The user then reads off six $(a/b, b/c)$ coordinates from this line, which can be either entered at run-time or as a data file into ELLIPS4. (c) Output giving the values ELLIPS4 returns for $(a/b, b/c)$ and the preexponential factors for each of the four starting estimates for $(a/b, b/c)$ and the final "best" result. A list of potential error warning estimates is also given

problems; (ii) repeat the operation allowing for experimental error in the constraining function; (iii) check for any concentration dependence of the returned parameters: and extrapolate if necessary to zero concentration.⁵⁴ These and other features have been

```

*****
          ELLIPS4 Output Data
*****

Protein: Bonington (True a/b, b/c) = (1.5, 1.5)

Constraining co-ordinates:
  a/b      b/c
( 1.000,   2.106)
( 1.092,   2.000)
( 1.509,   1.483)
( 1.676,   1.280)
( 1.836,   1.116)
( 1.937,   1.000)

Temperature 293.000 K
Solvent viscosity 0.01000 Poise
Intrinsic viscosity 2.700 ml/g
Molecular weight 0.72000E+05 Da

From a starting estimate of a/b = 1.09200
Best least squares value = 0.000258733747
a/b = 1.62197 A+ = 0.058319856203 A- = 0.061743342703

From a starting estimate of a/b = 1.50900
Best least squares value = 0.000258733747
a/b = 1.62199 A+ = 0.058319708277 A- = 0.061743502486

From a starting estimate of a/b = 1.67600
Best least squares value = 0.000258733747
a/b = 1.62198 A+ = 0.058319757156 A- = 0.061743449026

From a starting estimate of a/b = 1.83600
Best least squares value = 0.000258733747
a/b = 1.62199 A+ = 0.058319656141 A- = 0.061743559654

*****
Optimum best least squares value = 0.000258733747
          (a/b, b/c) = 1.622, 1.343
*****

In case of error output:
TYPE 2 or 3: No convergence - discard this result
TYPE 4: (unlikely) Overflow - discard this result
TYPE 5-8: some doubt about this result. The higher the
error no. the greater the doubt.
TYPE 5 means this value is almost certainly reliable;
TYPE 8 means this result is very doubtful
TYPE 9: There is probably a mistake in your constraining
data or your experimental values for mol. wt, temperature,
intrinsic viscosity or solvent viscosity

```

(c)

Figure 4 (Continued)

extensively explored with the earlier mainframe version of the program.^{36,49,53} A possible area of further improvement includes the additional constraint that $A'_+ + A'_- = \Delta n_{\max}$ although this may cause problems if the data are noisy and the $t=0$ position is not precisely defined.

9 Concluding Comment

The routines are downloadable from the NCMH web site <http://www.nottingham.ac.uk/ncmh> complete with full instructions. In any publication users are requested to acknowledge Salford Software and the Numerical Algorithms Group, Oxford for ELLIPS2-4 and BUSS Limited for use of SimplePlot Library routines in ELLIPS3.

References

1. S. E. Harding, H. Cölfen, and J. C. Horton, *Eur. Biophys. J.*, 1987, **25**, 347.
2. J. Garcia de la Torre, S. Navarro, M. C. Lopez Martinez, F. G. Diaz and J. J. Lopez Cascales, *Biophys. J.*, 1994, **67**, 530.
3. J. Garcia de la Torre, B. Carrasco and S. E. Harding, *Eur. Biophys. J.*, 1997, **25**, 361.
4. J. Garcia de la Torre and B. Carrasco, *Biopolymers*, 2002, **63**, 163.
5. J. Garcia de la Torre, this volume.
6. S. E. Harding, in *Dynamic Properties of Biomolecular Assemblies*, S. E. Harding and A. J. Rowe (eds), Royal Society of Chemistry, Cambridge, UK, 1989, 32–56.
7. C. Tanford, *Physical Chemistry of Macromolecules*, chap. 4, Wiley New York, 1961.
8. B. Carrasco, J. Garcia de la Torre, O. Byron, D. King, C. Walters, S. Jones and S. E. Harding, *Biophysical J.*, 1999, **77**, 2902.
9. B. Carrasco, J. Garcia de la Torre, K. G. Davis, S. Jones, D. Athwal, C. Walters, D. R. Burton and S. E. Harding, *Biophys. Chem.*, 2001, **93**, 181.
10. E. Longman, K. Kreusel, S. B. Tendler, I. Fiebrig, K. King, J. Adair, P. O'Shea, A. Ortega, J. Garcia de la Torre and S. E. Harding, *Eur. Biophys. J.*, 2003, **32**, 503.
11. S. E. Harding, E. Longman, A. Ortega, K. Kreusel, S. B. Tendler, K. King and J. Garcia de la Torre, *Prog. Coll. Polym. Sci.*, 2004, **127**, 113.
12. S. E. Harding, E. Longman, B. Carrasco, A. Ortega and J. Garcia de la Torre, *Method. Mol. Biol.*, 2004, **248**, 93.
13. E. Longman, Ph.D. Dissertation, University of Nottingham, 2005.
14. S. E. Harding, *Biophys. Chem.*, 1995, **55**, 69.
15. R. Simha, *J. Phys. Chem.*, 1940, **44**, 25.
16. N. Saito, *J. Phys. Soc. (Japan)*, 1951, **6**, 297.
17. A. Einstein, *Ann. Physik*, 1906, **19**, 289.
18. A. Einstein, *Ann. Physik*, 1911, **34**, 591.
19. F. Perrin, *J. Phys. Radium*, 1936, **7**, 1.
20. J. M. Rallison and S. E. Harding, *J. Coll. Int. Sci.*, 1985, **103**, 284.
21. P. D. Jeffrey, L. W., Nichol, D. R. Turner and D. J. Winzor, *J. Phys. Chem.*, 1977, **81**, 776.
22. K. E. Van Holde, *Physical Biochemistry*, 1st edn, Prentice-Hall, Englewood Cliffs, NJ, 1985, 171–172.
23. K. E. Van Holde, *Physical Biochemistry*, 2nd edn, Prentice-Hall, Englewood Cliffs NJ, 1985, 198–199.
24. F. Perrin, *J. Phys. Radium*, 1934, **5**, 497.
25. D. Ridgeway, *J. Am. Chem. Soc.*, 1966, **88**, 1104.

26. D. Ridgeway, *J. Am. Chem. Soc.*, 1968, **90**, 18.
27. W. A. Wegener, R. M. Dowben and V. J. Koester, *J. Chem. Phys.*, 1979, **70**, 622.
28. H. Scheraga and L. Mandelkern, *J. Am. Chem. Soc.*, 1953, **75**, 179.
29. S. E. Harding, *Int. J. Biol. Macromol.*, 1981, **3**, 340.
30. M. Wales and K. E. Van Holde, *J. Polym. Sci.*, 1954, **14**, 81.
31. A. J. Rowe, *Biopolymers*, 1977, **16**, 2595.
32. J. M. Creeth and C. G. Knight, *Biochim. Biophys. Acta*, 1965, **102**, 549.
33. A. J. Rowe, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. E. Harding, A. J. Rowe and J. C. Horton (eds), Royal Society of Chemistry, Cambridge, UK, 1992, 394–406.
34. S. E. Harding, *Biophys. J.*, 1987, **51**, 673.
35. S. E. Harding, *Biochem. J.*, 1980, **189**, 359; vol. **189** corrigenda (correction in the formula for τ_h).
36. S. E. Harding and A. J. Rowe, *Biopolymers*, 1983, **22**, 1813 and **23**, 843.
37. S. E. Harding and H. Cölfen, *Anal. Biochem.*, 1995, **228**, 131.
38. M. Kirkitadze, K. Jumel, S. E. Harding, D. T. F. Dryden, M. Krych, J. P. Atkinson and P. N. Barlow, *Prog. Coll. Polym. Sci.*, 1999, **113**, 164.
39. NAG, *Workstation Library Manual*, Numerical Algorithms Group, Jordan Hill, Oxford, UK, 1991.
40. R. L. Brady, D. J. Edwards, R. E. Hubbard, J. -S. Juang, S. M. Lange, S.M. Roberts, R. J. Todd, J. R. Adair, J. S. Emtage, D. J. King and D. C. Low, *J. Mol. Biol.*, 1992, **227**, 253.
41. W. R. Taylor, J. M. Thornton and W. G. Turnell, *J. Mol. Graphics*, 1983, **1**, 30.
42. S. E. Harding, *Prog. Biophys. Mol. Biol.*, 1997, **68**, 207.
43. E. Longman, N. Marheineke and S. E. Harding, *GIT Separation*, 2005, **2**, in press.
44. P. Nicolas, G. Batelier, M. Rholam and P. Cohen, *Biochemistry*, 1981, **19**, 3563.
45. S. E. Harding and A. J. Rowe, *Int. J. Biol. Macromol.*, 1982, **4**, 357.
46. D. Pörschke and A. Obst, *Rev. Sci. Instrum.*, 1991, **62**, 818.
47. S. E. Harding, *Biochem. Soc. Trans.*, 1986, **14**, 857.
48. E. W. Small and I. Isenberg, *Biopolymers*, 1977, **16**, 1907.
49. S. E. Harding, Ph.D. Dissertation, University of Leicester, UK, 1980.
50. J. W. Jost and C. T. O'Konski, *Mol. Electro-Optics*, 1978, **2**, 529.
51. D. V. O'Connor, W. R. Ware and J. C. Andre, *J. Phys. Chem.*, 1979, **83**, 1333.
52. R. M. Johnsen and W. Brown, in *Laser Light Scattering in Biochemistry*, S. E. Harding, D. B. Sattelle, V. A. L. Bloomfield (eds), Royal Society of Chemistry, Cambridge, UK, 1992, 77–91.
53. S. E. Harding, *Comput. Biol. Med.*, 1983, **13**, 89.
54. C. L. Riddiford and B. Jennings, *Biopolymers*, 1967, **5**, 757.