Eur Biophys J (1997) 25: 373–384

ARTICLE

Bruno Spotorno · Luca Piccinini · Giovanni Tassara
Carmelina Ruggiero · Marco Nardini
Francesco Molina · Mattia Rocco

BEAMS (BEAds Modelling System): a set of computer programs for the generation, the visualization and the computation of the hydrodynamic and conformational properties of bead models of proteins

Received: 1 September 1996 / Accepted: 1 December 1996

Abstract Spheres, cylinders or ellipsoids, whose hydrodynamic properties can be computed from analytical or semi-analytical expressions, have been traditionally used as low-resolution approximate descriptors of macromolecular size and shape. However, these simple geometrical bodies can seldom faithfully reproduce any detail of a macromolecular surface. A more sophisticated procedure utilizes instead ensembles of spheres ("beads") of various diameters in an appropriate spatial arrangement to model the macromolecule. This method has not yet gained widespread application, partially because of the difficulties involved both in the generation and in the handling of the models, and because of the rather complicated mathematics involved in the computation of the hydrodynamic parameters, requiring non-trivial dedicated computer software virtually unavailable in the public-domain. A notable exception was the "TRV" program and its predecessors developed by the García de la Torre's group, which have been recently updated and made available as the "HYDRO" package (García de la Torre et al. 1994).

To make accessible as many aspects as possible of this powerful modelling procedure, we have assembled a set of computer programs written in C language called BEAMS (BEAds Modelling System), which are described in this paper. The main BEAMS programs provide the user with a choice of four different methods for the computation of the hydrodynamic and structural parameters of ensembles of beads, with the option of automatically generating many random conformations of linear, branched and/or looped strings-of-beads. Selected models can be visualized from any desired point of view and manipulated interactively on a high-resolution colour monitor, and plotted as two-dimensional projections on an eight colour plotter. A further option permits the calculation of the parameters for segmentally flexible models composed of two subunits. Two ancillary programs, PROMOLP (written in Visual-Basic for Windows™) and GRUMB (written in C for general PC use) help the user in the definition of the number and radius of the beads to be used for the models, and in the interactive construction of spatially pre-defined models. BEAMS should be especially useful in the generation of low-resolution models of large-sized proteins which are difficult or impossible to solve with high-resolution techniques such as X-ray crystallography or NMR, and, in multidomain/modular proteins, in defining the overall spatial arrangement of the various domains/modules from their known 3D-structures.

Key words Translational diffusion · Rotational diffusion · Relaxation times · Intrinsic viscosity · Protein structure and dynamics

1. Introduction

The determination of the three-dimensional (3D) structure of proteins and other biomacromolecules is a crucial step in the understanding of their biological functions. High-resolution techniques, namely X-ray crystallography and, more recently, nuclear magnetic resonance (NMR), have been applied very successfully in solving the 3D structure of so far roughly 1,000 proteins. Unfortunately, these methods cannot always be employed to yield the desired information, owing to some inherent limitations still present. In particular, large-sized proteins ($M_r > 100{,}000$), while being for various reasons often difficult to crystallize, are well beyond the capabilities of today's most advanced NMR methods, which now have an upper limit of around M_r 30,000.

In the absence of high-resolution data, useful information on the size and shape of a protein in solution can be gained from hydrodynamic characterization. This involves the determination of some parameters which are size- and shape-related, such as the translational and rotational fric-

B. Spotorno · M. Nardini · F. Molina · M. Rocco (✉)
Gruppo di Biostrutture, Istituto Nazionale per la Ricerca
sul Cancro (IST), Centro per le Biotecnologie Avanzate (CBA),
Largo Rosanna Benzi 10, I-16132 Genova, Italy
(Fax: +39-10-5737-306; e-mail: rocco@ermes.cba.unige-it)

B. Spotorno · L. Piccinini · G. Tassara · C. Ruggiero
Dipartimento di Informatica, Sistemistica e Telematica (DIST),
Università di Genova, Viale Causa 13, I-16145 Genova, Italy

tional coefficients and the intrinsic viscosity, utilizing well-established techniques such as, for example, sedimentation velocity, fluorescence anisotropy decay, dynamic light scattering, dielectric relaxation, electric and flow birefringence and NMR (Cantor and Schimmel 1980b; Marshall 1978). Additional structural parameters, such as the radius of gyration, can be obtained by employing various elastic radiation scattering techniques, from Rayleigh light scattering to small-angle neutron or X-ray scattering (Cantor and Schimmel 1980b; Marshall 1978). Traditional modelling involves comparing the measured values of the hydrodynamic parameters with those that can be calculated for a simple geometrical solid body of appropriate dimensions, such as a sphere, a cylinder or an ellipsoid of revolution (Cantor and Schimmel 1980b; Marshall 1978). However, these procedures yield a rather approximate picture of the protein, which is often of little value. The situation has been improved by the development of the hydrodynamic theory for general tri-axial ellipsoids (see Harding 1989; Harding 1997), but multi-subunit or highly irregular shapes are still difficult to model with this approach. A more advanced method, first introduced in the biopolymers field by V. Bloomfield and collaborators (Bloomfield et al. 1967a, b), and subsequently expanded and improved by the same authors and by a number of other workers (see for example García de la Torre and Bloomfield 1977a, b, c; García de la Torre and Bloomfield 1978; Wegener 1981; García de la Torre and Rodes 1983; Goldstein 1985; the reviews by García de la Torre and Bloomfield 1981; García de la Torre 1989), consists of modelling the macromolecule as a set of spheres (beads) of various diameters in an appropriate spatial arrangement, of which the hydrodynamic and structural parameters can be computed and compared with the measured ones. This permits a much more accurate description of the shape of a protein, and when pushed to its limits, by "coating" the surface of small proteins of known 3D-structure with many atom-sized beads, has been shown to be able to reproduce quite satisfactorily the values of their measured translational diffusion coefficients (Teller et al. 1979). Later, the influence of tigthly bound water molecules on the calculated parameters for both the translational and rotational diffusion of lysozyme and ribonuclease has been studied with this method, with very encouraging results (Venable and Pastor 1988). Another area of intense work is the application of this procedure to the study of segmentally flexible macromolecules (Harvey 1979; Wegener et al. 1980; Wegener 1982a; Harvey et al. 1983; Mellado et al. 1988), which will allow a more accurate description of many real life proteins, such as antibodies or myosin (Yguerabide et al. 1970; Wegener 1982b; Iniesta et al. 1988). In addition, as sizeable body of large-size, mainly extracellular proteins, are "modular", e.g. they are made up by multiple copies of the same or different basic structural units ("domains" or "modules", see Doolittle 1995; Bork et al 1996; and references therein), whose relatively small size renders them amenable to high-resolution studies. Modelling the single structural units in modular proteins with beads could provide a powerful method for studying their overall 3D conformation.

Although this modelling procedure is very versatile and conceptually simple, it has not yet gained widespread application, for two main reasons: it requires the writing of non-trivial computer software for the computation of the hydrodynamic properties of the ensembles of beads, and a series of often non-straightforward decisions have to be made regarding the number, the dimensions and the spatial positions of the beads to be used in the modelling. The first hurdle has been at least partially removed by the recent appearance of a freeware version of the computer code written by the group led by one of the most active theoreticians in the field, Josè García de la Torre, namely the "HYDRO" routines package (García de la Torre et al. 1994). HYDRO is a suite of FORTRAN routines for the calculation of the hydrodynamic properties of rigid macromolecules in solution, which requires as input the number, coordinates and radii of the beads in the model, and calculates a number of hydrodynamic parameters (and the radius of gyration) of the bead model, via direct inversion of the coefficient "supermatrix". The authors also suggest how HYDRO could be used for the computation of the properties of flexible macromolecules, but additional dedicated routines must be provided by the user.

By comparison, BEAMS has been designed from the very beginning (Ruggiero et al. 1990) as a more comprehensive and flexible software package, which can be used in every step of the bead modelling procedure, from the choice of the number and dimensions of the beads, to the building of models (either with a pre-defined spatial arrangement or with random chain-generation), to the visualization of the models on a computer screen and their plotting; the computation of the hydrodynamic and conformational parameters for rigid models is done via four alternative methods, and via two complementary methods for segmentally flexible models composed ot two subunits. A comprehensive description of the main BEAMS programs, as well as of two important ancillary programs, and a brief introduction to the ASA and TRANS programs which are being developed for the generation of bead models starting from atomic coordinates, are presented here.

2. Theory

The hydrodynamic theory of the frictional properties of rigid ensembles of beads was developed by a number of workers and has been fully described in many excellent reviews (see for example García de la Torre and Bloomfield 1981; García de la Torre 1989); for this reason we will report here only a brief summary and some of the relevant equations that we have utilized in our programs. For a full description of an alternative treatment which has also been implemented in BEAMS, the variational treatment, the reader is referred to the original work by Goldstein (1985).

The hydrodynamics of segmentally flexible ensembles of beads has also received much attention from theoreticians,but the issue is more complicated and far from being resolved. In particular, two approaches, the "rigid body" and the "Harvey-Wegener" treatments, have been developed, and are somewhat complementary. They are both summarized in a recent very useful review (García de la Torre, 1994). In our programs, only the case of a segmentally flexible model with two subunits has been implemented, utilizing both approaches. Again, only a summary of the relevant equations implemented in our programs will be given here.

2.1 Rigid ensembles

The translational-rotational dynamics of a rigid particle of arbitrary shape can be described by a 6×6 resistance matrix $\mathbf{R}_O$ which, under low-Reynolds-number conditions, directly relates the three forces and the three torques acting on the particle to its linear and angular velocities (Happel and Brenner 1973):

$$\mathcal{F} = \mathbf{R}_O \cdot \mathbf{v} \qquad (1)$$

where $\mathcal{F}$ and $\mathbf{v}$ are the six-dimensional force and velocity vectors, respectively.

Usually, $\mathbf{R}_O$ is partitioned in four 3×3 blocks containing the translational, rotational, and roto-translational coupling frictional tensors Ξ_t, $\Xi_{O,r}$ and $\Xi_{O,c}$, and the transpose of the latter, $\Xi_{O,c}^T$, so Eq. (1) can be rewritten as:

$$\begin{pmatrix} \mathbf{F} \\ \mathbf{T}_O \end{pmatrix} = \begin{pmatrix} \Xi_t & \Xi_{O,c}^T \\ \Xi_{O,c} & \Xi_{O,r} \end{pmatrix} \begin{pmatrix} \mathbf{u}_O \\ \boldsymbol{\omega} \end{pmatrix} \qquad (2)$$

where $\mathbf{F}$ and $\mathbf{T}_O$ are the force and torque vectors, respectively, and $\mathbf{u}_O$ and ω are the corresponding linear and angular velocities, respectively. The O in the subscripts indicates that the values of the corresponding vectors or tensors are origin-dependent. The coupling frictional tensor is symmetric only at a particular point called the center of reaction R, and the distance vector $\mathbf{r}_{OR}$ relating $\Xi_{O,c}$ to $\Xi_{R,c}$ is given by Eq. (9) of García Bernal and García de la Torre (1980).

A relation analogous to Eq. (2) can be written for the diffusion tensors, and a center of diffusion D exists for which the diffusion coupling tensor $\mathbf{D}_{O,c}$ is symmetric. The position of D can be found by substituting Ξ_t with $\mathbf{D}_r$ and $\Xi_{O,r}$ with $-\mathbf{D}_{O,c}$ in Eq. (9) of García Bernal and García de la Torre (1980). The relation between the diffusion matrix $\mathbf{D}_O$ and the resistance matrix $\mathbf{R}_O$ is given by the generalized Stokes-Einstein equation (Brenner 1967):

$$\mathbf{D}_O = k_B T \mathbf{R}_O^{-1} \qquad (3)$$

where k_B is Boltzmann's constant and T is the absolute temperature.

For an ensemble of N beads, it is possible to calculate from Stokes' law the frictional force exerted on the solvent by each bead. However, the motion of each bead creates an internal velocity field in the solvent that must be added to the external one. This "hydrodynamic interaction" can be described by (García de la Torre and Bloomfield 1977a):

$$\mathbf{F}_i (6\pi \eta_0 \sigma_i)^{-1} + \sum_{j=1}^{N}{}' \mathbf{T}_{ij} \cdot \mathbf{F}_j = (\mathbf{u}_i - \mathbf{v}_i^0) \qquad (4)$$

where $6\pi\eta_0\sigma_i$ is the frictional coefficient of a sphere of radius σ_i in a solvent with viscosity η_0, $\mathbf{F}_i$ is the frictional force exerted on the solvent by the i^{th} element, $\mathbf{u}_i$ and $\mathbf{v}_i^0$ are respectively the velocity of the i^{th} element and the velocity of the unperturbed solvent, and the prime on the summation indicates the exclusion of the terms with $i=j$. $\mathbf{T}_{ij}$ is the Rotne-Prager-Yamakawa hydrodynamic interaction tensor as modified by García de la Torre and Bloomfield (Rotne and Prager 1969; Yamakawa 1970; García de la Torre and Bloomfield 1977a) for non-overlapping beads of different size,

$$\mathbf{T}_{ij} = (8\pi \eta_0 R_{ij})^{-1} \left[\mathbf{I} + \frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{R_{ij}^2} + \frac{(\sigma_i^2 + \sigma_j^2)}{R_{ij}^2} \left(\tfrac{1}{3}\mathbf{I} - \frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{R_{ij}^2} \right) \right] \qquad (5)$$

where R_{ij} and $\mathbf{R}_{ij}$ are, respectively, the scalar and the vectorial distances between the centers of beads i and j, and $\mathbf{I}$ is the unity tensor.

Equation (4) can be rewritten as (García de la Torre 1989):

$$\sum_{j=1}^{N} \mathbf{B}_{ij} \cdot \mathbf{F}_j = (\mathbf{u}_i - \mathbf{v}_i^0) \qquad (6)$$

where $\mathbf{B}_{ij}$ is a 3×3 matrix defined as:

$$\mathbf{B}_{ij} = \delta_{ij} \frac{\mathbf{I}}{6\pi \eta_0 \sigma_i} + (1 - \delta_{ij}) \mathbf{T}_{ij} \qquad (7)$$

where δ_{ij} is Kronecker's delta.

Equation (6) represents a system of N linear equations with 3N unknowns. In particular, by direct inversion of the "supermatrix" $\mathcal{B}$ of dimension 3N×3N, composed of N×N $\mathbf{B}_{ij}$ blocks, one obtains a new "supermatrix" $\mathcal{C}$ (García de la Torre 1989):

$$\mathcal{C} = \mathcal{B}^{-1} \qquad (8)$$

which contains N×N $\mathbf{C}_{ij}$ 3×3 blocks, leading to the elements of the resistance matrix $\mathbf{R}_O$:

$$\Xi_t = \sum_i \sum_j \mathbf{C}_{ij} \qquad (9)$$

$$\Xi_{O,c} = \sum_i \sum_j \mathbf{U}_i \cdot \mathbf{C}_{ij} \qquad (10)$$

$$\Xi_{O,r} = -\sum_i \sum_j \mathbf{U}_i \cdot \mathbf{C}_{ij} \cdot \mathbf{U}_j + 6\eta_0 V \mathbf{I} \qquad (11)$$

In Eqs. (10)–(11), $\mathbf{U}_i$ and $\mathbf{U}_j$ are two matrices whose elements are the components of the vectors joining the center of beads i and j with the origin of the reference system (Harvey et al. 1983; García de la Torre 1989). The last term on the right-hand side in Eq. (11), $6\eta_0 V \mathbf{I}$, is the "volume" correction introduced by García de la Torre and Rodes (1983), where V is the total volume of the beads.

A less precise but less computer-demanding way of solving Eq. (4), is by Gauss-Seidel iterations. The procedure is described in detail in García de la Torre and Bloomfield (1981), Eqs. (7)–(8), (10), (21)–(23), and directly yields Ξ_t; it utilizes the "shielding tensors" $\mathbf{G}_i$, defined by Eq. (16) of García de la Torre and Bloomfield (1977a), from which the coupling tensor $\Xi_{O,c}$ can also be obtained (Eq. (60) in García de la Torre and Bloomfield (1981)).

From the trace of Ξ_t^{-1}, the translational frictional coefficient $f_t^{(\eta_0)}$ can be derived:

$$f_t^{(\eta_0)} = 3/\mathrm{tr}(\Xi_t^{-1}) \tag{12}$$

The translational diffusion coefficient $D_t^{(T,\eta_0)}$ at the temperature T and solvent viscosity η_0, and the corresponding translational Stokes' radius R_S^t can be then obtained from $f_t^{(\eta_0)}$ as:

$$D_t^{(T,\eta_0)} = \frac{k_B T}{f_t^{(\eta_0)}} \tag{13}$$

$$R_S^t = f_t^{(\eta_0)}/6\pi\eta_0 \tag{14}$$

To calculate the rotational diffusion coefficients along the tree principal axes of the model in the presence of roto-translational coupling, the rotational frictional tensor must be evaluated at the center of reaction R, $\Xi_{R,r}$, from $\Xi_{O,r}$ (García Bernal and García de la Torre 1980):

$$\Xi_{R,r} = \Xi_{O,r} - \mathbf{r}_{OR} \times \Xi_t \times \mathbf{r}_{OR} + \Xi_{O,c} \times \mathbf{r}_{OR} - \mathbf{r}_{OR} \times \Xi_{O,c}^T \tag{15}$$

where the symbol "×" denotes the dyadic product (García Bernal and García de la Torre 1980).

$\Xi_{R,r}$ can be also obtained by Gauss-Seidel iterations as described by García de la Torre and Bloomfield (1981), utilizing the vectors $\mathbf{R}_i'$ joining the center of resistance R to the center of bead i. The "volume" correction (García de la Torre and Rodes 1983) is also applied here.

The three rotational diffusion coefficients along the axes x^k (k=1, 2, 3), $D_r^{k(T,\eta_0)}$, and the corresponding rotational Stokes' radii, $R_S^{r,k}$, are given by:

$$D_r^{k(T,\eta_0)} = k_B T[\Xi_{R,r}^{kk}]^{-1} \tag{16}$$

$$R_S^{r,k} = \left[\frac{\Xi_{R,r}^{kk}}{8\pi\eta_0}\right]^{1/3} \tag{17}$$

Moreover, from the eigenvalues of the rotational diffusion matrix, $\mathbf{D}_r$, five rotational relaxation times τ_k^r (k=1–5), and their unweighted harmonic mean τ_h^r, can be calculated in the general case when $D_1 \le D_2 \le D_3$ (Wegener et al. 1979):

$$\left.\begin{aligned} \tau_1^{r(T,\eta_0)} &= (6D - 2\Delta)^{-1} \\ \tau_2^{r(T,\eta_0)} &= (3D + 3D_1)^{-1} \\ \tau_3^{r(T,\eta_0)} &= (3D + 3D_2)^{-1} \\ \tau_4^{r(T,\eta_0)} &= (3D + 3D_3)^{-1} \\ \tau_5^{r(T,\eta_0)} &= (6D + 2\Delta)^{-1} \end{aligned}\right\} \tag{18}$$

$$\tau_h^{r(T,\eta_0)} = \left[\left(\sum_k \left(\tau_k^{r(T,\eta_0)}\right)^{-1}\right)/5\right]^{-1} \tag{19}$$

where:

$$D = (1/3)\,(D_1 + D_2 + D_3) \tag{20}$$

$$\Delta = (D_1^2 + D_2^2 + D_3^2 - D_1 D_2 - D_2 D_3 - D_1 D_3)^{(1/2)} \tag{21}$$

Finally, the intrinsic viscosity $[\eta]$ is obtained (García de la Torre and Bloomfield 1978; García de la Torre and Bloomfield 1981; García de la Torre 1989) from:

$$\begin{aligned} [\eta] = \frac{N_A}{M\eta_0} \sum_i \sum_j \Bigg[& \frac{1}{15} \sum_\alpha (x_i^\alpha - v^\alpha)\, \mathbf{C}_{ij}^{\alpha\alpha}\, (x_j^\alpha - v^\alpha) \\ & + \frac{1}{20} \sum\sum_{\alpha\ne\beta} (x_i^\alpha - v^\alpha)\, \mathbf{C}_{ij}^{\beta\alpha}\, (x_j^\beta - v^\beta) \\ & - \frac{1}{30} \sum\sum_{\alpha\ne\beta} (x_i^\alpha - v^\alpha)\, \mathbf{C}_{ij}^{\alpha\beta}\, (x_j^\beta - v^\beta) \\ & + \frac{1}{20} \sum\sum_{\alpha\ne\beta} (x_i^\alpha - v^\alpha)\, \mathbf{C}_{ij}^{\beta\beta}\, (x_j^\alpha - v^\alpha) \Bigg] \quad (\alpha, \beta = 1, 2, 3) \end{aligned} \tag{22}$$

Here N_A is Avogadro's number, M is the total mass of the beads, and x_i^α and v^α are, respectively, the coordinates of bead i and of the center of viscosity V, defined as the point where $[\eta]$ is at a minimum (García de la Torre and Bloomfield 1978). Thus, V can be found from a system of three linear equations whose coefficients are combinations of the $\mathbf{C}_{ij}$ and x_i^α (García de la Torre and Bloomfield 1978), the first of which is:

$$\begin{aligned} & v^x \sum_i \sum_j (8\mathbf{C}_{ij}^{xx} + 6\mathbf{C}_{ij}^{yy} + 6\mathbf{C}_{ij}^{zz}) \\ & \quad + v^y \sum_i \sum_j (\mathbf{C}_{ij}^{xy} + \mathbf{C}_{ij}^{yx}) + v^z \sum_i \sum_j (\mathbf{C}_{ij}^{xz} + \mathbf{C}_{ij}^{zx}) \\ & = \sum_i \sum_j [(x_i^x + x_j^x)(4\mathbf{C}_{ij}^{xx} + 3\mathbf{C}_{ij}^{yy} + 3\mathbf{C}_{ij}^{zz}) \\ & \qquad + x_j^y (3\mathbf{C}_{ij}^{yx} - 2\mathbf{C}_{ij}^{xy}) + x_j^z (3\mathbf{C}_{ij}^{zx} - 2\mathbf{C}_{ij}^{xz}) \\ & \qquad + x_i^y (3\mathbf{C}_{ij}^{xy} - 2\mathbf{C}_{ij}^{yx}) + x_i^z (3\mathbf{C}_{ij}^{xz} - 2\mathbf{C}_{ij}^{zx})] \end{aligned} \tag{23}$$

The other two equations are obtained by cyclic permutation of (x, y, z).

Recently, García de la Torre (1989) has advocated the general use in bead modelling procedures of a "viscosity" correction in the computation of $[\eta]$ similar to the "volume" correction included in Eq. (11) for the rotational frictional coefficients. It consists in adding to the right-hand side of Eq. (22) a term corresponding to the intrinsic viscosity that would be exhibited by a sphere having the a volume V and a mass M respectively equal to the total volume and total mass of the beads:

$$[\eta]^{\mathrm{corr}} = [\eta] + \frac{5 N_A V}{2M} \tag{24}$$

However, a numerical computation study utilizing shell-models of prolate ellipsoids of revolution that we have conducted (Spotorno et al., submitted), has cast some doubts on the validity of this correction for rigid ensembles of beads. Therefore, the two $[\eta]$ values (uncorrected and corrected) are separately reported.

By analogy with the Stokes' radii for translational and rotational diffusion (Eqs. (14) and (17)), it is possible to calculate from $[\eta]$ the so-called Einstein radius R_E as:

$$R_E = \left[\frac{3M[\eta]}{10\pi N_A}\right]^{\frac{1}{3}} \tag{25}$$

We end this section by recalling that a conformational property, the rms radius of gyration R_g, can be easily computed from the beads' masses m_i, and the distances r_i from the coordinates of the center of mass of the ensemble of beads:

$$R_g = \frac{\sum_i m_i r_i^2}{\sum_i m_i} \tag{26}$$

2.2 Segmentally flexible ensembles with two "subunits"

2.2.1 Harvey-Wegener treatment

The formalism of Harvey has been fully developed for the particular case of a segmentally flexible model composed of two independent subunits (Harvey et al. 1983; Mellado et al. 1988), and is the one that we have implemented in our programs. Such a model, in which the "joint" between the subunits can either act as a "hinge" (only bending, no torsions) or as a "swivel" (bending and torsions, see Fig. 1 of Harvey et al. 1983), has a maximum of nine degrees of freedom, three for instantaneous overall translation, three for instantaneous overall rotations and three for the instantaneous internal bending and torsions. Harvey et al. (1983) used three coordinates for the overall translation, three for the overall rotations, and used an angle α for the bending between the subunits, and two angles β_1 and β_2 for the torsions between the subunits. For each subunit, the principal axis passing from the center of the swivel is defined, and two unit vectors, $\hat{\imath}_1$ and $\hat{\imath}_2$, are associated with these axes. The model is pre-oriented so that $\hat{\imath}_1$ and $\hat{\imath}_2$ define the xy plane, and the bending between the subunits described by α, e.g. between $\hat{\imath}_1$ and $\hat{\imath}_2$, takes place in that plane. Two non-independent parameters, a_1 and a_2, define the proportion of the angular velocity $d\alpha/dt$ associated with each subunit:

$$\omega_z^{(1)} = \omega_z + a_1 \frac{d\alpha}{dt} \tag{27}$$

$$\omega_z^{(2)} = \omega_z - a_2 \frac{d\alpha}{dt} \tag{28}$$

$$a_1 + a_2 = 1 \tag{29}$$

where ω_z is the component along the z-axis of the angular velocity of the whole ensemble and $\omega_z^{(1)}$ and $\omega_z^{(2)}$ those of the two subunits.

The advantages of this notation, such as the choice of where to fix the molecular coordinate system through the parameter $a_{(k)}$, and the relative equations, are fully described by Harvey et al. (1983), and will not be repeated here.

Following Harvey et al. (1983), we will now recall that, by analogy with the case of rigid ensembles, it is possible to generalize Stokes' law:

$$\mathbf{F}_\mathrm{O} = -\mathbf{R}_\mathrm{O} \cdot \mathbf{u}_\mathrm{O} \tag{30}$$

where $\mathbf{F}_\mathrm{O}$ and $\mathbf{u}_\mathrm{O}$ are, respectively, the nine vector generalized force and velocity, and $\mathbf{R}_\mathrm{O}$ is the 9×9 resistance tensor which can be again partitioned into 3×3 submatrices:

$$\mathbf{R}_\mathrm{O} = \begin{pmatrix} \Xi_\mathrm{tt} & \Xi_\mathrm{O,tr}^\mathrm{T} & \Xi_\mathrm{O,ti}^\mathrm{T} \\ \Xi_\mathrm{O,tr} & \Xi_\mathrm{O,rr} & \Xi_\mathrm{O,ri}^\mathrm{T} \\ \Xi_\mathrm{O,ti} & \Xi_\mathrm{O,ri} & \Xi_\mathrm{O,ii} \end{pmatrix} \tag{31}$$

Here a slightly different notation from that used for rigid ensembles has been used in the subscripts to indicate the coordinates associated with each tensor, in which "t" stands for translation, "r" for rotation, and "i" for the three internal coordinates of bending and torsions ("tt" indicate pure translation, "tr" translation-rotation coupling, and so on). To obtain the 9×9 diffusion tensor $\mathbf{D}_\mathrm{O}$, the generalized Stokes-Einstein relation (Eq. (3)) can be applied again.

$\mathbf{R}_\mathrm{O}$ is obtained as we have seen for rigid ensembles, and in this way the hydrodynamic interaction between elements on different subunits can be taken into account. The expressions utilized for the computation of the various 3×3 blocks from the $\mathbf{S}_\mathrm{ij}$ obtained from the $\mathbf{Q}_\mathrm{ij}$ defined in Eqs. (7)–(8) of García de la Torre and Bloomfield (1981) by the direct inversion of the coefficient "supermatrix", are Eqs. (27)–(32) of Harvey et al. (1983), with the addition of the "volume" correction as reported in Mellado et al. (1988). Instead of the $\mathbf{Q}_\mathrm{ij}$ and $\mathbf{S}_\mathrm{ij}$, the $\mathbf{B}_\mathrm{ij}$ and $\mathbf{C}_\mathrm{ij}$ defined in Eqs. (6)–(11) can be used advantageously (Mellado et al. 1988; García de la Torre 1989), but this later innovation has not yet been introduced in our programs.

It is important to stress that the parameters obtained are instantaneous values for any given conformation, and are valid if the origin is set at the joint between the two subunits (Harvey et al. 1983). Once the diffusion tensor $\mathbf{D}_\mathrm{O}$ is obtained from $\mathbf{R}_\mathrm{O}$, the instantaneous translational diffusion coefficient $\mathrm{D}_\mathrm{t}^{(\mathrm{T},\eta_0)}(\alpha)$ for a segmentally flexible macromolecule with two subunits could be calculated. However, the correct point where $\mathrm{D}_\mathrm{t}^{(\mathrm{T},\eta_0)}(\alpha)$ must be evaluated is at the center of diffusion. Wegener (1985) has shown that an excellent approximation to the "true" center of diffusion of a segmentally flexible body is the instantaneous center of diffusion, which is found by applying a relation similar to Eq. (9) of García Bernal and García de la Torre (1980). A transformation relation from $\mathbf{D}_\mathrm{O}$ and $\mathbf{D}_\mathrm{D}$ is applied (see Harvey et al. 1983), and $\mathrm{D}_\mathrm{t}^{(\mathrm{T},\eta_0)}(\alpha)$ is then obtained:

$$\mathrm{D}_\mathrm{t}^{(\mathrm{T},\eta_0)}(\alpha) = \tfrac{1}{3}\,\mathrm{tr}\,(\mathbf{D}_\mathrm{D,tt}) \tag{32}$$

The origin-independent instantaneous values for the three rotational diffusion coefficients for overall rigid rotations can also be extracted from $\mathbf{D}_\mathrm{O}$ (see Eq. (8) of Mellado et al. 1988) if the parameter $a_{(k)}$ is properly set. However, some experimental techniques, such as fluorescence anisotropy decay, can yield the rotational relaxation times for the reorientation of one subunit with respect to the other. In the formalism proposed by Wegener (1982a, b), which uses different rotational coordinates, such relaxation times can be easily calculated. According to Wegener, individual torsions of the subunits around their principal axes constitute the first two rotations, $\mathrm{D}_\parallel^{(k)}$, where $k = 1, 2$ indicate the subunits, $\mathrm{D}_\mathrm{in}^{(k)}$ describes the end-over-end rotations of each subunit within the xy plane, and $\mathrm{D}_\mathrm{out}^{(k)}$ is for out-of-plane rotations around the perpendicular z-axis. The transformation relations from Harvey's to Wegener's coordinate systems (Harvey et al. 1983) can be derived from Eqs. (9a)–(9c) of Mellado et al. (1988).

The average properties of segmentally flexible ensembles are then obtained by integration of the instantaneous values weighted by an appropriate probability distribution function. For instance, for R_g we have:

$$R_g = \int_{\alpha_{\min}}^{\alpha_{\max}} R_g(\alpha)\, p(\alpha)\, d\alpha \tag{33}$$

and analogous relations apply for $D_t^{(T,\eta_0)}$, $D_{kk}^{(T,\eta_0)}$, $\bar{D}_{\parallel}^{(k)}$, $\bar{D}_{out}^{(k)}$, $\bar{D}_{in}^{(k)}$.

For the more general case of unrestricted motions with nine degrees of freedom, the spherical probability distribution function of the angle described by the subunits is proportional to $\sin(\alpha)$ (Harvey et al. 1983), leading to:

$$p(\alpha) = \frac{\sin(\alpha)}{\int_{\alpha_{\min}}^{\alpha_{\max}} \sin(\alpha)\, d\alpha} = \frac{\sin(\alpha)}{\cos(\alpha_{\min}) - \cos(\alpha_{\max})} \tag{34}$$

Finally, from Wegener's coefficients five rotational relaxation times for each subunit can be extracted in the more general case (Wegener et al. 1980), which reduce to three for axially symmetric ensembles (Wegener 1982a; Mellado et al. 1988):

$$\left.\begin{aligned}
\tau_1^{r(i)} &= (6D + 2\Delta)^{-1}\\
\tau_2^{r(i)} &= (3D + \bar{D}_{\parallel}^{(i)})^{-1}\\
\tau_3^{r(i)} &= (3D + \bar{D}_{out}^{(i)})^{-1}\\
\tau_4^{r(i)} &= (3D + \bar{D}_{in}^{(i)})^{-1}\\
\tau_5^{r(i)} &= (6D - 2\Delta)^{-1}
\end{aligned}\right\} \tag{35}$$

where:

$$D = (1/3)(\bar{D}_{out}^{(i)} + \bar{D}_{in}^{(i)} + \bar{D}_{\parallel}^{(i)}) \tag{36}$$

$$\begin{aligned}\Delta = [&(\bar{D}_{out}^{(i)})^2 + (\bar{D}_{in}^{(i)})^2 + (\bar{D}_{\parallel}^{(i)})^2\\ &- \bar{D}_{\parallel}^{(i)}\bar{D}_{out}^{(i)} - \bar{D}_{\parallel}^{(i)}\bar{D}_{in}^{(i)} - \bar{D}_{out}^{(i)}\bar{D}_{in}^{(i)}]^{(1/2)}\end{aligned} \tag{37}$$

2.2.2 *Rigid body approximation*

As first pointed out by Mellado et al. (1988), the Harvey-Wegener treatment is completely valid only for freely jointed subunits (e.g. when the joints act as swivels), and even its applicability to hinged segments has been recently questioned (García de la Torre 1994). Instead, in the most realistic case a "restoring force" opposing the bending can be present (Harvey et al. 1983; Mellado et al. 1988; Iniesta et al. 1988; García de la Torre 1994). For these reasons, Iniesta et al. (1988) suggested that treating each instantaneous conformation in a semiflexible model as being rigid and then averaging with proper weighting would be a reasonable approximation, especially for properties which are expressions of overall motions such as the translational diffusion coefficient and the intrinsic viscosity.

As pointed out by Harvey et al. (1983), in the case of bending restrained by an harmonic potential limited by a constant restoring force, the probability distribution function should include an exponential term depending on α^2/k_BT. Using the same form as in Eq. (6) of Iniesta et al. (1988), if the torsions of the two subunits are opposed by identical linear restoring torques we have:

$$V_t/k_BT = Q_t(\beta_1 + \beta_2)^2 \tag{38}$$

where V is the harmonic potential and Q is the elastic constant.

For a swivel-jointed model with two subunits in the presence of a linear restoring force opposing the bending, but in the absence of linear torques opposing the torsions of the subunits, Eq. (34) becomes:

$$p(\alpha) = \frac{\sin(\alpha)\, e^{(-Q\alpha^2)}}{\int_{\alpha_{\min}}^{\alpha_{\max}} \sin(\alpha)\, e^{(-Q\alpha^2)}\, d\alpha} \tag{39}$$

In the most general case, for a swivel-jointed model with two subunits in the presence of a linear restoring force and linear torques, we have:

$$p(\alpha) = \frac{\sin(\alpha)\, e^{(-Q\alpha^2)}\, e^{(-Q_t(\beta_1+\beta_2)^2)}}{\int_{\alpha_{\min}}^{\alpha_{\max}} \sin(\alpha) e^{(-Q\alpha^2)} d\alpha \int_{\beta_{1(\min)}}^{\beta_{1(\max)}} \int_{\beta_{2(\min)}}^{\beta_{2(\max)}} e^{(-Q_t(\beta_1+\beta_2)^2)} d\beta_1\, d\beta_2} \tag{40}$$

The factor $\sin(\alpha)$ is removed in both Eqs. (39) and (40) for a hinged model.

The computation of the hydrodynamic and conformational parameters for segmentally flexible bead models in the rigid-body approximation is then carried out for each conformation with the formalism seen for completely rigid models, and then the averaging of each property is carried out as is done in Eq. (33) for R_g with the proper weighting factor and probability distribution function.

3. Materials and methods

BEAMS was originally developed on a HP9000 series 300 microcomputer equipped with an MC68040 processor, MC68882 numerical coprocessor, and 8 Mb RAM. The programming language chosen was C++, for its greater flexibility, including graphic capabilities, compared to FORTRAN. The programs have since been successfully recompiled on various other platforms, including a PC with a Pentium 75 MHz processor and 16 Mb RAM, a DEC System 5100 with 16 Mb RAM, a Sun SPARCstation 10 with 32 Mb RAM, and a Silicon Graphics Indy with 32 Mb RAM.

The ancillary programs PROMOLP (PROtein MOLecular Parameters) and GRUMB (GRaphical Utilities for Modelling with Beads) were developed on various PC computers from the 386 to the 486 series. PROMOLP and GRUMB are written and compiled in VisualBasic for Windows™ and C++ for MS-DOS™, respectively.

PROMOLP stores, retrieves and analyzes a protein sequence, either entered manually or converted from the SWISSPROT data bank format (see Rodriguez-Tomé et al. 1996), including information about number and position of free -SH groups, phospho-serines, -threonines and

-tyrosines, and N- and O-linked carbohydrate chains, to yield the following physico-chemical parameters: a) molecular weight of the amino acids included in the sequence; b) molecular weight of the carbohydrate chains; c) global molecular weight (amino acids plus carbohydrates plus eventual phospho-groups); d) partial specific volume of the amino acids only; e) global partial specific volume; f) theoretical hydration due only to the amino acids; g) global theoretical hydration; h) unhydrated molecular volume and radius of the equivalent sphere; i) hydrated molecular volume and radius of the equivalent sphere; j) theoretical extinction coefficient at 280 nm; k) average hydrophobicity [H(ϕ)]; l) R3 ratio (see Cantor and Schimmel 1980a); m) discriminant function Z (see Cantor and Schimmel 1980a); n) isoelectric point.

The program also calculates and displays in a table the net charge of the ionizable groups (amino acids plus carbohydrates) and the net total charge, for each unit of pH from pH 4 to 12. It is also possible to analyze only parts of the sequence. A full description of PROMOLP will be reported elsewhere (Rocco et al., manuscript in preparation).

With GRUMB, every step in the geometric building of a bead model is facilitated. Beads or blocks of beads can be added or removed, translated along an axis, rotated along an axis (either by a user-determined amount, or until contact is made with another bead), their properties (radius, mass and "colour") can be modified and symmetries can be generated. Spatial analyses can be performed between the beads (overlap test, top-down distance, distance between centers and surfaces of selected beads), and the entire structure can also be automatically rotated in order to have two chosen beads "aligned" on an axis, or it can be "normalized" (e.g., it can be aligned by a principal axis transformation (PAT) routine (Goldstein 1950) so as to have the maximum dimension along the x-axis, and the minimum dimension along the z-axis). Various visualization modes are present: orthogonal projection (2D), three-dimensional incoherent vision, and prospectic vision; the model can be represented by segments connecting the centers of the beads, by empty circles, or by filled circles (sixteen colours are currently available). As this program was conceived for model building only, no high-resolution three-dimensional "shading" modes are present, nor are graphical outputs to plotters or laser printers. The x, y and z coordinates of the beads are saved as three columns of n rows in a file which contains on the first line the number n of beads used, and a second number which is the radius of the beads if all the beads have the same radius, or is set to "0" and used as a "flag" if the beads have different radii; in this case, it is followed by the filename of another file which contains in three columns of n rows the radius, the mass, and a number specifying the "colour" of the beads ("rmc" file). In this way, multiple conformations of the same ensemble of beads can be generated without having to unnecessarily duplicate common information each time. This file structure is utilized by all the BEAMS programs.

As can be seen in Fig. 1, BEAMS is structured so as to have a main routine from which the various subprograms can be called. The five subprograms are GRAPH, COEFF, FLEX, SUPFLEX and STRINGS.

GRAPH is a graphical utility, with limited model manipulation capabilities, the most notable of which is that the model can be "preconditioned" for segmental flexibility studies with two subunits (practically speaking, it is aligned along the x-axis, and the 0,0,0 coordinates are placed wher the hinge or swivel is placed by the user). The models can be visualized either as segments connecting the centers of the beads, as coloured disks or as spheres in either low- or high-resolution, the latter only in black and white. To visualize the models in the disks or spheres modes, GRAPH utilizes the "Starbase" graphical routines package from Hewlett-Packard on our HP9000 platform (where it is set up for either an HP-98789A or an HP-98754A high-resolution colour monitor), but it has already been adapted for use on a PC. On the HP9000, GRAPH is fully interfaced to a HP-7550A 8-colour plotter through another self-standing program, PLOT. On plotting, the beads are represented as circles, but no shadowing or filling is actually implemented. Among other plotting features are: a) the size (A4 or A3) and orientation (vertical or horizontal) of the paper can be chosen; b) more than one file can be plotted on the same sheet, conserving the same scale: c) the location in the sheet where each file is to be

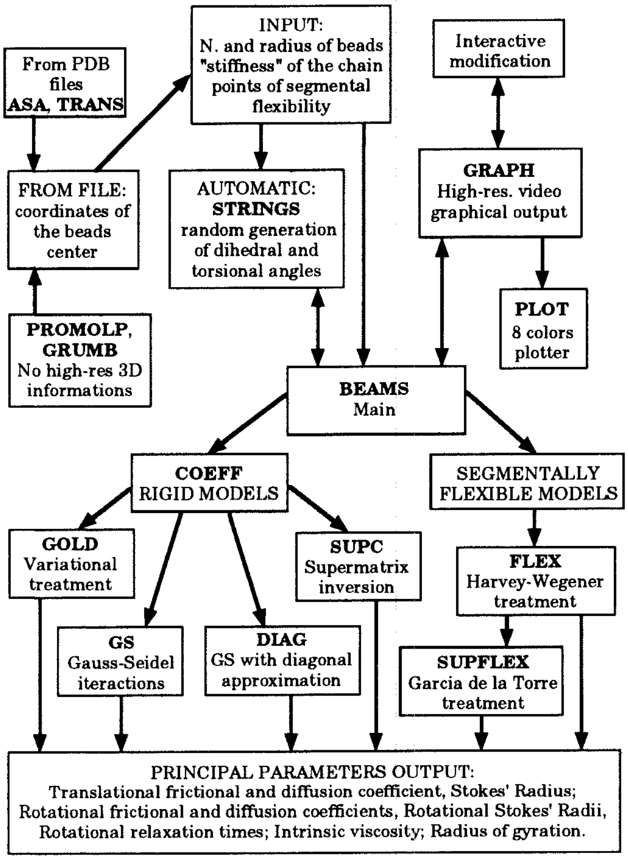

Fig. 1 Flow chart of the BEAMS package. Program names are in bold

plotted can be chosen (either in the center, in the upper/lower half, in one of four quadrants, or to manually enter the coordinates); d) for the whole plot, a scaling factor can be chosen; e) the speed of drawing can be varied, as well as the amount of "overlap" between contiguous beads; f) plotting the same files on up to 10 sheets is another option.

COEFF calls one of four programs, SUPC, GS, DIAG and GOLD, which compute the hydrodynamic and conformational parameters for rigid ensembles of beads. The main features common to all the COEFF subprograms are: a) they can process up to 100 files, either by entering the names one-by-one or as sequential files (common prefix + sequential number); b) the standard units in input for radii and coordinates are in nanometers (and so are the outputs such as the Stokes' radius). Therefore, there is a conversion feature that lets the user use any units as input units; c) choice of computation of hydrodynamic parameters at the Cartesian origin or at the diffusion center; d) choice between stick and slip boundary conditions; e) choice of the first and last bead to be included (allows the computation of the parameters of parts of models); f) the models are automatically "normalized" by a PAT routine (this is mandatory for the correct computation of the rotational diffusion coefficients); g) output of all the parameters is for standard conditions, e.g. water at 20 °C. The parameters calculated by the four methods and their units are reported in Table 1; h) in the case of multiple files, mean values and standard deviations of all the parameters are also computed; i) the "flag" in the coordinates' file can be used for two additional options: if set to "–1", the radii of the unhydrated beads must also be present in the "rmc" file, and are used for the computation of R_g; if the "flag" is set to "–2", the value of the partial specific volume $\bar{v}_2$ must be present after the "rmc" filename, and is used to calculate the sedimentation coefficient $s_0^{(20°,w)}$ via the Svedberg equation (see pp. 605–607 of Cantor and Schimmel 1980b); if the "flag" is set to "–3", both options are active.

In addition, some extra features are present only in SUPC: j) the volumetric correction for the computation of the rotational diffusion coefficients, and the mass of the model can be also entered manually (in GS, DIAG and GOLD, the total volume and the total mass are taken only from file as the sums of the volumes and of the masses of all the beads, respectively); k) a "colour filter" is provided. If selected, all the beads having a certain colour associated are excluded from the hydrodynamic computations (but not in the calculation of the radius of gyration). For instance, this is very useful when checking for the effects on the overall hydrodynamic parameters of "buried" beads in compact models.

Among the four methods, SUPC is the most precise and most computationally demanding. It calculates the forces by the direct inversion of the coefficient "supermatrix" $\mathcal{B}$ (Eqs. (6)–(8)), which is done by Cholesky decomposition (see pp. 96–98 of Press et al. 1992). Then the Ξ_t, $\Xi_{O,r}$ and $\Xi_{O,c}$ matrices are calculated (Eqs. (9)–(11)) and the vector $\mathbf{r}_{O,R}$ is evaluated from Eq. (9) of García Bernal and García de la Torre (1980). Next $\Xi_{R,r}$ is obtained from Eq. (15), and the diffusion matrix $\mathbf{D}_O$ is obtained from Eq. (3).

Table 1 Parameters calculated by the COEFF sub-programs, and their units

Parameter[a]	Units	SUPC	GS, DIAG and GOLD
Molecular weight	g mol^{-1}	+	+
Total volume	nm^3	+	+
Surface area	nm^2	+	–
R_g	nm	+	+
$f_t^{(w)}$	g s^{-1}	+	+
$D_t^{(20°,w)}$	cm^2 s^{-1}	+	+
R_S^t	nm	+	+
$s_0^{(20°,w)}$	S (10^{-13} s)	+	+
$f_r^{k(w)}$ (3×)	g cm^2 s^{-1}	+	+
$D_r^{k(20°,w)}$ (3×)	s^{-1}	+	+
$R_S^{r,k}$ (3×)	nm	+	+
$f_r^{(w)}$	g cm^2 s^{-1}	+	+
$D_r^{(20°,w)}$	s^{-1}	+	+
$\tau_{1-5}^{r(20°,w)}$ (5×)	ns	+	+
$\tau_h^{r(20°,w)}$	ns	+	+
$[\eta]$	cm^3 g^{-1}	+	–
$[\eta]^{corr}$	cm^3 g^{-1}	+	–
$[\eta]$ (DS)	cm^3 g^{-1}	+	+
$[\eta]$ (Tsuda CM)	cm^3 g^{-1}	+	+
$[\eta]$ (Tsuda CV)	cm^3 g^{-1}	+	–
R_E	nm	+	+
CR (x, y, z)	nm	+	+
CD (x, y, z)	nm	+	+
CM (x, y, z)	nm	+	+
CV (x, y, z)	nm	+	–
Max. Ext. (x, y, z)	nm	+	+
Axial ratios $x:y$ & $x:z$	/	+	+

[a] Symbols and abbreviations: $f_t^{(w)}$, $D_t^{(20°,w)}$ and R_S^t, translational frictional coefficient, diffusion coefficient and Stokes' radius, respectively; $s_0^{(20°,w)}$, sedimentatklkkjlnklnkljnnkkl,ion coefficient; $f_r^{k(w)}$, $D_r^{k(20°,w)}$ and $R_S^{r,k}$, rotational frictional coefficient, diffusion coefficient and Stokes' radii along the k (=x, y, z) axes, respectively; $f_r^{(w)}$ and $D_r^{(20°,w)}$, "global" rotational frictional and diffusion coefficients (mean of the individual values), respectively; $\tau_{1-5}^{r(20°,w)}$, five rotational relaxation times; $\tau_h^{r(20°,w)}$, unweighted harmonic mean of the five rotational relaxation times; $[\eta]$ and $[\eta]^{corr}$, intrinsic viscosity uncorrected and corrected, respectively; DS, double-sum; R_E, Einstein's radius (calculated for all the $[\eta]$ values); CR, CD, CM, and CV, centers of resistance, diffusion, mass and viscosity, respectively; Max. Ext., maximum extension along the individual axes. The superscripts (20°, w) and (w) indicate that the parameter is calculated at 20 °C and for the viscosity of water at 20 °C, or only for the latter, respectively

The various parameters are calculated from Eqs. (12)–(14) and (16)–(21), the centre of viscosity is evaluated from the system of equations defined by Eq. (23), and $[\eta]$ is obtained from Eq. (22). Two less precise methods for obtaining $[\eta]$, the double-sum (García de la Torre et al. 1983; see Eq. (30) in García de la Torre 1989) and the method of Tsuda (1970), are also implemented for comparison purposes.

GS utilizes an iterative approach, namely the well-known procedure of Gauss-Seidel (see Theory section), to get the forces from the coefficient "supermatrix", while DIAG utilizes the same method but with the diagonal approximation, e.g. it neglects the off-diagonal components of the hydrodynamic interaction tensor (García de la Torre and Bloomfield 1977a). The various parameters are then calculated as above except for [h], which in GS, DIAG and GOLD is calculated at the center of mass via the double-sum and Tsuda's methods (see above). The last subprogram, GOLD, is based on the variational method of Goldstein (1985), and the Cholesky method is again used to invert the "supermatrix".

The hydrodynamics of segmentally flexible ensembles of beads with two subunits with just one hinge or swivel point are computed by FLEX. FLEX uses the Harvey-Wegener treatment, hence all the caveats pointed out by García de la Torre (1994) apply. Its main applications, beside the generation of the sequentially flexed models, is in the calculation of the rotational relaxation times for each subunit, for comparison with those that can be derived from fluorescence anisotropy decay or electron spin resonance experiments when the fluorophore or spin label is present on only one subunit. The main features of FLEX are: a) choice between 7, 8 or 9 degrees of freedom (7 = bending, 8 = bending plus one torsion, 9 = bending plus two torsions; no restoring forces on bending and/or torsion are allowed, since they are outside the Harvey-Wegener treatment); b) inclusion or exclusion of hydrodynamic interactions between subunits during bending and/or torsion; c) choice of axis and direction of bending ($x+, x-, y+, y-$); d) setting between 0 and 1 of $a_{(k)}$ (Eqs. (27)–(29)); e) user-selectable initial and final angle, and bending step; f) user-selectable ratio between bending and torsion(s); g) inclusion or exclusion of the volume correction; h) choice between diffusion or origin as the center to which all the computations are referred to; i) computed parameters: translational frictional and diffusion coefficient, translational Stokes' radius, radius of gyration, five rotational relaxation times for each subunit, diagonal components of the rotational diffusion matrix ($D_{44}D_{55}D_{66}$).

The models need to be pre-aligned, for instance with the option present in GRAPH, before being passed to FLEX. After entering the starting and ending angles for bending and the relative torsional angles, the step angle and the direction of bending, the part of the model which is to remain fixed on the x-axis through the parameter $a_{(k)}$ (Eqs. (27)–(29)) must be defined. FLEX will then begin to generate the successive conformations for which it will compute the instantaneous hydrodynamic parameters as described in the Theory section. These conformations can be also saved in sequential files to be analyzed later with the rigid-body approximation, if wanted. The forces and torques are obtained by the direct inversion of the coefficient "supermatrix" utilizing the LU decomposition method with Crout's algorithm employed to find the lower and upper triangular matrices (see pp. 43–50 of Press et al. 1992). The instantaneous center of diffusion $\mathbf{D_D}$ is determined from $\mathbf{D_O}$ as described by Harvey et al. (1983), and the evaluation of the matrix $\mathbf{W}$ (Eq. (9) of Harvey et al. 1983) is done by approximating the partial derivatives with finite differentials. Then $D_t^{(T,\eta_0)}(\alpha)$ is obtained from Eq. (32), and Wegener's coefficients are calculated. The integration of the instantaneous values (Eq. (33)) with the proper probability distribution function (Eq. (34)) is done by either the rectangles or the Cavalieri-Simpson methods. Equations (35)–(37) are then applied to get the rotational relaxation times of each subunit.

The sequential files generated by FLEX can also be entered in another program, SUPFLEX, which calculates the hydrodynamic parameters for segmentally flexible models using the rigid-body approximation. The parameters are calculated for each sequential model using the same algorithms of SUPC, but at the end they are averaged according to the probability distribution functions. In absence of restoring forces, the average is done as seen for FLEX. We have already also implemented the case in which there is a linear restoring force opposing the bending but no linear torques opposing the torsions of the subunits. The probability distribution function is described by Eq. (39), in which the integral is approximated by the Cavalieri-Simpson method. The numerical values of $p(\alpha)$ for each value of α are then entered into the averaging functions for each parameter, whose integrals are again approximated by the Cavalieri-Simpson method.

The fifth BEAMS program, STRINGS, was the first program we actually developed. Its original core was derived from a FORTRAN program written by J. Hermans and M. Carson (Rocco et al. 1983) for the random generation of strings-of-beads of equal size and the computation of their translational frictional coefficient, but it has evolved into a much more powerful "generator" of bead models. The main features of STRINGS are: a) it accepts as input a file with beads coordinates (the same kind as the ones described above for GRUMB) for quick calculation of R_S^t (by Gauss-Seidel iterations with the diagonal approximation), R_g, and the end-to-end distance h_0; it can also produce a list of the distances between the centers and the surfaces of specified beads; b) alternatively, it accepts as input only the "rmc" file and generates at random a number of conformations (up to 200), and calculates the same parameters for each conformation. Conformations in which beads overlap are rejected (and not counted), and at the end the program gives the mean values and standard deviations of R_S^t, R_g and h_0. The randomness is achieved by random generation of the "bond" angles (between the lines connecting the centers of three successive beads) and the "dihedral" angles (defining the "rotation" on one bead with respect to the preceeding one around the line connecting their centers), and then calculating the coordinates of the beads. The degree of "compactness" (or "rigidity", so to speak) of the chain is controlled by a user-defined parameter k which works by restraining the "bond" angles from moving from the "rest" value of 180°, utilizing a harmonic potential in a Boltzmann distribution, e.g. for high k all the "bond" angles are 180° and the beads are all aligned, for low k all the angles which do not lead to bead overlap are accessible (for more details, see Rocco et al. 1983). "Dihedral" angles can always freely assume any value between 0° and 360°; c) the coordinates of every model or selected models generated (with no overlaps) can be automatically saved in a file for better analysis by the COEFF programs; d) the program can generate the model as a continuous chain (string-of-beads) or introduce branches of any length at selected points (more than one chain is allowed at any branch point); e) the user can select only part of the beads' coordinates to be generated at random, by providing a file with the "bond" and "dihedral" angles of the part(s) to be kept fixed. An ancillary program, ANGLES, calculates the angles from a previous bead model and stores them in a file. Again, the beads need not to be on a string, e.g., there can be "side chains". STRINGS can also directly save in a file the angles for each model it randomly generates; f) the user can select different *k* values for different parts of the chain, achieving different degrees of "compactness" for different parts of the chain; in this case, the *k* values need to be in another separate file; g) looped branches can be generated between any two points in the chain.

4. Results and discussion

The first problem in the bead modelling procedure is the definition of the number and size of the beads to be used. There are two main possibilities: presence or absence of high-resolution data (e.g. 3D-coordinates). Even if 3D-coordinates are available, the problem is not trivial, and a brief mention of a series of programs under development in our laboratory to tackle this problem will be made at the end of this section.

In the absence of high-resolution data, one needs to define each segment of the protein that has to be represented by a bead with a defined radius. The program PROMOLP is very useful for defining the number and size of beads starting from information generated by electron microscopy, differential scanning calorimetry, proteolytic digestion studies, or according to modular arrangement as revealed by sequence alignment, or gene duplication as indicated by intron/exon boundaries. One chooses parts of the sequence according to the information available and with the visual help of additional information, if present, displayed with the sequence (proteolytic cleavage sites, intron/exon boundaries), one obtains the volume, and thus the radius, of the corresponding bead, plus data about the net charge and hydrophobicity of that bead. In this way, there's a direct relationship between sequence and beads, allowing for interacting parts of a molecule to be placed close in space in bead models. Most importantly, the "hydration problem" is addressed in a "statistical" manner, by including in each bead the theoretically bound water of hydration calculated according to Kuntz and Kauzmann (1974).

The second problem in bead modelling procedures in the definition of the spatial position of the beads. Still in the absence of high-resolution data, two main options are possible, depending on the availability of independent information about the overall shape of the protein to be modelled: random generation or "manual" definition of the Cartesian coordinates of the beads. The programs STRINGS and GRUMB provide the user with a full spectrum of options that cover both possibilities and their combination.

Having defined number, size, and the spatial position of the beads, the next step is the computation of their solution parameters. BEAMS offers four options for "rigid" models and two for "segmentally flexible" models (see Fig. 1). The "rigid" models options are grouped under the COEFF sub-main, and can be chosen as a function of the number of the beads used, the precision needed, and the

Table 2 Performances of SUPC, GS and DIAG on various platforms[a]

Method	HP 9000	PC-Pentium	DEC 5100	Sun SPARC	Silicon IRIX
	Elapsed time for N = 300				
DIAG	50′ 02″	22″	1′ 49″	49″	30″
GS	1 h 00′ 10″	1′ 05″	3′ 56″	1′ 17″	50″
SUPC	6 h 24′ 45″	2′ 45″	39′ 00″	17′ 01″	5′ 45″
	Elapsed time for N = 400				
DIAG	1 h 54′ 09″	38″	3′ 26″	1′ 27″	54″
GS	2 h 19′ 39″	2′ 31″	7′ 56″	2′ 47″	1′ 39″
SUPC	/	6′ 29″	1 h 37′ 13″	39′ 52″	14′ 00″
	Elapsed time for N = 500				
DIAG	3 h 15′ 45″	1′ 00″	7′ 53″	2′ 39″	1′ 41″
GS	3 h 27′ 04″	3′ 33″	11′ 08″	4′ 20″	2′ 58″
SUPC	/	/	3 h 14′ 11″	1 h 17′ 39″	28′ 03″
	Elapsed time for N = 800				
DIAG	/	/	4 h 11′ 31″	8′ 22″	6′ 06″
GS	/	/	4 h 51′ 44″	15′ 15″	10′ 47″
SUPC	/	/	13 h 57′ 18″	5 h 33′ 21″	2 h 10′ 07″
	Elapsed time for N = 1000				
DIAG	/	/	7 h 17′ 18″	/	5 h 14′ 46″
GS	/	/	7 h 31′ 25″	6 h 24′ 56″	4 h 47′ 25″
SUPC	/	/	/	/	/

[a] Single run tests: times reported are only indicative

time required for the computation. In Table 2, the speed of execution for three of the four COEFF programs is reported for various platforms and for increasing numbers of beads. The enormous increase in computer power that can be made available to investigators in the average Institution now in comparison to only a decade ago, has pushed the limits of the practical, routine computation of the hydrodynamics of bead models to about 1,000 beads with the Gauss-Seidel methods, and to about 800 with the most precise method, the supermatrix inversion. For single structures for which there is no high-resolution information available, that is usually enough, and assemblies or oligomers can also be tackled reasonably. For example, some of us have recently published a model of integrin $\alpha_{IIb}\beta_3$ ($M_r \approx 230{,}000$) solubilized in octyl glucoside, in which 36 beads were employed (Rocco et al. 1993). This model is currently being used, for instance, to study $\alpha_{IIb}\beta_3$ oligomerization in solution: side-by-side vs. "rosette-style" complexes of dimers, trimers, and tetramers can be easily modelled and their hydrodynamic parameters computed and compared with ultracentrifugation and dynamic light scattering data, providing sound tests for current theories of integrin aggregation based mainly on EM images (Hantgan and Rocco, in progress). As another example, we are developing a bead model of fibrinogen ($M_r \approx 340{,}000$) consisting of 75 beads (Rocco et al. 1991); it will then become feasible to model the association of fibrin oligomers and $\alpha_{IIb}\beta_3$, its "natural" receptor on the surface of platelets.

However, if flexibility is an issue, the time required for analysis of the hydrodynamic properties increases dramatically: while there is no difference in the maximum number of beads that can be examined in respect to the rigid ensembles case, each successive configuration must be completely analyzed. For a bending of 30° in steps of 1°, that means a factor of ~30 to ~100 times more computer time, depending on whether torsions are allowed or not. This time doubles if the rigid-body approximation is then also utilized.

As a last point, we would like to discuss the possibility and the utility of generating accurate bead models from high-resolution three-dimensional structures. Apart from its employment in the verification of the solution behaviour of single proteins and the role of hydration on it (for instance, see Venable and Pastor 1988), this kind of modelling is very appealing for its potential applications in the structure determination of multidomain and/or modular proteins. In fact, the proteins belonging to this class, while usually being well outside the NMR range and often difficult to crystallize in intact form, lend themselves naturally to a strategy of structure resolution one piece at a time, where the pieces can be as small as single modules or can span up to few units. Strategies involving the combined use of small-angle X-ray or neutron scattering and electron microscopy data (Baron et al. 1991), with the addition of sedimentation velocity data and a simplified from of bead modeling (see Perkins 1994, and references therein), have already been proposed as a way to infer the overall structure of a multidomain/modular protein from those of its single modules/domains. However, we believe that a more sophisticated bead modelling procedure, which would also allow the computation of additional hydrodynamic parameters, such as the rotational relaxation times or the intrinsic viscosity, would greatly help in discriminating alternative structures. In this direction, we have developed and are refining a pair of programs, ASA and TRANS, that convert automatically a protein structure from the Protein Data Bank (PDB) format (Bernstein et al. 1977) into a bead model, utilizing as few non-overlapping beads as possible, and taking directly into account the theoretically bound water of hydration (Rocco et al., in preparation). Preliminary tests on small, well characterized proteins, such as BPTI, RNase A and lysozyme, have produced very encouraging results, with all the available parameters calculated within a few percent of the experimental ones. The reduction in bead number achieved with our strategy from that needed for an atomic description, is about a factor of seven, meaning in practice that a module of 10–14,000 mol.wt. could be represented with no more than 100–140 beads, and strategies for a further reduction without loss of accuracy are being explored.

Although recently there has appeared in the literature alternative, more direct methods for the computation of the hydrodynamic properties of proteins from their three-dimensional structures (Brune and Kim 1993; Antosiewicz 1995; Zhou 1995 a,b), we believe that the versatility of the bead modelling procedure, with its capability of mixing together models based on high-resolution and low-resolution data, will still be well positioned to play an important role in the elucidation of biomacromolecular structures, dynamics and interactions. In this respect, we hope that the BEAMS suite of modelling programs will favour its further diffusion among the biophysical and structural biology communities.

Acknowledgements Part of this work was supported by a grant from the EEC program "Biomed II" to C. R. (# BMH4 CT95 0396), by grants from the Italian Space Agency (ASI) to M. R., and by a grant from the Associazione Italiana per la Ricerca sul Cancro (AIRC) to Prof. M. Bolognesi, to whom we are grateful. The development of BEAMS over the years was also made possible by the work of, among others, S. Fiore, C. De Maestri and F. Cauglia. We also thank D. Scanarotti for the installation and performance verification of BEAMS on various platforms. The BEAMS and GRUMB source codes and the PROMOLP compiled version for either Windows 3.1x™ or Windows 95™ are available free on request from M. R. (rocco@ermes.cba.unige.it) or C. R. (carmel@dist.unige.it), and they will be soon also made available for direct downloading from a www page (http://wwwz.cba.unige.it/MacroHydro).

References

Antosiewicz J (1995) Computation of the dipole moments of proteins. Biophys J 69:1344–1354

Baron M, Norman DG, Campbell ID (1991) Protein modules. Trends Biochem Sci 16:13–17

Bernstein F, Koetzle T, Williams G, Meyer E Jr, Brice M, Rodgers J, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 112:535–542

Bloomfield VA, Dalton WO, Van Holde KE (1967a) Frictional coefficients of multisubunit structures. I. Theory. Biopolymers 5:135–148

Bloomfield VA, Van Holde KE, Dalton WO (1967b) Frictional coefficients of multisubunit structures. II. Application to proteins and viruses. Biopolymers 5:149–159

Bork P, Downing AK, Kieffer B, Campbell ID (1996) Structure and distribution of modules in extracellular proteins. Q Rev Biophys 29:119–167

Brenner H (1967) Coupling between the translational and rotational Brownian motions of rigid particles of arbitrary shape. J Colloid Interface Sci 23:407–436

Brune D, Kim S (1993) Predicting protein diffusion coefficients. Proc Natl Acad Sci USA 90:3835–3839

Cantor CR, Schimmel PR (1980a) Biophysical chemistry. Part I: The conformation of biological macromolecules. WH Freeman and Co, San Francisco, CA, USA

Cantor CR, Schimmel PR (1980b) Biophysical chemistry. Part II: Techniques for the study of biological structure and function. WH Freeman and Co, San Francisco, CA, USA

Doolittle RF (1995) The multiplicity of domains in proteins. Annu Rev Biochem 64:287–314

García Bernal JM, García de la Torre J (1980) Transport properties and hydrodynamic centers of rigid macromolecules with arbitrary shapes. Biopolymers 19:751–766

García de la Torre J (1989) Hydrodynamic properties of macromolecular assemblies. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. The Royal Society of Chemistry Special Publication No 74, Cambridge, UK, pp 3–31

García de la Torre J (1994) Hydrodynamics of segmentally flexible macromolecules. Eur Biophys J 23:307–322

García de la Torre J, Bloomfield VA (1977a) Hydrodynamic properties of macromolecular complexes. I. Translation. Biopolymers 16:1747–1763

García de la Torre J, Bloomfield VA (1977b) Hydrodynamics of macromolecular complexes. II. Rotation. Biopolymers 16:1765–1778

García de la Torre J, Bloomfield VA (1977c) Hydrodynamics of macromolecular complexes. III. Bacterial viruses. Biopolymers 16:1779–1793

García de la Torre J, Bloomfield VA (1978) Hydrodynamic properties of macromolecular complexes. IV. Intrinsic viscosity theory, with applications to once-broken rods and multisubunit proteins. Biopolymers 17:1605–1627

García de la Torre J, Bloomfield VA (1981) Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. Q Rev Biophys 14:81–139

García de la Torre J, Rodes V (1983) Effects from bead size and hydrodynamic interactions on the translational and rotational coefficients of macromolecular bead models. J Chem Phys 79:2454–2460

García de la Torre J, Lopez MC, Tirado MM, Freire JJ (1983) Approximate methods for calculating hydrodynamic properties of macromolecules in dilute solution. Theory and application to rigid structures. Macromolecules 16:1121–1127

García de la Torre J, Navarro S, Martinez MCL, Diaz FG, Cascales JJL (1994) HYDRO: A computer program for the prediction of hydrodynamic properties of macromolecules. Biophys J 67:530–531

Goldstein RF (1950) Classical mechanics. Addison-Wesley. Chapter 5, pp 143–163

Goldstein RF (1985) Macromolecular diffusion constants: A calculational strategy. J Chem Phys 83:2390–2397

Happel J, Brenner H (1973) Low Reynolds number hydrodynamics, ch 5. Nordhoff, Leyden

Harding SE (1989) Modelling the gross conformation of assemblies using hydrodynamics: the whole body approach. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. The Royal Society of Chemistry Special Publication No 74, Cambridge, UK, pp 32–56

Harding SE (1997) The ELLIPS suite of macromolecular conformation PC algorithms. Eur Biophys J 25:347–359

Harvey SC (1979) Transport properties of particles with segmental flexibility. I. Hydrodynamic resistance and diffusion coefficients of a freely hinged particle. Biopolymers 18:1081–1104

Harvey SC, Mellado P, García de la Torre J (1983) Hydrodynamic resistance and diffusion coefficients of segmentally flexible macromolecules with two subunits. J Chem Phys 78:2081–2090

Iniesta A, Diaz FG, García de la Torre J (1988) Transport properties of rigid bent-rod macromolecules and of semiflexible broken rods in the rigid-body treatment. Analysis of the flexibility of myosin rod. Biophys J 54:269–275

Kuntz ID, Kauzmann W (1974) Hydration of proteins and polypeptides. Adv Prot Chem 28:239–345

Marshall AG (1978) Biophysical chemistry. Principles, techniques and applications. Wiley, New York, NY, USA

Mellado P, Iniesta A, Diaz FG, García de la Torre J (1988) Diffusion coefficients of segmentally flexible macromolecules with two subunits: a study of broken rods. Biopolymers 27:1771–1786

Perkins SJ (1994) High-flux X-ray and neutron solution scattering. In: Jones C, Mulloy B, Thomas AH (eds) Microscopy, optical spectroscopy, and macroscopic techniques. Methods in molecular biology, vol. 22. Humana Press, Totowa, NJ, pp 39–60

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in C, 2nd edn. Cambridge University Press, UK

Rocco M, Cabib D, Pioli F, Spotorno B, Molina F, Cuniberti C (1991) An hydrodynamic model of fibrinogen fragment X. Thromb Haemost (abstract) 65:898

Rocco M, Carson M, Hantgan R, McDonagh J, Hermans J (1983) Dependence of the shape of the plasma fibronectin molecule on solvent composition. Ionic strength and glycerol content. J Biol Chem 258:14545–14549

Rocco M, Spotorno B, Hantgan RR (1993) Modeling the $\alpha_{\mathrm{IIb}}\beta_3$ integrin solution conformation. Protein Sci 2:2154–2166

Rodriguez-Tomé P, Stoehr PJ, Cameron GN, Flores TP (1996) The European Informatics Institute (EBI) databases. Nucleic Acids Res 24:6–12

Rotne J, Prager S (1969) Variational treatment of hydrodynamic interaction on polymers. J Chem Phys 50:4831–4837

Ruggiero C, Piccinini L, Spotorno B, Tassara G, Molina F, Rocco M (1990) A set of computer programs for the construction, the evaluation and the representation of hydrodynamic models of proteins. Proc 12th Annu Int Conf Eng Med Biol Soc 7:2–5

Teller DC, Swanson E, De Haën C (1979) The translational friction coefficient of proteins. Meth Enzymol 61:103–124

Tsuda K (1970) Hydrodynamic properties of rigid complex molecules. Polym J 1:616–631

Venable RM, Pastor RW (1988) Frictional models for stochastic simulations of proteins. Biopolymers 27:1001–1014

Wegener WA, Dowben RM, Koester VJ (1979) Time-dependent birefringence, linear dichroism, and optical rotation resulting from rigid-body rotational diffusion. J Chem Phys 70:622–632

Wegener WA, Dowben RM, Koester VJ (1980) Diffusion coefficients for segmentally flexible macromolecules: General formalism and application to rotational behavior of a body with two segments. J Chem Phys 73:4086–4097

Wegener WA (1981) Diffusion coefficients for rigid macromolecules with irregular shapes that allow rotational-translational coupling. Biopolymers 20:303–326

Wegener WA (1982a) Bead models of segmentally flexible macromolecules. J Chem Phys 76:6425–6430

Wegener WA (1982b) A swivel-jointed formalism for segmentally flexible macromolecules and its application to the rotational behavior of myosin. Biopolymers 21:1049–1080

Wegener WA (1985) Center of diffusion of flexible macromolecules. Macromolecules 18:2522–2530

Yamakawa H (1970) Transport properties of polymer chains in dilute solutions. Hydrodynamic interaction. J Chem Phys 53:436–443

Yguerabide J, Epstein HF, Steyer L (1970) Segmental flexibility in an antibody molecule. J Mol Biol 51:573–590

Zhou HX (1995a) Calculation of translational friction and intrinsic viscosity. 1. General formulation for arbitrarily shaped particles. Biophys J 69:2286–2297

Zhou HX (1995b) Calculation of translational friction and intrinsic viscosity. 2. Application to globular proteins. Biophys J 69:2298–2303

Eur Biophys J (1997) 26: 417

ERRATUM

Bruno Spotorno · Luca Piccinini · Giovanni Tassara
Carmelina Ruggiero · Marco Nardini
Francesco Molina · Mattia Rocco

BEAMS (BEAds Modelling System): a set of computer programs for the generation, the visualization and the computation of the hydrodynamic and conformational properties of bead models of proteins

Eur Biophys J (1997) 25: 373–384

The above article unfortunately contained three errors:

1. The e-mail address on the title page should read rocco@vega.cba.unige.it

2. The Internet site in the Acknowledgements should be www2.cba.unige.it/MacroHydro

3. The correct Equation (25) should read

$$R_E = \left[\frac{3M[\eta]}{10\pi N_A} \right]^{\frac{1}{3}}$$

We apologize for these mistakes.