

## Advances in modelling solution properties of macromolecules and particles

J. García de la Torre<sup>\*1</sup>, S. E. Harding<sup>†</sup> and B. Carrasco<sup>\*</sup>

<sup>\*</sup>Departamento de Química Física, Universidad de Murcia, 30071 Murcia, Spain, and <sup>†</sup>Department of Applied Biochemistry and Food Science, University of Nottingham, Sutton Bonington LE12 5RD, U.K.

716

### Introduction to bead modelling

The size and detailed shape of rigid macromolecules or particles can be inferred from the properties of their solutions or suspensions. Some of these are hydrodynamic properties, such as the sedimentation coefficient,  $s$ , the translational diffusion coefficient,  $D_t$ , rotational diffusion coefficients,  $D_r$ , and relaxation times,  $\tau$ , and the intrinsic viscosity  $[\eta]$ . There are also equilibrium solution properties, including the angular dependence of radiation scattering,  $P(\theta)$ , the radius of gyration,  $R_g$  and the covolume,  $u$ . Theoretical formalisms, implemented in adequate computer procedures, are required to correlate global structure (size and shape) with solution properties. In some instances, the overall shape can be conveniently described by revolution or triaxial ellipsoids, for which theory and simple procedures are available [1,2], but in many other cases the ellipsoidal shape is clearly inadequate to model the very specific shape of biological macromolecules. For such cases, Bloomfield et al. [3,4] designed the so-called bead models, in which the shape of a rigid particle is represented by an array of spherical elements. By using a sufficient number of beads of (possibly) varying size, the intricate details of macromolecular shape can be deduced. Two illustrative examples are displayed in Figure 1.

The theory on which bead modelling is based comes from the pioneering works on macromolecular hydrodynamics by Kirkwood and Riseman [5,6], of which Bloomfield et al. adopted an approximate, simplified form. Garcia de la Torre, Bloomfield and co-workers improved the theoretical treatment and took advantage of the increasing availability of computing power, developing methods for efficient and accurate calculation of hydrodynamic properties of bead models of arbitrary complexity [7–13]. A few years ago the computer program HYDRO was published [14] as a useful tool to calculate some hydrodynamic properties of bead models.

Abbreviations used: DNMR, dynamic NMR; EBD, time decay of electric birefringence; FAD, fluorescence anisotropy decay.

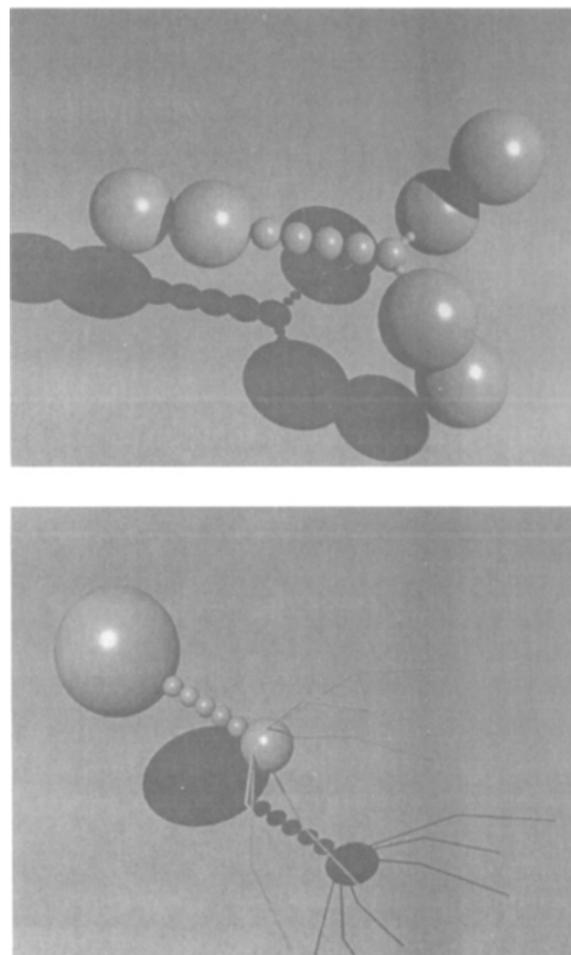
<sup>1</sup>To whom correspondence should be addressed.

It was soon evident that the utility of bead modelling could be pushed further in various directions. (1) Prediction of complex, more involved solution properties, usually related to rotational brownian diffusion of the rigid solute particles, that are monitored in some electro-optical and spectroscopic techniques. (2) Calculation of compound quantities involving two solution properties in such a way that the combination depends on the shape but not on the size of the particles. Such quantities had been

Figure 1

#### Examples of bead models

Upper panel: model for the human immunoglobulin IgG3, according to Gregory et al. [31]. Lower panel: model for a T-even bacteriophage, from García de la Torre and Bloomfield [32].



used customarily with ellipsoidal models but not with bead models. (3) Extension to other, non-hydrodynamic properties, such as those related to radiation scattering (angular variation, distribution of distances, radius of gyration) and to concentration dependences (covolume and virial coefficient). In recent work we have described the theory of bead-modelling calculation of such properties, and a new computer program, SOLPRO [15,16].

Here we focus on the computer procedures HYDRO and SOLPRO, summarizing their main features, possible applications and availability.

### HYDRO: simple, 'single-valued', hydrodynamic properties

The computer program HYDRO implements all the theoretical developments in the hydrodynamics of bead models. The input for HYDRO is simple: it just consists of a list of the Cartesian coordinates and radii of the beads in the model, complemented by a list of basic data such as temperature, and the viscosity and density of the solvent. The output from HYDRO is also simple; basically, it contains the numerical

results for the sedimentation and translational friction coefficients,  $s$  and  $D_i$ , the (up to five) rotational relaxation times  $\tau_i$ ,  $i, \dots, 5$  and the intrinsic viscosity  $\eta$ . Examples of input and output files are presented in Figure 2. We say that these properties are 'single-valued', in contrast with more involved solution properties that are functions of time, scattering angle or whatever (see below).

In HYDRO, we have implemented the so-called volume correction, which attempts to remove a deficiency in the Kirkwood-Riseman theory, which was present in the earliest treatment of bead models. Recently we have discussed the theoretical basis of this correction [17]. The correction improves the results more or less when it is applied to what we call 'bead models in the strict sense', such as those in Figure 1. However, the correction produces incorrect results in the case of what we call 'filling models', in which many (hundreds or even thousands) small beads are densely packed, filling the volume occupied by the macromolecule. Such models are used by workers who attempt to predict the hydrodynamic properties of biopolymers (mainly globular proteins) from the

Figure 2

#### HYDRO: input and output files

(A) Input file for the HYDRO calculation for the IgG3 model; (B) results from HYDRO for the 'single-valued' properties.

IMMUNOGLOBULIN IGG3	!Title	(A)
298.,	!Temperature, K	
0.01,	!Solvent viscosity, poise (water, 20C)	
1.E-07,	!Unit of length for coordinates and radii, cm (1 A)	
1.58E5,	!Molecular weight (-1=unknown)	
0.239,	!Bouyancy factor for sedimentation (-1=unknown)	
15,	!Number of beads	
-12.9000	0.0000 0.0000 2.05	
-8.8	0. 0. 2.05	
.....		
1.9	0. 3.2909 2.05	
3.95	3.5507 3.2909 2.05	
		(B)
	Number of beads 15	
Translational diffusion coefficient	3.821E-07 cm <sup>2</sup> /s	
Radius of gyration	7.492E-07 cm	
Volume	2.256E-19 cm <sup>3</sup>	
Relaxation time (1)	4.321E-07 s	
Relaxation time (2)	3.445E-07 s	
Relaxation time (3)	3.287E-07 s	
Relaxation time (4)	2.024E-07 s	
Relaxation time (5)	2.023E-07 s	
Harmonic relaxation time	2.756E-07 s	
Intrinsic viscosity	9.823E+00 cm <sup>3</sup> /g	
Sedimentation coefficient	5.826E+00 svedberg	

crystallographic atomic coordinates in the Brookhaven Protein Data Bank or similar files [18,19]. We have discussed this problem elsewhere ([20]; B. Carrasco and J. García de la Torre, unpublished work). In the latest version of HYDRO the volume correction is optional.

We have also recently included in HYDRO a provision for models with overlapping beads. In principle we do not advise the use of such models because in the case of bead overlapping, the hydrodynamic theory is not well developed. Anyway, some authors employ models (again, mainly from atomic coordinates of proteins) with a considerable amount of bead overlap. To avoid serious failures of HYDRO, an *ad hoc* correction has been implemented (see [19,20]; B. Carrasco and J. García de la Torre, unpublished work).

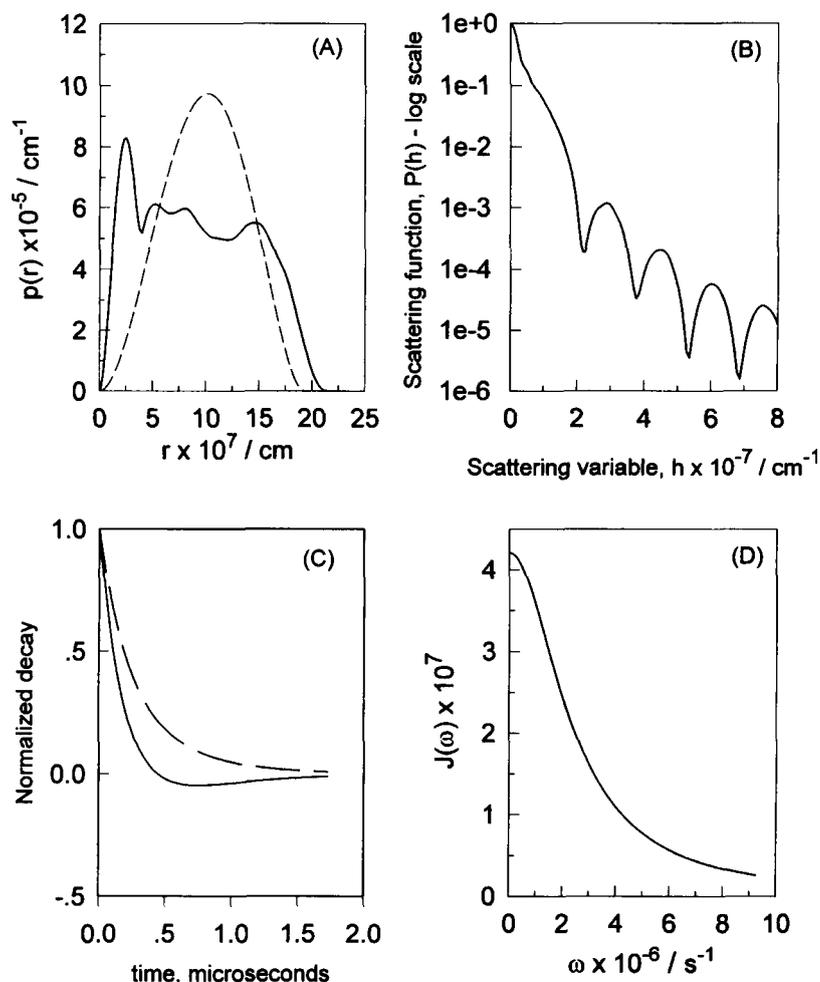
**SOLPRO: properties related to rotational dynamics**

As mentioned above, the dynamics of the rigid macromolecule in solution, particularly the rotational dynamics, is usually described not by a single value of some molecular coefficient, but by some quantity that is a function of time or frequency. Such quantities depend also on some electro-optical or spectroscopic properties of the macromolecule. Examples are the transient electric birefringence (more specifically, the time decay of electric birefringence, EBD), given by the time-dependence of birefringence  $\delta n(t)$ ; the fluorescence anisotropy decay, FAD, given by the time-dependence of the anisotropy,  $r(t)$ ; and the dynamic nuclear magnetic resonance, DNMR, where the relaxation times are calculated by the

**Figure 3**

Results from SOLPRO calculations for the IgG<sub>3</sub> model

(A) Distribution of distances; (B) wide-angle scattering form factor; (C) decay of electric birefringence (continuous line) and fluorescence anisotropy (broken line); (D) spectral density function in NMR.



spectral density function,  $\mathcal{Y}(\omega)$ , which is a function of frequency,  $\omega$ , calculated from the Fourier transform of another time-dependent correlation function  $P_{\frac{1}{2}}^{(\text{NMR})}(t)$ .

For an arbitrary particle there are up to five rotational relaxation times [21], which are calculated by HYDRO. The functions  $\delta n(t)$ ,  $r(t)$  and  $P_{\frac{1}{2}}^{(\text{NMR})}(t)$  are all functions of time ( $f(t)$ ) having as a generic form a sum of up to five negative exponentials, each containing one of the relaxation times,  $\tau_i$ :

$$f(t) = \sum_{k=1}^5 a_k \exp(-t/\tau_k) \quad (1)$$

where the amplitudes  $a_k$  are specific for the technique, FAD, EBD or DNMR. The additional information required for calculating the  $a_k$  in these functions are: for EBD, the dipole moment and the electric and optical polarizabilities [22]; for EBD, the orientation of the absorption and emission dipoles in fluorescence [23]; for DNMR, the orientation of the dipole (atom-to-atom) vector corresponding to the pair of nuclei whose coupling is being monitored [24,25].

If the electro-optical and spectroscopic quantities are available, SOLPRO can compute the time or frequency functions. It also gives a useful quantity that characterizes the overall rotational dynamics: the mean relaxation time  $\tau_{\text{mean}}$ , also called correlation time,  $\tau_c$ , which given in general is:

$$\tau_{\text{mean}} = \sum_{k=1}^5 a_k \tau_k. \quad (2)$$

For a rigid macromolecule with a given size and shape, the set of the (up to five)  $\tau_k$  values is fixed; however, the  $\tau_{\text{mean}}$  values for FAD, EBD and DNMR are different. Other mean times, namely the initial mean and harmonic mean relaxation times, can be defined [15].

### SOLPRO: other solution properties

Bead models were initially intended for the calculation of hydrodynamic properties, such as  $s$ ,  $D_t$ ,  $\tau$  values, which are basically the properties calculated by HYDRO. However, other properties can be more or less easily calculated from the bead model. A most trivial property is the particle's volume  $V$ , simply calculated as the sum of the volumes of the spherical beads. Another simple quantity is the radius of gyration,  $R_g$ , which can be easily calculated for a rigid array of spheres [15,26]. Experimentally,  $R_g$  is deter-

mined from the low-angle part of the angular dependence of radiation scattering, given by the scattering form factor,  $P(h)$ , where  $h$  is the angular variable containing the scattering half-angle. The calculation of  $P(h)$  for a bead model is feasible [15,27] and very interesting, for this angular function embodies much more structural information than the 'single-valued' property  $R_g$ . In X-ray scattering it is a common practice to transform  $P(h)$  into the so-called distance distribution function  $p(r)$  [27], which expresses the distribution of distances,  $r$ , between any pair of points within the particle. We have found analytical expressions that allow the calculation of  $p(r)$  for an arbitrary bead model [16]. The range of  $r$  goes from zero to a quantity  $L$ , which is the maximum dimension of the particle and serves as another useful indicator of the macromolecular structure;  $L$  can be obtained quite trivially. Examples of the various functions calculated by SOLPRO are given in Figure 3.

Other important solution properties refer to the concentration dependence of scattering intensity, sedimentation or diffusion coefficients, and so on. Such dependences are usually expressed in the form of the virial coefficient, which in turn can be formulated in terms of a molecular property, the covolume  $u$ . This property measures the volume excluded by a fixed particle to a second particle, averaged over orientations. The determination of  $u$  for an arbitrary shape is difficult: the concept is simple but there is no immediate result. We have developed a Monte Carlo procedure for the determination of  $u$  in a general case, and have implemented the procedure when the particle is a rigid array of beads [16]. Even for very elongated structures, for which the efficiency is low, we have seen that accurate results can be achieved in a very short computing time.

As described elsewhere [15,16], all these non-hydrodynamic properties  $V$ ,  $R_g$ ,  $P(h)$ ,  $p(r)$ ,  $L$  and  $u$  can be calculated with the SOLPRO computer program, passing to it the same bead model (list of coordinates and radii of spheres) used in the HYDRO calculation of the hydrodynamic quantities.

### SOLPRO: dimensionless, shape-dependent compound quantities

From the early days in macromolecular science, it has been known that some combinations of two or more properties can be formulated in such a way that the resulting compound quantity is

dimensionless and depends on the conformation or shape of the macromolecule but not on its size. Because of this they are sometimes called 'universal functions'. The combined properties may be any equilibrium and hydrodynamic properties in solution, the molecular volume and covolume, and so on. Well-known examples are the Perrin function,  $P$ , which combines the translational friction coefficient,  $f$ , and the molecular volume,  $V$  [28]:

$$P = \frac{f}{6\pi\eta_0(3V/4\pi)^{1/3}} \quad (3)$$

and the Scheraga–Mandelkern parameter,  $\beta$ , which combines  $f$  with the intrinsic viscosity  $[\eta]$  and the molecular weight,  $M_r$ :

$$\beta \equiv \frac{N_A s [\eta]^{1/3} \eta_0}{M_r^{2/3} (1 - \bar{v} \rho_0) 100^{1/3}} \quad (4)$$

where  $\eta_0$  and  $\rho_0$  are the viscosity and density of the solvent respectively,  $\bar{v}$  is the specific partial volume,  $N_A$  is the Avogadro number and  $s$  is the sedimentation coefficient.

For ellipsoids of revolution (prolate and oblate), such combinations are functions of the axial ratio and have been widely employed to determine the shape of rigid biological macromolecules. In more recent years, more shape-dependent functions have been formulated, and their applicability was extended to triaxial ellipsoids [1,2,29].

The possibility of using the size-independent, shape-dependent functions is not exclusive to ellipsoids. If two rigid particles have the same shape but differ in size (more specifically, if one can be derived from the other by a uniform expansion) then they must have the same values for such functions, which can therefore be used as shape indicators for particles of any geometry. Thus in SOLPRO we have programmed the various combinations of the hydrodynamic properties resulting from HYDRO and the other solution properties related to scattering and covolume obtained in SOLPRO itself. We have included shape-dependent functions taken from the domain of flexible macromolecules, the Flory functions  $P_0$  and  $\Phi_0$  [30], which combine  $f$  and  $[\eta]$  respectively with the radius of gyration  $R_g$ . It is noteworthy that each of the shape functions involving rotation admits a number of subcases: one for each of the five  $\tau_i$  values and one for each mean relaxation time, which, as commented

above, can in turn be formulated for each of the instrumental techniques EBD, FAD or NMR.

Having included in SOLPRO all those possible combinations, the list of the computed shape-dependent compound quantities is too long to be reproduced here. The interested reader can find a complete description of all of them in our more specialized publications [15,16].

### Computer programs

HYDRO and SOLPRO are public-domain programs, with the form of FORTRAN subroutines which will be linked to the user's main program. For users who wish to avoid programming, or if a FORTRAN compiler is not available, we also supply executables for various platforms. All the program files and further information can be found in the World Wide Web site <http://leonardo.fcu.um.es/macromol>.

This work was supported by grant PB96-1106 from Dirección General de Enseñanza Superior, M.E.C. (Spain). B.C. is the recipient of a predoctoral fellowship from the same source. Support from the B.B.S.R.C. and E.P.S.R.C. (U.K.) is also acknowledged.

- 1 Harding, S. E. (1995) *Biophys. Chem.* **55**, 69–93
- 2 Harding, S. E., Horton, J. C. and Cölfen, H. (1997) *Eur. Biophys. J.* **25**, 347–360
- 3 Bloomfield, V. A., Dalton, W. O. and Van Holde, K. E. (1967) *Biopolymers* **5**, 135–148
- 4 Bloomfield, V. A., Dalton, W. O. and Van Holde, K. E. (1967) *Biopolymers* **5**, 149–159
- 5 Kirkwood, J. G. and Riseman, J. (1948) *J. Chem. Phys.* **16**, 565–573
- 6 Riseman, J. and Kirkwood, J. G. (1949) *J. Chem. Phys.* **17**, 442–446
- 7 García de la Torre, J. and Bloomfield, V. A. (1977) *Biopolymers* **16**, 1747–1763
- 8 García de la Torre, J. and Bloomfield, V. A. (1978) *Biopolymers* **17**, 1605–1627
- 9 Wilson, R. W. and Bloomfield, V. A. (1979) *Biopolymers* **18**, 1205–1211
- 10 García Bernal, J. M. and García de la Torre, J. (1980) *Biopolymers* **19**, 751–766
- 11 García de la Torre, J. and Bloomfield, V. A. (1981) *Quart. Rev. Biophys.* **14**, 81–139
- 12 García de la Torre, J. and Rodes, V. (1983) *J. Chem. Phys.* **79**, 2454–2460
- 13 García de la Torre, J. (1989) in *Dynamic Properties of Biomolecular Assemblies* (Harding, S. E. and Rowe, A. J., eds.), pp. 3–31, Royal Society of Chemistry, Cambridge

- 14 García de la Torre, J., Navarro, S., López Martínez, M. C., Díaz, F. G. and López Cascales, J. J. (1994) *Biophys. J.* **67**, 530–531
- 15 García de la Torre, J., Carrasco, B. and Harding, S. E. (1997) *Eur. Biophys. J.* **25**, 361–372
- 16 García de la Torre, J., Harding, S. E. and Carrasco, B. (1998) *Eur. Biophys. J.*, in the press
- 17 García de la Torre, J. and Carrasco, B. (1998) *Eur. Biophys. J.* **27**, 549–557
- 18 Byron, O. (1997) *Biophys. J.* **72**, 408–415
- 19 Zipper, P. (1998) Presented at the 5th UK Analytical Ultracentrifuge Users Meeting, University of Nottingham, Sutton Bonington, UK
- 20 Carrasco, B. (1998) Ph.D. thesis, Universidad de Murcia
- 21 Favro, L. D. (1960) *Phys. Rev.* **119**, 53–62
- 22 Wegner, W. A., Dowben, R. M. and Koester, V. J. (1979) *J. Chem. Phys.* **70**, 622–632
- 23 Belford, G. G., Belford, R. L. and Weber, G. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1392–1393
- 24 Woessner, D. E. (1962) *J. Chem. Phys.* **40**, 647–654
- 25 Huntress, W. T. (1968) *J. Chem. Phys.* **48**, 3524–3533
- 26 Sólvez, A., Iniesta, A. and García de la Torre, J. (1988) *Int. J. Biol. Macromol.* **9**, 39–43
- 27 Glatter, O. and Kratky, O. (1982) *Small Angle X-ray Scattering*, Academic Press, New York
- 28 Perrin, F. (1936) *J. Phys. Radium* **7**, 1–11
- 29 Harding, S. E. (1989) in *Dynamic Properties of Biomolecular Assemblies* (Harding, S. E. and Rowe, A. J., eds.), pp. 32–35, Royal Society of Chemistry, Cambridge
- 30 Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, NY
- 31 Gregory, L., Davis, K. S., Sheth, B., Boyd, J., Jefferis, R., Naves, C. and Burton, D. (1987) *Mol. Immunol.* **24**, 821–829
- 32 García de la Torre, J. and Bloomfield, V. A. (1977) *Biopolymers* **16**, 1779–1773

Received 29 June 1998

### Models for the multisubunit conformation of oil-seed globulins

B. Carrasco\*, J. García de la Torre\*<sup>1</sup> and S. E. Harding<sup>†1</sup>

\*Departamento de Química Física, Universidad de Murcia, 30071 Murcia, Spain, and <sup>†</sup>Department of Applied Biochemistry and Food Science, University of Nottingham, Sutton Bonington LE12 5RD, U.K.

#### Introduction

Seed globulins provide a nitrogen reserve for the plant [1,2], and have a considerable commercial interest for food use. From a structural point of view, these globulins are an interesting example of multi-subunit proteins for which the determination of the spatial arrangement of the subunits can be attempted from measurements and modelling of solution properties.

The native structure of sunflower and rapeseed globulin is the so-called 11 S form, despite the fact that the precise value of their sedimentation coefficients is approx. 12 S at 20.0 °C. The 11 S globulins dissociate into subunits according to the following scheme [3–5]:



where the last step requires much stronger dissociating agents (6 M guanidinium chloride or 8 M urea) and so to a reasonable extent the 11 S sunflower and rapeseed globulins can be regarded as being made up of six subunits [5].

A detailed study of the conformation of these proteins has been undertaken by Plietz [5],

who used small-angle X-ray scattering to show the most likely arrangement of the subunits as a trigonal bipyramid. It would be most interesting to carry out the structural determination from other solution properties, which have been studied in detail by K. D. Schwenke and co-workers [4,6] and V. Prakash and co-workers [1,2]. A summary of published values is presented in Table 1.

Oligomeric proteins, composed of moderate number of subunits arranged in a rather specific manner, with a polygonal or polyhedral geometry, are good examples of biological macromolecules whose solution properties cannot be described by simple ellipsoidal models. Instead, this important class of structures was among those that motivated the development of bead models for solution properties [7,8], for which the existing theory [9,10] has been implemented in the HYDRO computer program [11]. Indeed, oligomeric structures have, over the years, been the subject of bead modelling work [9,10,12–14].

From the early studies based on ellipsoidal shapes, it is known that it is possible to model the shape of the particle directly without ambiguities induced by having to include particle size as well [15]. This is done by combining two or

<sup>1</sup>To whom correspondence should be addressed.