

Hydrodynamic Properties of Biomacromolecules and Macromolecular Complexes: Concepts and Methods. A Tutorial Mini-review

J. García de la Torre and J.G. Hernández Cifre

Departamento de Química Física, Facultad de Química, Universidad de Murcia, 30071 Murcia, Spain

Correspondence to J. García de la Torre: Fax: +34 968 364148. jgt@um.es

<https://doi.org/10.1016/j.jmb.2019.12.027>

Edited by Pau Bernadó

Abstract

This article is intended to be an overview of the current possibilities offered by hydrodynamic methods in the calculation of properties and prediction of the behavior of biomacromolecules in solution. First, we briefly cover the experimental techniques, mentioning their fundamentals and current status. Using a tutorial approach, we provide basic hints to understand conceptual aspects of macromolecular hydrodynamics which underlie the instrumental methods and the modeling and computational procedures. The description is focused on the bead model methodology as implemented in the HYDRO suite of computer programs. For rigid particles, we cover simple models, such as ellipsoids and cylinders, to the most detailed models with atomic resolution. The fundamentals and applicability of the basic random coil and wormlike models for flexible particles are also briefly described. In addition to the simple linear, more or less flexible chain, flexibility may appear in other more specific forms. These cases can be treated by Monte Carlo and Brownian simulation methods, for which computational tools are also available. Finally, we present in some detail the applicability of these tools for unfolded and intrinsically disordered proteins. For the particular case of partially disordered proteins comprising both globular domains and flexible linkers or tails, solution properties can be accurately predicted, and this validation makes the methodology quite promising for future work.

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

For over a century, hydrodynamic properties have been essential sources of information about the structure of biological macromolecules. Historically, the first landmark was the doctoral thesis of Albert Einstein, who developed the theory of Brownian diffusion and carried out its first application to the determination of the size of a biomolecule – sucrose [1,2]. His pioneering work prompted the invention by T. Svedberg of analytical ultracentrifugation (AUC), which made it possible to demonstrate that a biological macromolecule, hemoglobin, is a single species with a well-defined molecular weight, which was determined by observing its sedimentation in the ultracentrifuge [3]. In the early years of modern molecular biology, AUC was used in the famous Meselson–Stahl experiment [4], demonstrating the semiconservative replication of DNA, and thus supporting the Watson-Crick double helix model. A

particular mention is relevant here to the very first paper published in the *Journal of Molecular Biology*, in which Zubay and Doty [5] described the characterization of nucleohistones by hydrodynamic methods, namely intrinsic viscosity and flow birefringence.

For many years, hydrodynamic properties such as the diffusion and sedimentation coefficients and the intrinsic viscosity were the main choice (along with light scattering) to ascertain structural information about size, overall shape, flexibility, etc., of biomacromolecules. During the 1980s and 1990s, the emergence of high-resolution techniques such as X-ray crystallography, solution NMR spectroscopy, cryoelectron microscopy, etc., to some extent relegated hydrodynamic techniques to a secondary role. But recently, mainly over the last two decades, macromolecular hydrodynamics has emerged again as a most valuable structural approach. There are several reasons for this resurgence.

Complex, bulky machinery for analytical ultracentrifugation or dynamic light scattering has been replaced by much more simple and compact instruments, with data collection, handling, and analysis facilitated by user-friendly software. High-performance computational methods allow numerical approaches for difficult problems in hydrodynamic theory. Such methods are implemented in hydrodynamic modeling procedures that allow the direct calculation of hydrodynamic properties from the high-resolution structures specified by, for instance, atomic coordinates in a PDB (protein data bank) file [6] or an electron density map [7]. Thus, it is possible to confirm, from simple measurements and calculations, that structures obtained from crystal diffraction or cryoelectron microscopy are compatible with the true structure in solution.

The technicalities in modern instrumentation, software for data acquisition and analysis, and computational tools for hydrodynamic studies may blur essential concepts that are necessary for a proper interpretation of experiments and computations. With this idea in mind, the scope of this article is essentially that of a tutorial, intending to provide background for such an interpretation. We first describe the basic concepts underlying the various techniques for the experimental determination of hydrodynamic properties. Then, we describe fundamental aspects that relate the experimentally determined properties (e.g., diffusion and sedimentation coefficients, intrinsic viscosity, etc.) to the structure of biomacromolecules. We emphasize that the primary outcome of any hydrodynamic theory are some hydrodynamic radii, which are computed directly from the geometry of the hydrodynamic model. Observable properties such as diffusion or sedimentation coefficients are only obtained later by combining those hydrodynamic radii with data of the system under study (e.g., temperature, solvent viscosity, etc.). A description of the hydrodynamics of rigid particles in terms of simple models such as ellipsoids and cylinders provides, in addition to practical applications, notable insights into such hydrodynamic concepts. Modeling procedures, at different levels of detail, are available for arbitrarily shaped particles. The hydrodynamic description of macromolecules that have any kind of flexibility (continuous or segmental) depends on hydrodynamic (frictional) aspects and also on their conformational variability. Over the years, our group has been active in the development of the theory and computational tools to enable the calculation of hydrodynamic properties of rigid and flexible macromolecules. Thus, along this article, we mention and provide an overview of those tools. Certainly, in addition to the bead modeling methodologies that we describe, there are other hydrodynamic approaches, and there are also a variety of programs for bead model calculations, as mentioned below. None-

theless, we do not intend to review all the various alternative approaches and programs. Instead, we concentrate our tutorial review on the bead modeling methodology and computer programs of the HYDRO suite, which is an integrated, easily accessible resource for understanding hydrodynamic modeling and calculating and analyzing solution properties of both rigid and flexible macromolecules.

Hydrodynamics is particularly helpful in the study of protein structure. The case of quasirigid globular proteins is nowadays well covered by a variety of computational methodologies. The new field of intrinsically disordered (ID) proteins presents a new challenge: the description of macromolecules which are partially ordered, partially unfolded, and partially flexible. The importance that macromolecular hydrodynamics has for structural determination of folded proteins can and should be extended to ID proteins. Thus, throughout this article, we describe here some recent contributions in this regard, where the hydrodynamic properties of partially disordered proteins have been predicted with remarkable success.

Hydrodynamic techniques and solution properties

Translational friction and diffusion

The translational diffusion coefficient, D_t , determines the microscopic translational Brownian motion of the solute molecules and is experimentally observable from the spread of a solution/solvent boundary [8], or indirectly observable in techniques that depend on diffusion, as described below. The fundamental theory of Einstein, which will be outlined in the next section, relates D_t to the frictional coefficient f_t , as $D_t = k_B T / f_t$, where k_B is the Boltzmann's constant and T is the absolute temperature. The frictional coefficient, f_t , depends on the size and shape of the particle. For spherical particles of radius a , Stokes law is $f_t = 6\pi\eta_0 a$, where η_0 is the solvent viscosity, so the particle size can be determined as

$$a = \frac{k_B T}{6\pi\eta_0 D_t} \equiv R_{H,t} \quad (1)$$

For nonspherical particles, eq. (1) is still used as a definition of the so-called hydrodynamic radius, $R_{H,t}$ (sometimes called Stokes radius), which corresponds to the radius of a spherical particle that has the same values of the translational frictional and diffusion coefficients as the nonspherical particle.

Diffusion coefficients can be experimentally determined by a variety of techniques which monitor in some way the Brownian motion of the solute molecules. Such is the case of dynamic light

scattering (DLS), which monitors fluctuations of the intensity of the light scattered, related to the diffusive behavior of the solute. The monograph by Berne and Pecora [9] is the classical reference, and a particularly readable description of theory, instrumentation and data analysis can be found in the review of Stetefeld et al. [10].

The data analysis in DLS allows the consideration of heterogeneity, resulting in a distribution of diffusion coefficients or hydrodynamic radii. Notably, the distribution may cover a wide range of sizes, from about 1 nm up to over 1000 nm. Thus, DLS is most suitable for the characterization of widely polydisperse samples, and the detection of aggregation. The disadvantage of such a dynamic range is the lack of fine resolution. Furthermore, even well-defined species present a peak with some instrumentally caused width. Thus, two species differing in D_t or $R_{H,t}$ by a factor of, say, about 2 (for instance monomer and dimer), could not be distinguished and their peaks would overlap.

Thanks to technical advances and the development of control and data treatment software, DLS is the most frequently used technique for macromolecular characterization based on diffusion, but other methodologies, based on spectroscopic phenomena, are available. Such is the case of pulsed-field gradient (PFG)-NMR spectroscopy [11], which seems particularly suitable for proteins and small peptides [12], or fluorescence correlation spectroscopy [13].

Sedimentation in analytical ultracentrifugation

As indicated in the Introduction, AUC has been for many years the most valued hydrodynamic tool for macromolecules; until the 1970s determination of the sedimentation coefficient was a standard part in the characterization of proteins and nucleic acids, and in the last two decades it has emerged again as an advance thanks to instrumental and computational developments [14,15]. The movement of particles in the ultracentrifuge results from the balance of two forces. One of them is the centrifugal force, $m^{(b)}\omega^2r$, where ω is the rotor angular velocity, r is the distance to the axis of rotation and $m^{(b)} = m(1 - \bar{v}\rho)$ is the buoyant mass of the particle of mass $m = M/N_A$, partial specific volume \bar{v} and molecular weight M , in a solution of density ρ ; $M^{(b)} = M(1 - \bar{v}\rho)$ is the buoyant molecular weight. The other one is the frictional force experienced by the particles, $f_t v$, whose frictional coefficient is f_t when they move with velocity v . The sedimentation coefficient is defined as the ratio of velocity to centrifugal acceleration.

$$s = v / (\omega^2 r) \quad (2)$$

From the balance of forces, it follows that s is determined by f_t and $M^{(b)}$, with f_t in turn related to

the diffusion coefficient, D_t , as per the Einstein equation, so that

$$s = \frac{m^{(b)}}{f_t} = \frac{M^{(b)}}{N_A f_t} = \frac{M^{(b)} D_t}{RT} \quad (3)$$

where N_A is Avogadro's number. Thus, s may provide, such as D_t , information on the size and shape of the solute molecules. In the case of spherical particles, the Stokes-Einstein relationship leads to

$$s = \frac{M(1 - \bar{v}\rho)}{6N_A \pi \eta_0 a} \quad (4)$$

The sedimentation velocity is monitored during the time course of sedimentation by the displacement of the boundary that separates clear solvent from sedimenting solute. For small solutes, or low rotor speeds, the boundary is not sharp but broadened by a concurrent effect of diffusion. The concentration gradient created by sedimentation causes a diffusional counterflow. Thus, the description of the particles' motion is more complex but feasible with present methods of data analysis [16–18], which allow not only the determination of s but also the estimation of M and D_t . AUC is also a separative methodology which is particularly useful for heterogeneous samples. The concentration or signal profile along the cell is the superposition of contributions corresponding to the various species with different s and M , so that their molecular properties can be obtained with proper data analysis methods. Furthermore, thanks to the sensitivity of the technique, species with not-too-different parameters can be differentiated.

Intrinsic viscosity

The viscosity of a solution, η , is greater than that of the pure solvent, η_0 . The increase, due to the solute, can be expressed as a relative viscosity $\eta_r = \eta/\eta_0$ or specific viscosity

$$\eta_{sp} = \frac{\eta_r - 1}{c} = \frac{\eta - \eta_0}{\eta_0 c} \quad (5)$$

which depends on concentration and is related to the size and conformation of the solute molecules. At sufficiently low concentrations, the specific viscosity is proportional to the mass concentration, c , and the proportionality constant is the intrinsic viscosity, $[\eta]$. The concentration dependence can be expressed as

$$\frac{\eta_{sp}}{c} = [\eta] + k_H [\eta]^2 c \quad (6)$$

where k_H is the so-called Huggins constant. The primary result, namely the intrinsic viscosity $[\eta]$, can be obtained from a series of measurements of solution viscosity at several concentrations, by linear

extrapolation of η_{sp}/c vs. c . There is a useful shortcut to obtain $[\eta]$ from just one measurement at single concentration, based on the Solomon-Ciuta-Gotesman equation [19,20]:

$$[\eta] = [2(\eta_{sp} - \ln\eta_r)]^{1/2} / c \quad (7)$$

Actually, the numerator of the right-hand side of eq. (7) depends on concentration so weakly that even at moderate concentration it allows a precise single-point determination of $[\eta]$.

Solution viscosities have traditionally been measured using conventional capillary glass viscometers. To obtain sufficiently accurate viscosity increments, measurements had to be carried out at moderate concentrations that require the above-mentioned extrapolations. Nowadays, these shortcomings are overcome with modern instruments, such as those based on the rolling-ball design or microfluidic devices [21,22]. Detailed information on measurement, interpretation, and applications of the intrinsic viscosity in molecular biology can be found in the review by Harding [23].

As shown by Einstein, the increase of viscosity over that of the solvent, caused by the solute, depends on the size and shape of its molecules. As detailed below, the intrinsic viscosity can be formulated as $[\eta] = \nu V N_A / M$, where the size is given by the particle's volume V , and the numerical factor ν depends on the shape. For a spherical solute, Einstein found the numerical factor to be $\nu = 5/2$, so for a sphere of radius a , with $V = 4\pi a^3/3$, we have

$$[\eta] = \frac{10\pi N_A a^3}{3M} \quad (8)$$

Knowing the particle's molecular weight, its size can be determined as

$$a = \left(\frac{3M[\eta]}{10\pi N_A} \right)^{1/3} \equiv R_{H,\eta} \quad (9)$$

For particles of arbitrary shape, eq. (9) can be used to express an effective hydrodynamic radius, $R_{H,\eta}$, which is the radius of a sphere having the same value of $[\eta]$.

Rotational diffusion and related techniques

Rotational diffusion is studied by means of techniques that monitor the reorientation due to Brownian motion of some characteristic vector within the particle, such as the fluorescence transition dipole moment in fluorescence polarization [24,25], or $^{13}\text{C-H}$ or $^{15}\text{N-H}$ bonds in NMR relaxation [26–30]. While translational dynamics is characterized by a single diffusion coefficient, which corresponds to the Brownian motion of the molecule's center of mass, averaged over orientations, the

theoretical description of rotational diffusion in those techniques is made in terms of a rotational diffusion tensor (a symmetric 3×3 matrix, \mathbf{D}_{rr}). The observed dynamics depends on the eigenvalues of this tensor and the orientation of the dipole of the bond with respect to its eigenaxes. For a particle of arbitrary shape, the time or frequency dependence of observed properties depends on up to five relaxation times (for details, see equations (1)–(12) in Ref. [31]). When many dipoles within the molecule are contributing simultaneously to the observations, one single rotational diffusion coefficient, D_r (the trace of \mathbf{D}_{rr}) suffices. Such is also the case when only one dipole is acting, as when the particle has spherical shape, with fully isotropic rotational motion. The primary outcome of the fluorescence polarization or NMR relaxation is a rotational correlation time $\tau_c = 1/(6D_r)$.

The rotational diffusion coefficient of a sphere of radius a is related to a rotational frictional coefficient f_r , as in the case of translation, with $D_r = k_B T / f_r$, where $f_r = 8\pi\eta_0 a^3 = 6V\eta_0$, where $V = 4\pi a^3/3$ is the volume of the sphere. Therefore, the correlation time for a spherical particle, which is the primary outcome of fluorescence or NMR techniques, is given by

$$\tau_c = \frac{4\pi\eta_0 a^3}{3k_B T} = \frac{V\eta_0}{k_B T} \quad (10)$$

and the radius of the spherical particle can be obtained from τ_c .

$$a = \left(\frac{3k_B T \tau_c}{4\pi\eta_0} \right)^{1/3} = R_{H,r} \quad (11)$$

For nonspherical particles, τ_c is greater than the value for a sphere of the same volume; instead of the particle's volume, eq. (10) requires then a greater effective volume that depends on the particle's size and shape. Nonetheless, eq. (11) can then be used to transform a value of τ_c for any arbitrarily shaped particle into another hydrodynamic radius, $R_{H,r}$.

Hydrodynamic theory and simple models

Basic hydrodynamic theory

One of the fundamental insights of Einstein was the connection between the macroscopically observable diffusion of solute molecules, their microscopic Brownian motion, and the friction that they experience when moving in a viscous solvent, expressed by the Einstein equation, $D_t = k_B T / f_t$. The frictional coefficient f_t is the proportionality constant between the translational velocity of a particle, v , in a viscous liquid (the solvent) and the frictional force $F = f_t v$

that it experiences. The problem of viscous friction belongs to the field of fluid mechanics, which shows that f_t is proportional to the solvent viscosity η_0 , and takes the general form that we indicate as

$$f_t = \eta_0 f_t^*(size, shape) \quad (12)$$

where $f_t^*(size, shape)$ is a quantity, with dimension of length, that depends only on the size and shape or conformation of the solute molecules, as described below. For spherical particles of radius a , the frictional coefficient is given by Stokes law, $f_t = 6\pi\eta_0 a$, such that $f_t^*(size, shape) = 6\pi a$, where a is the measure of size, and 6π is the numerical factor for a spherical shape. For any general model of arbitrary geometry, f_t^* can be put in the form of a hydrodynamic radius,

$$R_{H,t} = f_t^* / (6\pi) = 0.053 f_t^* \quad (13)$$

Likewise, for rotational diffusion around a given axis, the rotational diffusion coefficient and the corresponding relaxation time are proportional to solvent viscosity, with a factor $f_r^*(size, shape)$, which in the case of spherical particles is $8\pi a^3$, and can be calculated for any arbitrary shape, and transformed into the (rotational) hydrodynamic radius,

$$R_{H,r} = [f_r^* / (8\pi)]^{1/3} = 0.341 (f_r^*)^{1/3} \quad (14)$$

For the solution viscosity, Einstein's theory tells us that the intrinsic viscosity can be expressed as the ratio of an effective, hydrodynamic volume V_η^* which depends on the size and shape of the particle, and its mass, $m = M/N_A$. The effective volume is the particle volume V multiplied by a numerical factor, ν ,

$$[\eta] = V_\eta^*(size, shape) / m = \nu V N_A / M \quad (15)$$

The dependence on particle size is through V , while ν depends only on the shape of the particle. For spherical particles, the famous Einstein result is $\nu = 5/2$. For an arbitrarily shaped rigid particle, hydrodynamic theory provides, simply, a result for ν from which the hydrodynamic radius comes out immediately as

$$R_{H,\eta} = \left(\frac{3\nu V}{10\pi} \right)^{1/3} = 0.457 (\nu V)^{1/3} \quad (16)$$

To summarize, the quantities f_t^* , f_r^* , and ν , or alternatively the radii $R_{H,t}$, $R_{H,r}$, and $R_{H,\eta}$, must be considered as the primary results of the hydrodynamic theory that is embodied in the various computational tools intended to calculate hydrodynamic properties from structural models; these are the only quantities than can be obtained from the model geometry. Of course, one would like to predict values for the measurable properties, D_t , s , $[\eta]$, D_r 's, and τ 's, to compare them with experimental data. The computer tools do this from the R_H 's using eqs.

(1), (9) and (11), but require additional experimental data of the solute and solvent. Thus, temperature, T , and the strongly temperature-dependent solvent viscosity, η_0 , are needed for D_t , molecular weight M is required for $[\eta]$, and all these, along with solute specific volume \bar{v} , are needed to obtain the sedimentation coefficient, s . The result would accumulate not only the inaccuracies in the hydrodynamic model and the hydrodynamic calculations, but also the uncertainties in all these additional data.

As an example, consider the calculation of the sedimentation coefficient according to

$$s = \frac{M(1 - \bar{v}\rho)}{6\pi N_A \eta_0} \frac{1}{R_{H,f}} \quad (17)$$

from a computed hydrodynamic radius $R_{H,f} = 3.58$ nm. Suppose that the values of the quantities involved in the first term of the r.h.s. of eq. (17) were as follows: $M = 66 \pm 1$ (1.5%) kDa; $\bar{v} = 0.733 \pm 0.004$ (1.5%) cm^3/g , with $\rho = 0.9983 \pm 0.0001$ (0.1%) g/cm^3 , so that $1 - \bar{v}\rho = 0.267 \pm 0.004$ (1.6%); $\eta_0 = 1.002 \pm 0.012$ (1.2%) Poise (values and uncertainties in ρ and η_0 corresponding to a temperature of 20.0 ± 0.05 °C). Then, eq. (17) gives the following relationship between s in Svedberg units and $R_{H,t}$ in nanometers: s (S) = $[15.52 \pm 0.65(4.3\%)] / [R_{H,t} \text{ (nm)}]$. Thus, even if the computed $R_{H,f}$ value were absolutely exact, the uncertainty in the predicted sedimentation coefficient would be about 4%. Furthermore, the uncertainty in the experimental value of sedimentation coefficient determined by conventional protocols can be as large as 4%, or at least about 1% with special calibration procedures [32].

Accordingly, we suggest that the analysis of hydrodynamic properties—for example, when comparing calculated and experimental values, or results of calculation from different procedures—should be made better in terms of the hydrodynamic radii.

Equivalent radii and ratios of radii

The concept of equivalent hydrodynamic radii can be extended to practically any solution properties. Ortega and García de la Torre [33] proposed a systematic form of expressing a set of solution properties in the form of equivalent radii. Radius a_X is the value of the radius of a sphere that would have the same value of property X. This is just a definition of a different way to express the value of property X, and does not mean at all that the real particle would be spherical. The use of these equivalent radii has important advantages: (i) they all have the dimension of length, (ii) their values are not identical, but not too different from each other, and (iii) they jointly express the size and conformation of the solute in solution, independent of solvent and temperature.

For instance, the equivalent radius for translational diffusion and sedimentation, a_T in this notation, is the Stokes hydrodynamic radius, $R_{H,t}$, obtained from eq. (1). Similarly, for the intrinsic viscosity, a_I is $R_{H,\eta}$ and for rotational dynamics, a_R is $R_{H,r}$. Yet other useful equivalent radii are those corresponding to the geometry of the particle; thus, a_G is the radius of a sphere having the same value of the radius of gyration, R_g , and a_V is that which equalizes the particle's volume, which may, in turn, be expressed in terms of mass and specific volume. Table 1 lists expressions for various equivalent radii.

The use of equivalent radii permits a simple and systematic way to compare values of solution properties. The various properties depend in different ways on particle size (translational properties depend on a linear size, while rotational ones and intrinsic viscosity are related to volume). Furthermore, various physical-chemical properties of solute and solvent enter in several ways. Thus, to compare values for two sets of properties (e.g., computed vs. experimental s , D , $[\eta]$, τ 's, etc.), we prefer to transform all of them into hydrodynamic radii, a_X , and then express the differences as the sum of relative square deviations of the equivalent radii:

$$\Delta^2 \equiv \frac{1}{n_{\text{species}}} \sum_{\text{species}} \frac{1}{n_X} \sum_X \left(\frac{a_X - a_{X,\text{ref}}}{a_{X,\text{ref}}} \right)^2; \quad (18)$$

$$\Delta_{\%} = 100 \sqrt{\Delta^2}$$

The ref values could be those obtained from the experimental data when comparing them with the computed ones. As indicated by \sum_{species} in eq. (18),

this scheme can be applied to analyze data for a collection of species, samples, fractions, etc. Note that the value of the percent root mean square deviation $\Delta_{\%}$ is a useful indicator of the agreement between the two sets. We have written a computer code, HYDFIT [34], to obtain the overall deviation of $\Delta_{\%}$ for a set of experimental or reference data from values of the properties calculated by other programs of the HYDRO suite. Examples are described below.

The a_X radii depend on both the size and conformation of the particle. The dependence on size can be eliminated by formulating ratios of

equivalent radii for two properties, $XY = a_X/a_Y$, which depends only on conformation. As the equivalent radii for different properties are just slightly different, their ratios are of the order of unity, and present a smooth, moderate dependence on conformation. The ratio $TV = a_T/a_V$ is the same as the so-called frictional ratio classically used in macromolecular hydrodynamics, f/f_0 . Other ratios are related to other classical expressions involving two properties, such as the Flory P and Φ ratios combining R_g and either f or $[\eta]$, respectively, or the Scheraga-Mandelkern β coefficient, that combines M , f , and $[\eta]$. The advantage of the ratios of radii is that they are defined in a more systematic way, and take values that are easily comparable and not far from unity. Thus, for a fully flexible chain in a good solvent, $P = 5.3$, $\Phi = 1.9 \times 10^{23}$ and $\beta = 2.19 \times 10^6$ [35], while the corresponding ratios of radii are $GT = 1.87$, $GI = 1.69$, $IT = 1.11$. Some examples of values of the ratios of radii are listed in Table 2. Note that the ratios are easily interconvertible, since

$$XZ = \frac{a_X}{a_Z} = \frac{a_X/a_Y}{a_Z/a_Y} = \frac{XY}{ZY} = XY \cdot YZ \quad (19)$$

As the ratios of the radii depend only on conformational aspects, their main utility lies in obtaining conformational information. The combinations most sensitive to conformation are those involving one geometric property, either the volume V or the radius of gyration. The hydrodynamic volume is only well defined for compact particles, and is possibly influenced by hydration. The combinations in the GX ratios are most suitable for this purpose. On the other hand, we note that the combinations of two properties produce ratios that may be quite insensitive, and very close to unity—the only exceptions are very long rods and some stiff, not-too-long wormlike macromolecules. This happens for the ratio $IT = a_I/a_T = Rh,\eta/Rh,t$ of the Stokes and Einstein hydrodynamic radii, which—with the above exceptions—is usually in the range 1.00–1.10. However, this situation offers an interesting property—the approximate estimation of molecular weight from the measurements of viscosity and diffusion or sedimentation. From equations (1) and (9),

Table 1. Expressions for the equivalent radii.

Radius	From	Definition
$a_T, R_{H,t}$	s and M , or D	$f/(6\pi\eta_0) = k_B T/(6\pi\eta_0 D) = M^{(b)}/(6\pi\eta_0 s)$
$a_I, R_{H,\eta}$	$[\eta]$ and M	$[(3M[\eta])/(10\pi N_A)]^{1/3}$
a_G	R_g	$\sqrt{5/3} R_g$
a_V	V , or \bar{v} , and M	$[3V/(4\pi N_A)]^{1/3} = [(3M\bar{v})/(4\pi N_A)]^{1/3}$

Table 2. Ratios of radii for some typical conformations of rigid and flexible particles.

Conformation, axial ratio (p)	GT	GI	IT	TV	IV
Ellipsoid, $p = 2^a$	1.075	1.067	1.007	1.044	1.052
Cylinder, $p = 3^b$	1.30	1.25	1.04	1.46	1.13
Cylinder, $p = 18^c$	2.42	1.96	1.24	1.85	2.29
Flexible chain, no-EV ^d	1.65	1.53	1.08	—	—
Flexible chain, EV ^e	1.87	1.69	1.11	—	—
Wormlike chain, $L/P = 15^f$	2.61	2.08	1.25	—	—
Wormlike chain, $L/P = 1090^g$	1.96	1.74	1.12	—	—

^a An appreciably prolate globular protein, $p \approx 2$.

^b A 20 base pair B-DNA with $L \approx 6.4$ nm, $d \approx 2.1$ nm, $p \approx 3$.

^c Tobacco mosaic virus, with $L \approx 318$ nm, $d \approx 17$ nm, $p \approx 19$ (ref. [33] and references therein).

^d Very long random coil, no EV (excluded volume).

^e Very long random coil with EV, e.g., high-M single-stranded DNA.

^f A 2311 basepair double-stranded DNA, with $M_L = 1950$ Da/nm, $a \approx 55$ nm, $d \approx 2.1$ nm (ref. [36]).

^g T2 bacteriophage double-stranded DNA, with $M = 1.17 \times 10^8$ Da (ref. [37]).

$$IT = \frac{6\pi\eta_0 D_t}{k_B T} \left(\frac{3M[\eta]}{10\pi N_A} \right)^{1/3} \quad (20)$$

with the proper value for the constants, we have

$$M = 2.47 \times 10^{-27} (IT)^3 \left(\frac{T}{\eta_0 D_t} \right)^3 \frac{1}{[\eta]} \quad (21)$$

If, in the absence of information on the particle's conformation, one accepts a consensus value $IT = 1.05 \pm 0.05$, with an uncertainty of just 5%, then eq. (21) gives an estimate of the molecular weight with an uncertainty of 15% from data for D_t and $[\eta]$. A similar approach can be carried out with viscosity and sedimentation data.

Ellipsoids and cylinders

The first, classic model for anisometric rigid particles [38,39], mainly used to describe globular proteins, was the revolution ellipsoid, which has two identical, b , and one distinct, a , semiaxes, with axial ratio $p = a/b$, such that $p > 1$ for prolate and $p < 1$ for oblate ellipsoids, respectively. Expressions for the frictional coefficient and intrinsic viscosity derived by Perrin [40,41] and Simha [42] can be found in the literature [39,43–45]. The volume and radius of gyration of revolution ellipsoids are $V = 4\pi ab^2$ and $R_g = [(a^2 + 2b^2)/5]$. Harding et al. extended the model to consider triaxial ellipsoids with three unequal semiaxes [23,46,47].

For the representation of rigid, rodlike macromolecules and nanoparticles whose cross-section is uniform (e.g., oligonucleotides, nanotubes, tobacco mosaic virus, etc.), the ellipsoid—being thicker at its center than at the ends—is not appropriate, and a cylindrical model is preferred. The hydrodynamic properties of a cylindrical particle are determined by its length, L , and diameter, d , with aspect ratio $p = L/d$. The asymptotic expressions, for very large p have been known for many years; however, long

rods are likely curved (wormlike, *vide infra*) rather than straight, and the accurate evaluation of properties of moderately long cylinders was made possible more recently, using shell-modeling [48], as described below. The volume and radius of gyration of a cylinder are $V = \pi d^2 L/4$ and $R_g = (L^2/12 + d^2/8)^{1/2}$.

We have written a simple computer program, EllipCylin, for the calculation of hydrodynamic properties of revolution ellipsoids and cylinders. Using just the value of p , the program can already evaluate the ratios of radii. With the whole size information, i.e., the values of a and b or L and d , it calculates all the pertinent equivalent radii. As indicated above, the appropriate hydrodynamic calculation ends here, obtaining the equivalent radii from the particle's geometry. If values of M , η_0 , \bar{v} , etc., are supplied, the program will report finally values for the solution properties s , D , $[\eta]$, etc.

Fig. 1 displays the values of several ratios of radii of cylindrical particles, both rodlike $p > 1$, and disklike $p < 1$. We note the different sensitivity of the various ratios to the shape of the particle. For rods, the combinations involving a geometric property, present an appreciable dependence on the aspect ratio. As the molecular volume V may be somehow undefined, the most suitable combinations are those involving the radius of gyration. With the value of GT from measurements of R_g and s or D_t , or GI from R_g and $[\eta]$, the dimensions of the particle can be readily evaluated. However, as anticipated above, the IT ratio which combines two hydrodynamic properties remains very close to unity for disks and moderately long rods, and so the two hydrodynamic radii are very similar. A similar plot for ellipsoids covering both oblate and prolate (not shown) presents the same tendencies, such that the procedure for estimating molecular weight based on IT would apply for structures representable as ellipsoids, disks, or moderately long rods.

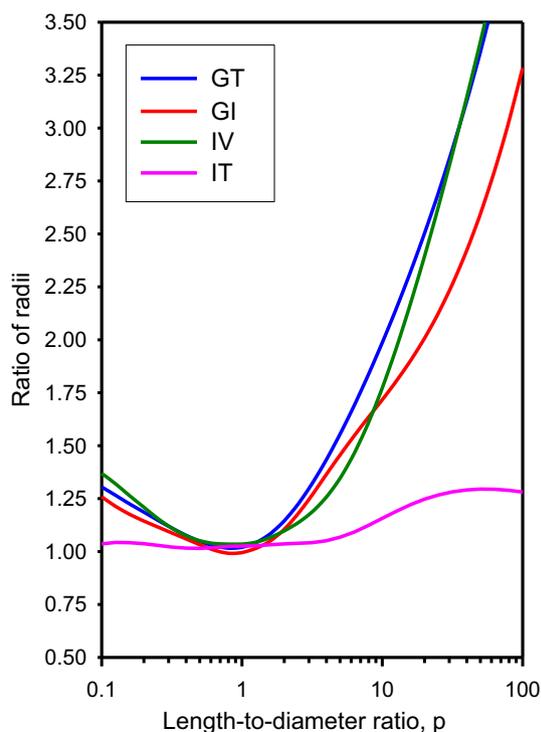


Fig. 1. Ratios of radii GT, GI, IV, IT for cylinder of varying aspect ratio $p = L/d$, from thin disks to moderately long rods. Results from program `EllypCylin`.

Hydrodynamic properties from rigid molecular structures. Detailed models and computational tools

Bead models

The simple ellipsoidal or cylindrical models can be appropriately applied in various cases, as an approximate representation of globular proteins, or short rigid oligonucleotides. However, they are certainly inadequate in many other circumstances dealing with complex multisubunit proteins and macromolecular complexes. In primitive theories of macromolecular hydrodynamics, long macromolecular chains, either rigid rods or flexible coils, were represented by models (like the “pearl necklace”) whose frictional elements are identical spherical beads [49,50], treated with hydrodynamic theories that involve several approximations with the purpose of arriving at analytical expressions. Bloomfield et al. [51] had the remarkable insight of modeling particles of complex shapes as arrays of not necessarily identical beads. With a moderate number N of beads of arbitrary sizes, a sufficiently detailed model can be constructed for any shape. In addition to this procedure for bead modeling in a strict sense,

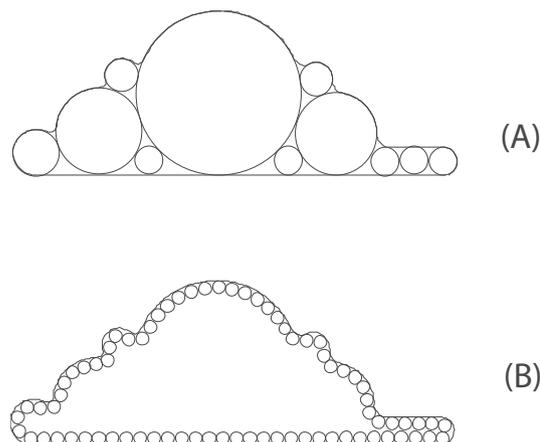


Fig. 2. Two kinds of bead model. (A) Bead model in strict sense, with a moderate number of beads of varying size. (B) Shell model with a large number of minibeats describing the particle's contour.

Bloomfield et al. [52] devised a variant in which the contour of the particle is represented by a shell of small, identical minibeats. The number of elements in the shell model is greater (usually, a few thousand), but allows for a more detailed description as displayed in Fig. 2.

With the primitive computing resources available fifty years ago, the hydrodynamic calculations for bead models had to be carried out with various physical and mathematical approximations. Years later, the advent of more powerful computers made it possible to incorporate in the computational scheme a more rigorous hydrodynamic treatment, particularly in regard to the concept of hydrodynamic interaction (HI). The frictional force experienced by one element in the model does depend linearly on those acting on all the others, and the HI is mediated by the so-called HI-tensors, which had been formulated for polymer models with identical beads [53,54]. Garcia de la Torre and Bloomfield [55–57] extended the treatment for unequal beads and developed a computational scheme for treating the HI effect. The calculation of the three components of the forces acting at the N elements in the model requires the inversion of a matrix of dimension $3N \times 3N$. For an overview of theoretical aspects and modeling strategies, see Ref. [31].

These advances were implemented in the computer program `HYDRO` [58], and upgraded later on to version `HYDRO++` [59], intended for bead models in the strict sense. The primary input is just a user-supplied list of Cartesian coordinates and radii of the N beads, from which the hydrodynamic radii are obtained. Then, the hydrodynamic coefficients and other solution properties are computed with the additional data for the physical properties of solvent and solute.

Solution properties from molecular structure

The advent of detailed, atomic-level structures of biomacromolecules experimentally determined by crystallography or NMR, or constructed with bioinformatic tools, motivated the development of the computer program `HYDROPRO`, whose main input is just a PDB-formatted file containing atomic coordinates. From the PDB file, a primary hydrodynamic model (PHM) is generated, replacing each nonhydrogen atom by a bead. A bead radius $a = 2.9 \text{ \AA}$ gives the best fit of calculated properties to a large set of experimental data of proteins and nucleic acids. Because of the bead size, which is larger than typical bond lengths, beads in the PHM overlap appreciably. Then, in the first version of `HYDROPRO` [6,60], the PHM is internally replaced by a shell model constructed with a large number of identical minibeads. This process is illustrated in Fig. 3. Thanks to the optimization of the computational methods for parallel computing in multicore computers, the CPU time is just a few seconds in a personal computer.

The present version of `HYDROPRO` [61] also implements additional, alternative computational procedures. The problem about bead overlapping has been circumvented using advances in the hydrodynamics of multisphere systems [59,62,63]. Thus, calculations can be made directly from the PHM.

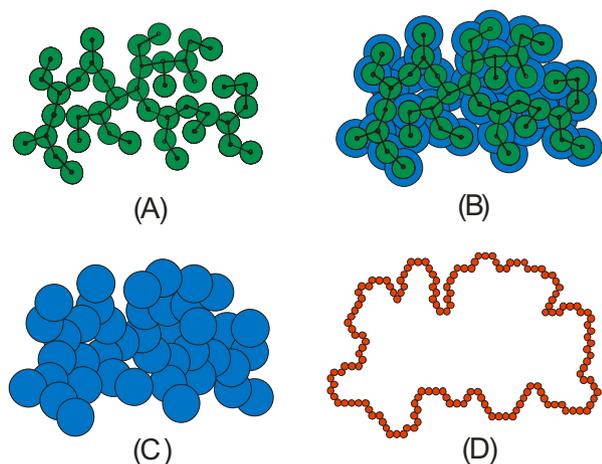


Fig. 3. Schematic description of the construction of hydrodynamic models for a protein. (A) Indicates an atomic-level structure. (B) Beads representing atoms are expanded to account for hydration. (C) The array of overlapping beads is the primary hydrodynamic model, PHM. The contour of this array is filled with closest-packed minibeads (not shown). This “filled model” is used for calculating scattering properties. (D) Removing all the internal beads (those which are surrounded by the maximum number of neighbors) the shell model is obtained.

As the number of beads is the same as the number of non-H atoms, for structures having up to a few thousand atoms, this way of computing is faster than shell modeling, which requires up to 2000 beads. Conversely, for very large biomacromolecules (say, over 200 kDa) the number of atomic beads would be excessive, and then the option is a shell-model calculation. Another possibility in the new version is to make the calculation at the residue-level, with one bead per amino acid residue in the PHM; the optimum bead radius is then $a_{PHM} = 6.1 \text{ \AA}$. This allows us to obtain properties from low-resolution structures, or to make calculations faster as the number of elements in the model is much smaller. We note, also, that the same approach used for proteins can be applied to small oligonucleotides [64].

The accuracy of `HYDROPRO` was tested [61] by comparison of calculated equivalent radii with values from diverse experimental properties (D , s , $[\eta]$, and τ), for a large set of proteins, in terms of the overall deviation $\Delta\%$ formulated in eq. (18) using the `HYDFIT` method. The result varies slightly among the various working modes of `HYDROPRO`; the overall deviation is within 4–6%.

Following the same aim as `HYDROPRO`, other computer programs have been presented for the calculation of hydrodynamic properties from atomic coordinates: `ZENO` [65], `SOMO` [66], `BEST` [67], `US-SOMO` [68], `HullRad` [69], `GRPY` [70], and `Finitelements` [71]. They differ in the modeling strategy or the hydrodynamic formalism for the computation. Comparative descriptions [69,72–74] reveal that the accuracy of these methods does not improve appreciably on that of `HYDROPRO`. This was, in a way to be expected, because, as described above, the experimental equivalent radii carry an uncertainty that comes from (i) the values of M , η_0 , \bar{v} required in their evaluation from the properties (e.g., 4% in R_H from s) and (ii) the experimental error of the property itself (e.g., at least 1%, for s , which results in an expected deviation of 5%). As is obvious, the hydrodynamic model and theories are never perfect.

Based on the same approach as `HYDROPRO`, our program `HYDRONMR` [75–77] extends the hydrodynamic calculations to NMR relaxation times. In addition to the single value of the correlation time, τ_c , which describes overall rotation, as outlined above, NMR measurements provide results for the ratio of spin-spin, T_1 , and spin-lattice, T_2 , residue-specific relaxation times. `HYDRONMR` predicts the sequence of values of the T_1/T_2 ratio; this series, with one value for each amino acid residue, has much more information content than a single-valued property, and depends not only on the shape of the protein molecule through its anisotropic rotational diffusion, but also on the orientation of the ^{15}N –H bonds within the molecule. An example of the results is shown in Fig. 4.

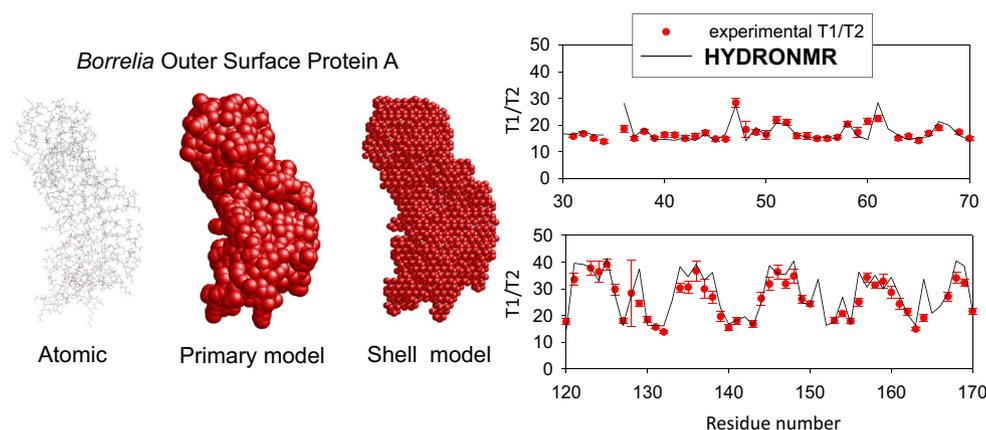


Fig. 4. NMR relaxation results (T_1/T_2 vs. residue number) for *Borrelia burgdorferi* Outer Surface Protein A (600 MHz, 318 K), compared with HYDRONMR calculation with a shell model calculation from atomic coordinates in 1osp.PDB [78].

Another source of biomolecular structures is cryoelectron microscopy, which is particularly valuable for large macromolecular complexes that are difficult to crystallize. The output of this technique is an electron density 3D map, from which the contour of the particle is determined by a cut-off value of the density. The HYDRO suite includes the computer program HYDROMIC [7] which—instead of atomic coordinates, such as in HYDROPRO—takes the density map and constructs a primary hydrodynamic model for which the solution properties are calculated similarly.

Yet another useful tool to predict properties from structural information is HYDROSUB [79] which is intended for situations when a detailed structure is not available. Such is the case for large, partially flexible or labile multisubunit proteins and complexes, such as antibodies [80]. With information on the overall structure of the individual subunits or domains, or their hydrodynamic properties, these can be modeled as ellipsoids or cylinders to obtain their a and b , or L and d dimensions. The geometry of the global model for HYDROSUB is specified by a set of values including coordinates of the subunit centers, dimensions, and orientations specified by a pair of polar angles. The output of HYDROSUB can be transferred to HYDFIT to find the geometry which best fits the solution properties [81].

Although mainly intended for the prediction of hydrodynamic properties, it should be noted that the programs in the HYDRO suite also calculate other solution properties, such as the molecular covolume, which determines the steric contribution to the second virial coefficients. Particularly useful can be the calculation of scattering-related properties; all the programs report the radius of gyration, R_g , the angular dependence of scattering intensity and the distribution of distances, $p(r)$. The calculations are based on particular methods adapted to bead

models [82]. Although tools are available for making such calculations directly from atomic coordinates, e.g., the program CRY SOL [83], the estimations provided by our programs may be suitable and sufficient for some purposes.

Flexible macromolecules

So far, we have considered the case of rigid macromolecules, with a unique conformation, whose shape can be modeled, with more or less detail, as a rigid body for which a hydrodynamic model can be constructed. Theoretical schemes of fluid mechanics are implemented in computational tools that allow the calculation of properties. In the case of flexible particles, able to adopt a variety of conformations, the conformational variability is important, and additional aspects about conformational statistics and internal motion have to be introduced in the theory and computations. In the rest of this article, we describe simple models for flexible macromolecules as well as methods and tools than can be applied in general situations.

The random coil model

The skeleton of macromolecular chains is, in most cases, flexible because of the allowed internal rotation around single bonds. In the absence of strong intramolecular interactions—such as those determining the rigid, secondary structure of proteins or nucleic acids—the conformation looks like a more or less flexible filament. It should be kept in mind that the conformation does not depend just on the local rigidity of the macromolecule but also on the chain length. For very long chains, segments which are far apart in the chain contour behave quite independently, and the molecule adopts the typical random

coil conformation, with the only long-range interaction being the excluded-volume (EV) effect. The conformational statistics of this model [84,85] conclude that the mean (conformationally averaged) radius of gyration R_g of this model can be described by a power-law dependence on the number of segments, N , or molecular weight M ($= NM_1$ for a chain of identical elements of molecular weight M_1). In the absence of excluded volume effects, which is the case for moderately long chains, or long chains in good solvents, the mean square radius of gyration R_g^2 is proportional to N , with R_g proportional to $N^{1/2}$ or $M^{1/2}$. A general expression, which accounts for possible EV effect, is $R_g = K_g M^{\alpha_g}$, where $\alpha_g = 1/2$ if EV is absent, and slightly larger, $\alpha_g \approx 0.6$ with EV effects in good solvents. The quantity K_g combines properties depending only on the local molecular structure of the chain.

Over the years, polymer physicists have devised a number of theoretical treatments for the conformational statistics and hydrodynamic properties of random coil macromolecules. A classical, conceptually clear, and still applicable treatment is the Flory theory [86,87], based on the “equivalent sphere” model. The random coil chain, with the solvent molecules wrapped up by the coil, behaves hydrodynamically as a spherical particle, whose radius R_H should be, simply, proportional to, and not too different from, its radius of gyration, $R_H = QR_g$, where Q is some numerical factor to be determined. From the general relationships between the R_H 's and the hydrodynamic properties (eqs. (1) and (9)), their molecular weight dependence turns out to be given by power laws, $D = k_D M^{\alpha_D}$, $s = k_s M^{\alpha_s}$, and $[\eta] = k_\eta M^{\alpha_\eta}$, with exponents $\alpha_D = -\alpha_g$, $\alpha_s = 1 - \alpha_g$, $\alpha_\eta = 3\alpha_g - 1$, which take the values $-1/2$, $1/2$, and $1/2$, respectively, if EV effect is absent, or -0.6 , 0.4 , and 0.8 in presence of intense EV effect in good solvents. The power-law exponents followed by experimental data are a convenient test to check the random coil behavior of a macromolecule.

The concepts of equivalent radii and ratios of radii are inherent in Flory's theory. The proportionality $R_H = QR_g$, with Q being a universal constant, applicable to any random coil polymer, implies that the ratios of radii $GT = a_G/a_T$ and $GI = a_G/a_I$ (recall that $a_T \equiv R_{H,T}$, $a_I \equiv R_{H,I}$, and $a_G \equiv \sqrt{5}/3 R_g = 1.29 R_g$) should also be universal constants valid for random coil macromolecules in the long-chain limit. For years after Flory's work, the values of these ratios were uncertain because of theoretical difficulties that required various approximations. With the advent of computer simulation methods [35], refined numerical values were reported. The currently accepted values are listed in Table 2.

Thus, the way to evaluate hydrodynamic properties of long random coil polymers begins with the value of R_g . Using the GT and GI ratios, the equivalent radii a_T and a_I are evaluated. Finally,

the values of properties, s , D and η can be calculated from eqs. (1) and (9). It should be kept in mind that adherence to the random coil method requires that the molecular chain be sufficiently flexible and long. As described below, stiff chains can be represented by the wormlike model and, for short chains, one could use model-specific, numerical simulation procedures, which will be presented below.

Wormlike chains

The wormlike chain model, as originally proposed by Kratky and Porod [88], can be envisioned as a continuously curved filament, whose direction of curvature is random, and the degree of curvature depends on the stiffness of the molecular chain that is being represented. It is characterized by three parameters: (i) the contour length, L , which is related to the molecular weight as $M = M_L L$ where the mass per unit length, M_L , is determined by the local molecular structure; (ii) the thickness diameter d of the filament; and (iii) a conformational parameter, the persistence length, P , depending on the filament stiffness, which gauges the chain curvature and overall conformation, such that when $L \gg P$ the chain behaves as a fully flexible coil, and when $L < P$ it behaves as a rigid, straight rod.

The radius of gyration R_g of an infinitely thin ($d \ll L$) wormlike chain is given by an exact expression derived by Benoit and Doty [89] (Paul Doty was author of the first paper published in J. Mol. Biol. [5], as indicated in the Introduction):

$$R_g^2 = 2PL \left[\frac{1}{6} - \frac{1}{2X} + \frac{1}{X^2} - \frac{1}{X^3} (1 - e^{-X}) \right] \quad (22)$$

where $X = L/P$ is the ratio of contour length to persistence length, which is very large for coillike conformations and very small for rodlike ones.

The first, classical but useful, treatment of hydrodynamic properties of wormlike treatments was presented by Yamakawa and Fujii [90,91]. As it happened with the random coil model, their results were influenced by some theoretical approximations regarding hydrodynamic interaction effects, and restrictions about the hydrodynamic diameter, d . These deficiencies were overcome recently by means of computer simulation. Amorós et al. [92] have obtained numerical results for the equivalent radii and hydrodynamic properties of wormlike chains, which have been implemented in the computer program `WormCyl`. With data for the set of parameters, L , (or M and M_L), d , and P , which fully determine the conformation of the wormlike chain, the program computes the equivalent radii, and finally obtains the values of the solution properties R_g , D , s , and $[\eta]$. Furthermore, the core of the `WormCyl` code has been implemented in program `HYDFIT`, which analyzes experimental values of the

solution properties in terms of the global deviation, Δ^2 , defined in eq. (18). Considering the deviation as a function of the structural parameters M_L , d , and P , their optimum values are obtained by minimization.

The wormlike model is useful to interpret solution properties of a variety of biomacromolecules, such as polysaccharides whose molecular chain is rather stiff. But the paradigmatic application has been, for many years now, for double-helical, double-stranded DNA (dsDNA). Amorós et al. [92] undertook the analysis of solution properties for a very large set of experimental data of R_g , D , s , and $[\eta]$ for a number of samples in a very wide range of molecular weight: from oligonucleotides with 8 base pairs, to bacteriophage with M over 10^8 Da and about 200 000 base pairs. HYDFIT found an excellent fit of the whole data set with $P = 560$ Å, $M_L = 195$ Da/Å, and $d = 23$ Å, with a root mean square percent deviation of 6%. The value of P agrees with previous estimates from light-scattering data of R_g analyzed in terms of eq. (22). This gives confidence to the procedures for modeling the wormlike chain. The values found for the other two parameters are particularly relevant with regard to the structure of dsDNA. Note that $L = n_{bp}r$, where n_{bp} is the number of base pairs and r is the rise-per-base pair of the helix, and $M = n_{bp}M_{bp}$, where $M_{bp} \approx 660$ Da is the average molecular weight of a base pair, such that $r = M_{bp}/M_L \approx 3.3$ Å, which is in very good agreement with the crystallographic value for B-DNA. Also, the hydrodynamic diameter agrees very well with the diameter of the bare double helix (20 Å) increased by a monolayer of hydration water. The noteworthy conclusion is that such precise structural information about the double helix can be gathered simply from measurements and computations of hydrodynamic properties.

General model and computer simulations. MONTEHYDRO and SIMUFLEX

Although the random coil and the wormlike model are applicable in many relevant situations, they are just models for more or less flexible linear-chain macromolecules, and do not cover many other relevant situations, which can include, for instance, branched topologies, localized (not uniform) segmental flexibility, etc. The treatment of such cases requires a general physical model which, in addition to beads (obviously required as the frictional elements for hydrodynamics) would take in a variety of intramolecular interactions that account for connectivity (bonds), local angular interactions (bending, torsions), excluded volume effects, electrostatic interactions, etc.

The conformational statistics of such a general model can be simulated by standard Monte Carlo procedures, in which a number of possible conformations are sequentially generated. The simulation is based on the evaluation of the total potential

energy of the model, as the sum of the contributions from all the intramolecular interactions at each simulation step. Conformational properties, such as the radius of gyration R_g , are evaluated as the mean over the values for the individual conformations. However, Monte Carlo is not a dynamic simulation; the successive conformations do not correspond to the real time evolution of the model. Nonetheless, following a proposal by B. Zimm [93] in the so-called rigid-body Monte Carlo method (RBMC), hydrodynamic properties are also evaluated, as was R_g , as conformational averages over the individual values for each conformation, obtained as if they were rigid particles [94,95]. Indeed, this was the procedure used to obtain the abovementioned computational results for random coils and wormlike chains. We have developed a computer program, namely MONTEHYDRO [96] which carries out all the tasks in this scheme: Monte Carlo generation of conformations, hydrodynamic calculations, and conformational averaging.

A most detailed way to simulate the time evolution of biomacromolecules is molecular dynamics (MD) simulation with atomic-level models and explicit solvent. Conformations of the macromolecular solute can be extracted from the MD trajectory, and in the spirit of the RBMC method, they can serve as the starting data for hydrodynamic calculation. Until recently, computing costs for running sufficiently large trajectories (to reach the long times proper of overall diffusion and conformational changes) were prohibitive, but the ever increasing performance of modern computers is making MD increasingly tractable. References [97–99] are just three nice examples of a combined approach in which conformations generated by MD simulations are the basis for building bead models whose hydrodynamic properties are evaluated with HYDROPRO.

Still, there are many situations in which MD is not appropriate. For large macromolecules and macromolecular complexes, MD may be still prohibitive. In addition, in the absence of detailed, atomic-level information, one may use more coarse-grained models in a continuous solvent.

A rigorous way to simulate the time evolution of a macromolecular model is Brownian dynamics (BD) simulation, in which the new conformation after a time step is calculated from the element positions before the step, adding (i) a deterministic displacement caused by the total force acting on the element, balanced by the frictional resistance, and (ii) a random displacement depending on the model's diffusivity, which obeys the Einstein laws for Brownian motion. A simple, popular implementation is the Ermak-McCammon algorithm [100], and other complex but more efficient algorithms are available [101]. In brief, BD simulation requires the evaluation of internal forces (instead of potentials as in RBMC

simulation) and hydrodynamic properties, and the outcome of the simulation is a trajectory of the model, i.e., a succession of conformations corresponding to real time-steps. The determination of solution properties requires a subsequent process to extract them from the trajectory. For example, the diffusion coefficient, D_t , is evaluated from the trajectory followed by the center of mass, according to the Einstein equation,

$$\langle [\mathbf{r}_{CM}(t_0 + t) - \mathbf{r}_{CM}(t_0)]^2 \rangle_{t_0} = 6D_t t \quad (23)$$

where $\mathbf{r}_{CM}(t_0)$ and $\mathbf{r}_{CM}(t_0 + t)$ are, respectively, the position vectors of the center of mass at some time t_0 , and after a time t has elapsed, and $\langle \dots \rangle_{t_0}$ means an average over the choice of the initial time t_0 along the simulated trajectory.

We have developed other computer software, SIMUFLEX [102] for BD simulations of flexible macromolecular models, which comprises two separate tools. First, SIMUFLEX is in charge of simulating the Brownian trajectory, using the same force field as MONTEHYDRO, with the possibility of including also the action of external agents, such as electric fields, flows, obstacles, etc., and so it is useful to predict (corresponding) properties, and for the simulation of single-molecule experiments. For instance, it has been possible to simulate the famous experiments of stretching DNA in elongational flows [103]. The other tool, ANAFLEX contains the procedures, similar to that for D in eq. (23), which analyzes the BD trajectory to obtain hydrodynamic properties and describe various aspects of the dynamic behavior of the macromolecule in solution.

Disordered proteins

Generalities

As is now widely recognized, intrinsically disordered (ID) proteins, which in their native state present a more or less unfolded conformation, play essential roles in a number of physiological aspects. These roles are related to their peculiar conformational and dynamic behavior, and therefore the possibility of analyzing or predicting their conformation and solution properties is evident. Indeed, clear evidence that the native state of these properties is somehow disordered comes from the analysis of solution properties—particularly hydrodynamic radii—in terms of classical macromolecular models, as shown by Uversky [104]. A detailed prediction of functional aspects by simulation would obviously require a detailed, complex model and costly computation, but it seems that simple, coarse-grained models could be sufficient for the overall aspects [105] and even functional aspects [106]. As

ID proteins are more or less flexible, the simulation methodologies implemented in our computer program MONTEHYDRO and SIMUFLEX are readily applicable to their simulation. Our group has proposed a minimal, coarse-grained model which satisfactorily predicts their solution properties, not only the single-valued coefficients, but also more involved and informative data, such as SAXS or NMR relaxation. In this section we summarize some modeling and computing aspects of our work; more details can be found in Ref. [107].

For a minimalistic description of the polypeptide chain, we use a model that contains identical beads, one per each amino acid residue. Beads are joined by bonds corresponding to the $C^\alpha-C^\alpha$ virtual bonds, whose equilibrium length is $b_e = 3.8 \text{ \AA}$. The force field, inspired by the so-called Go model successfully used in protein folding [108,109], includes only potentials for bonds, angles and pairwise non-bonded interaction.

$$V = \sum_{i=1}^{N-1} V_{bond}(b_i) + \sum_{i=1}^{N-2} V_{angle}(\alpha_i) + \sum_{i,j>i+2} V_{nb}(r_{i,j}) \quad (24)$$

A Hookean potential, $V_{bond}(b_i) = (1/2)K_b(b_i - b_e)^2$, quadratic in the instantaneous bond length, b_i , is used for bonds. For bond angles in the virtual chain, another quadratic potential $V_{angle}(\alpha_i) = (1/2)k_\alpha(\alpha_i - \alpha_e)^2$ represents an approximately Gaussian distribution of bond angles α_i in the coarse-grained model chain, centered at α_e and with a width determined by K_α . A torsional potential for dihedral angles is usually included in this force field, but preliminary calculations showed that it had no important effect on the final results. Values for the parameters in all these potential are extracted from statistics of a data base of protein structures [110]. The last term in eq. (24) is a function of the distance r_{ij} between the i and j residues, which differs depending on the region where the residues are placed, as described below.

Unfolded proteins

The simulation methods can be readily applied to the fully unfolded conformation of denatured proteins, for which there are abundant solution property data. A minimal model for fully unfolded proteins would include the potentials for bonds and bond angles, and an excluded-volume repulsion for non-bonded residues. The latter could be simply a hard-spheres potential, $V_{nb}(r_{ij}) = 0$ if $r_{ij} \geq d_{HS}$, where d_{HS} is the hard-sphere diameter, and $V_{nb}(r_{ij}) = K_{HS}$ if $r_{ij} < d_{HS}$, where K_{HS} is an arbitrarily large value. This potential is useful for Monte Carlo simulations but, because of its discontinuity, it

cannot be used in Brownian dynamics simulation, where we can use another simple continuous potential which, in the Go model, is

$$V_{nb}(r_{i,j}) = \epsilon_2(\sigma/r_{i,j})^{12} \quad (25)$$

with $\sigma = 4 \text{ \AA}$ in the original Go model. The main excluded-volume parameter can be determined by comparison of computed and experimental results.

A compilation of data for the radius of gyration of denatured proteins has been reported by Kohn et al. [111]. The simulation of R_g can be made efficiently with MONTEHYDRO, with the results shown in Fig. 5. The optimum value of the hard-spheres diameter, which provides a very good fit of the experimental R_g , was $d_{HS} = 4.5 \text{ \AA}$. Also, hydrodynamic properties of denatured proteins have been profusely studied, with several compilations of data available in the literature [104,112,113]. The prediction of hydrodynamic properties requires, in addition to the force-field parameters, a value for the hydrodynamic

radius of the beads representing the amino acid residues, which has the same meaning as the bead radius in the primary hydrodynamic model of globular proteins at the residue level used in HYDROPRO [61]. Then, we take the same $a_{PHM} = 6.1 \text{ \AA}$ value for unfolded, flexible protein chains. Results for predictions of hydrodynamic properties carried out with MONTEHYDRO or SIMUFLEX showed a satisfactory agreement with experimental data for the translational (Stokes) radius in Fig. 5.

Along with the data for denatured proteins, Fig. 5 also includes data points for some ID proteins which are extensively unfolded, [114]. Although there should be a variety of specific intramolecular interactions in ID proteins that are not present in denatured proteins, it is noticeable that they seem to follow the same tendency.

Unfolded proteins are assumed to behave as random coils. When experimental solution properties are analyzed in terms of their power-law on chain length, as described in a previous section, values of the exponents typical of the random coil model with excluded-volume effects are found: $\alpha_g = 0.59$ [111], $\alpha_D = -0.52$, $\alpha_s = 0.48$, and $\alpha_\eta = 0.66$ [113]. However, this concordance does not discard such intramolecular interaction [115]. Certainly, the minimalistic model does not include features that may be present in either denatured or intrinsically disordered proteins, but looking at the overall good agreement with experimental measurements we consider that it can be reliably used for modeling flexible, unfolded domains in partially disordered proteins, as described in next section.

Partially disordered proteins

Our work on ID proteins [92] focused on a particular kind of ID protein possessing several distinct regions, with some being globular, quasi-rigid domains, and others acting as flexible linkers or tails. The function and the solution properties and dynamic behavior of these proteins must be related to the presence of domains of both types, with different local structure and dynamics. Therefore, the development of a methodology for their simulation with mesoscale, coarse-grained models using MONTEHYDRO and SIMUFLEX was challenging.

Although rigid bead models are adequate for globular proteins, the coexistence of globular and flexible domains requires treating the globules as quasi-rigid parts of the polypeptide chain of marginal flexibility. To maintain the order in the globular domain, an intramolecular potential based on the concept of essential, native contacts can be used. Native contacts that are essential for the three-dimensional tertiary structure of a protein are determined by an algorithm implemented in the Contacts of Structural Units (CSU) software [116].

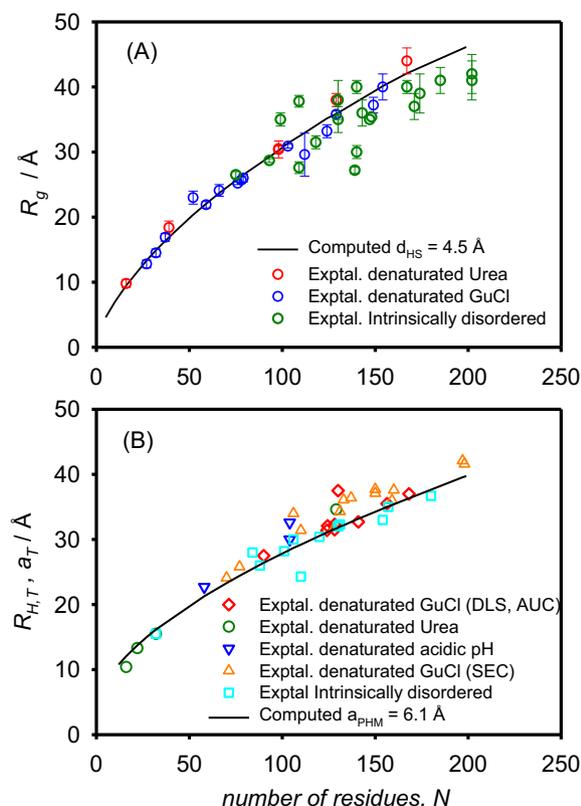


Fig. 5. (A): Radius of gyration of denatured and intrinsically disordered proteins. Denaturation by urea and guanidinium chloride (GuCl). Calculations with the coarse-grained force-field and the adjusted value of the hard-spheres diameter. (B): Hydrodynamic radius from AUC, DLS, and size-exclusion chromatography. Calculations as for R_g with the same bead hydrodynamic radius as for globular proteins.

From PDB files for the globular domains in the partially disordered proteins, program CSU identifies the (i, j) pairs involved in native contacts, and assigns them a short-ranged attractive potential, given by a 10–12 Lennard-Jones equation, as suggested by other authors [117,118]:

$$V_{ne}^{native}(i, j) = \epsilon_1 \left[5 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{10} \right] \quad (26)$$

The potentials for bonds, angles, and all the pairs except those in native contacts are the same as for unfolded chains. All these potentials are implemented in MONTEHYDRO and SIMUFLEX. Values of the parameters can be found in Ref. [92], where this methodology is applied to a variety of partially disordered proteins. A summary of the results is presented in Table 3. Considering the complexity of the diverse aspects involved in the computation (modeling, simulation, conformational statistics, hydrodynamics, etc.), the performance of this scheme, based on our modeling and simulation tools, is certainly satisfactory. In addition to single-valued properties, other more detailed results have been obtained for the angular dependence of scattering intensities, distribution of distances, residue-specific and global correlation time, etc.

Our work on unfolded and partially disordered proteins was based on rather simplified models of the polypeptide chain, with identical elements, neglecting charge effects, etc. Nonetheless, MONTEHYDRO and SIMUFLEX allow—even within a coarse-grained description—to include in the model a variety of intramolecular interactions, such as competing repulsive/attractive van der Waals forces and screened electrostatic interactions, which have successfully used to model the coil-globule collapse transition [126] and polyelectrolyte effects [127] in polymer models. Thus, these tools could be helpful to simulate relevant intramolecular effects in disordered proteins such as formation of blobs [128,129], temporary intrachain links [115] or col-

lapsed regions, in which hydrodynamic interactions may play an important role.

Concluding remarks and perspectives

Hydrodynamic properties were classically, and are again today, extremely useful for the determination of the structure of macromolecules in solution. The availability of readily available instrumentation—like DLS for measuring the diffusion coefficient, and simple viscometers for intrinsic viscosity—and computational tools to predict properties from structure, should encourage their use to confirm structures obtained by techniques like crystallography or microscopy, where the biomacromolecule in not in its natural environment, the aqueous solution. In cases when such techniques are not applicable, data for solution properties—preferably combining two of them—suffice to obtain information on the overall structure, or to fit parameters of a suitable model.

Hydrodynamic calculation can start directly from atomic coordinates, but this is by no means necessary. More coarse-grained hydrodynamic models with, say, one element per amino acid or nucleotide residue may suffice. Instead of atomistic molecular dynamics simulation, more efficient Monte Carlo or Brownian dynamics simulation can yield sufficient structural and dynamic knowledge to analyze properties and predict physiologically relevant behavior in solution. Available tools for modeling and simulation at this level, as described in this article, have found promising applications for intrinsically disordered proteins, whose properties and function are determined by their heterogeneous structure. These tools predict their dilute solution properties with remarkable accuracy, and may, hopefully, be applied to predict also their behavior (dynamics and interactions) in more complex, physiologically relevant systems.

Table 3. Results for some partially disordered proteins.

Protein	Topology ^a	Properties ^b		Reference
		Experimental	Calculated	
pX, Sendai virus	95 = 42(L)+53(G)	Rg = 29.7 or 25	Rg = 26.7	[119,120]
ZipA	306 = 163(L)+143(G)	s = 2.0	s = 2.1	[121]
S4	336 = 193(g)+50(L)+193(G)	Rg = 39.8	Rg = 38.3	[122]
MMP-12	365 = 155(G)+30(L)+180(G)	Rg = 25	Rg = 30.3	[123]
BTK	659 = 169(G) + 51(L) + 55(G) +8(L) + 96(G) + 22(L)	Rg = 50 s = 3.3	Rg = 49 s = 3.6	[124] [125]

^a Topology: total number of amino acid residues, indicating type globular (G) or flexible linker or tail (L).

^b Properties: sedimentation coefficient in S, radius of gyration in Å.

Computer programs

The computer programs mentioned in this article are of public domain and can be downloaded free from our website, <http://leonardo.inf.um.es/macromol>.

Acknowledgments

Financial support was provided by Fundación Séneca, Región de Murcia, Spain, grant 20933/PI/18; Ministerio de Economía y Competitividad, Spain, grant CTQ2017-85425P; FEDER funds, European Commission.

Received 7 October 2019;

Received in revised form 30 November 2019;

Accepted 13 December 2019

Available online 24 December 2019

Keywords:

solution properties;
hydrodynamic models and calculations;
disordered proteins;
Brownian dynamics

References

- [1] A. Einstein, On the movement of small particles suspended in a stationary liquid, demanded by the molecular-kinetic theory of heat, *Ann. Phys.* 17 (1905) 549–560.
- [2] A. Einstein, A new determination of molecular dimensions, *Ann. Phys.* 19 (1906) 289–306.
- [3] T. Svedberg, R. Fahraeus, A new method for the determination of the molecular weight of proteins, *J. Am. Chem. Soc.* 48 (1925) 430–438.
- [4] M. Meselson, F.W. Stahl, The replication of DNA in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A.* 44 (1958) 671–682.
- [5] G. Zubay, P. Doty, The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules, *J. Mol. Biol.* 1 (1959) 1–20.
- [6] J. García de la Torre, M.L. Huertas, B. Carrasco, Calculation of hydrodynamic properties of globular proteins from their atomic-level structures, *Biophys. J.* 78 (2000) 719–730.
- [7] J. García de la Torre, O. Llorca, J.L. Carrascosa, J.M. Valpuesta, HYDROMIC: prediction of hydrodynamic properties of rigid macromolecular structures obtained from electron microscopy, *Eur. Biophys. J.* 30 (2001) 457–462.
- [8] K.E. van Holde, W. Johnson, P. Ho, *Principles of Physical Biochemistry*, second ed., Prentice Hall, Upper Saddle River, N.J., 1998.
- [9] B. Berne, R. Pecora, *Dynamic Light Scattering*, John Wiley and Sons, New York, 1976.
- [10] J. Stetefeld, S.A. McKenna, T. Patel, Dynamic light scattering: a practical guide and applications in biomedical sciences, *Biophys. Rev.* 8 (2016) 409–427.
- [11] P. Stilbs, Fourier transform pulsed-gradient spin-echo studies of molecular diffusion, *Prog. Nucl. Magn. Reson. Spectrosc.* 19 (1987) 1–45.
- [12] S. Yao, D.K. Weber, F. Separovic, D.W. Keizer, Measuring translational diffusion coefficients of peptides and proteins by PFG-NMR, *Eur. Biophys. J.* 43 (2014) 331–339.
- [13] Z. Petrasek, P. Schwille, Precise measurement of diffusion coefficients using fluorescence correlation spectroscopy, *Biophys. J.* 94 (2008) 1437–1448.
- [14] P. Schuck, H. Zao, C. Brautigam, R. Ghirlando, *Basic Principles of Analytical Ultracentrifugation*, CRC Press, Boca Raton, 2016.
- [15] D.J. Winzor, T.R. Patel, D.J. Scott, Analytical ultracentrifugation: a versatile tool for macromolecular complexes in solution, *Methods* 95 (2016) 55–61.
- [16] H.P. Brown, P. Schuck, Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation, *Biophys. J.* 90 (2006) 4651–4665.
- [17] E. Brooks, W. Cao, B.A. Demeler, Two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape, *Eur. Biophys. J.* 39 (2010) 405–418.
- [18] J. García de la Torre, J.G. Hernández Cifre, A.I. Díez Peña, Prediction and analysis of analytical ultracentrifugation experiments for heterogeneous macromolecules and nanoparticles based on Brownian dynamics simulation, *Eur. Biophys. J.* 47 (2018) 845–854.
- [19] O.F. Solomon, B.S. Gotesman, Calculation of viscosity number from a single measurement, *Makromol. Chem.* 104 (1967) 177–184.
- [20] R. Pamies, J.G. Hernández Cifre, M.C. López Martínez, J. García de la Torre, Determination of intrinsic viscosities of macromolecules and nanoparticles. Comparison of single-point and dilution procedures, *Colloid Polym. Sci.* 286 (2008) 1223–1231.
- [21] J. Lee, A. Tripathi, Intrinsic viscosity of polymers and biopolymers measured by microchip, *Anal. Chem.* 77 (2005) 7137–7147.
- [22] S. Gupta, W.S. Wang, S.A. Vanapallia, Microfluidic viscometers for shear rheology of complex fluids and biofluids, *Biomicrofluidics* 10 (2016), 043402.
- [23] S.E. Harding, The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution, *Prog. Biophys. Mol. Biol.* 68 (1997) 207–262.
- [24] M.G. Badea, L. Brand, Time-resolved fluorescence measurements, *Methods Enzymol.* 61 (1979) 378–425.
- [25] J.R. Lakowicz (Ed.), *Principles of Fluorescence Spectroscopy*, third ed., Springer, Berlin, 2006.
- [26] A.G. Palmer, J. Williams, A. McDermott, Nuclear magnetic resonance studies of biopolymer dynamics, *J. Phys. Chem.* 100 (1996) 13293–13310.
- [27] N. Tjandra, D.S. Garrett, A.M. Gronenborn, A. Bax, G.M. Clore, Defining long range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy, *Nat. Struct. Biol.* 4 (1997) 443–449.
- [28] D.M. Korzhnev, M. Billeter, A.S. Arseniev, V.Y. Orekhov, NMR studies of Brownian tumbling and internal motions in proteins, *Prog. Nucl. Magn. Reson. Spectrosc.* 38 (2001) 197–266.

- [29] A.G. Palmer, NMR characterization of the dynamics of biomacromolecules, *Chem. Rev.* 104 (2004) 3623–3640.
- [30] J. García de la Torre, P. Bernadó, M. Pons, Hydrodynamic models and computational methods for NMR relaxation, *Methods Enzymol.* 394 (2005) 419–430.
- [31] B. Carrasco, J. García de la Torre, Hydrodynamic properties of rigid particles. Comparison of different modelling and computational procedures, *Biophys. J.* 76 (1999) 3044–3057.
- [32] H. Zhao, et al., A multilaboratory comparison of calibration accuracy and the performance of external references in analytical ultracentrifugation, *PLoS One* 10 (2015), e0126420.
- [33] A. Ortega, J. García de la Torre, Equivalent radii and ratios of radii from solution properties as indicators of macromolecular conformation, shape, and flexibility, *Biomacromolecules* 8 (2007) 2464–2475.
- [34] A. Ortega, D. Amorós, J. García de la Torre, Global fit and structure optimization of flexible and rigid macromolecules and nanoparticles from analytical ultracentrifugation and other dilute solution properties, *Methods* 54 (2011) 115–123.
- [35] J.M. García Bernal, M.M. Tirado, J. García de la Torre, Monte Carlo calculation of hydrodynamic properties of linear and cyclic chains in good solvents, *Macromolecules* 24 (1991) 593–598.
- [36] S.S. Sorlie, R. Pecora, A dynamic light scattering study of four DNA restriction fragments, *Macromolecules* 23 (1990) 487–497.
- [37] J.A. Harpst, J.R. Dawson, Low angle light scattering studies on whole, half and quarter molecules of T2 bacteriophage DNA, *Biophys. J.* 55 (1989) 1237–1249.
- [38] I.N. Serdyuk, N.R. Zaccai, J. Zaccai, *Methods in Molecular Biophysics. Structure, Dynamics, Function*, Cambridge University Press, New York, 2007.
- [39] C. Tanford, *Physical Chemistry of Macromolecules*, J. Wiley and Sons, New York, 1961.
- [40] F. Perrin, Mouvement Brownien d'un ellipsoïde. I. Dispersion diélectrique pour des molécules ellipsoïdales, *J. Phys. Radium* 5 (1934) 497–511.
- [41] F. Perrin, Mouvement Brownien d'un ellipsoïde. II. Rotation libre et dépolarisation des fluorescences. Translation et diffusion de molécules ellipsoïdales, *J. Phys. Radium* 7 (1936) 1–11.
- [42] R. Simha, The influence of Brownian movement on the viscosity of solutions, *J. Phys. Chem.* 44 (1940) 25–34.
- [43] H.A. Scheraga, Non-Newtonian viscosity of solution of ellipsoidal particles, *J. Chem. Phys.* 23 (1955) 1526–1532.
- [44] T. Yoshizaki, H. Yamakawa, Dynamics of spheroid-cylindrical molecules in dilute solution, *J. Chem. Phys.* 72 (1980) 57–79.
- [45] S.E. Harding, A.J. Rowe, Modelling biological macromolecules in solution: 1. The ellipsoid of revolution, *Int. J. Biol. Macromol.* 4 (1982) 160–164.
- [46] S.E. Harding, On the hydrodynamic analysis of macromolecular conformation, *Biophys. Chem.* 55 (1995) 69–93.
- [47] S.E. Harding, M. Dampier, A.J. Rowe, The viscosity increment for a dilute suspension of triaxial ellipsoids in dominant Brownian motion, *J. Colloid Interface Sci.* 79 (1981) 7–13.
- [48] A. Ortega, J. García de la Torre, Hydrodynamic properties of rodlike and disklike particles in dilute solution, *J. Chem. Phys.* 119 (2003) 9914–9919.
- [49] J.G. Kirkwood, J. Riseman, The intrinsic viscosities and diffusion constants of flexible macromolecules in solution, *J. Chem. Phys.* 16 (1948) 565–573.
- [50] J.G. Kirkwood, P.L. Auer, The visco-elastic properties of solutions of rod-like macromolecules, *J. Chem. Phys.* 19 (1951) 281–287.
- [51] V.A. Bloomfield, W.O. Dalton, K.E. van Holde, Frictional coefficients of multisubunit structures, I. Theory, *Biopolymers* 5 (1967) 135–148.
- [52] V.A. Bloomfield, D.P. Filson, Shell model calculations of translational and rotational frictional coefficients, *J. Polym. Sci. Part C* 25 (1968) 73–83.
- [53] J. Rotne, S. Prager, Variational treatment of hydrodynamic interaction on polymers, *J. Chem. Phys.* 50 (1969) 4831–4837.
- [54] H. Yamakawa, Transport properties of polymer chains in dilute solution: hydrodynamic interaction, *J. Chem. Phys.* 53 (1970) 436–443.
- [55] J. García de la Torre, V.A. Bloomfield, Hydrodynamic properties of macromolecular complexes. I. Translation, *Biopolymers* 16 (1977) 1747–1763.
- [56] J. García de la Torre, V.A. Bloomfield, Hydrodynamic properties of macromolecular complexes. IV. Intrinsic viscosity theory with applications to once-broken rods and multisubunit proteins, *Biopolymers* 17 (1978) 1605–1627.
- [57] J. García de la Torre, V.A. Bloomfield, Hydrodynamic properties of complex, rigid, biological macromolecules. Theory and applications, *Q. Rev. Biophys.* 14 (1981) 81–139.
- [58] J. García de la Torre, S. Navarro, M.C. López Martínez, F.G. Díaz, J.J. López Cascales, HYDRO: a computer software for the prediction of hydrodynamic properties of macromolecules, *Biophys. J.* 67 (1994) 530–531.
- [59] J. García de la Torre, G. Del Río Echenique, A. Ortega, Improved calculation of rotational diffusion and intrinsic viscosity of bead models for macromolecules and nanoparticles, *J. Phys. Chem. B* 111 (2007) 955–961.
- [60] J. García, de la Torre, Hydration from hydrodynamics. General considerations and applications of bead modelling to globular proteins, *Biophys. Chem.* 93 (2001) 159–170.
- [61] A. Ortega, D. Amorós, J. García de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models, *Biophys. J.* 101 (2011) 892–898.
- [62] B. Carrasco, J. García de la Torre, Improved hydrodynamic interaction in macromolecular bead models, *J. Chem. Phys.* 111 (1999) 4817–4826.
- [63] J. García de la Torre, D. Amorós, A. Ortega, Intrinsic viscosity of bead models for macromolecules and nanoparticles, *Eur. Biophys. J.* 39 (2010) 381–388.
- [64] M.X. Fernandes, A. Ortega, M.C. López Martínez, J. García de la Torre, Calculation of hydrodynamic properties of small nucleic acids from their atomic structures, *Nucleic Acids Res.* 30 (2002) 1782–1788.
- [65] E.H. Kang, M.L. Mansfield, J.F. Douglas, Numerical path integration technique for the calculation of transport properties of proteins, *Phys. Rev. E* 69 (2004), 031918.
- [66] N. Rai, M. Nöllman, B. Spottorno, B. Tassara, O. Byron, M. Rocco, SOMO (SOLUTION MOdeller): differences between x-ray and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics, *Structure* 13 (2005) 723–734.
- [67] S.A. Aragón, D.K. Hahn, Precise boundary element computation of protein transport properties: diffusion

- tensors, specific volume and hydration, *Biophys. J.* 91 (2006) 1591–1603.
- [68] B. Demeler, E. Brooks, C. Rosano, M. Rocco, Developments in the US-SOMO bead modelling suite, *Eur. Biophys. J.* 39 (2010) 423–435.
- [69] P.J. Fleming, K.G. Fleming, Hullrad: fast calculations of folded protein and nucleic acid hydrodynamic properties, *Biophys. J.* 114 (2018) 856–869.
- [70] P. Zuk, B. Cichocki, P. Szymczak, GRPY: an accurate bead method for calculation of hydrodynamic properties of rigid macromolecules, *Biophys. J.* 115 (2018) 782–800.
- [71] R.S. Sedeh, G. Yun, J.Y. Lee, K.J. Bathe, D.N. Kim, A framework of finite element procedures for the analysis of proteins, *Comput. Struct.* 196 (2018) 24–35.
- [72] M. Rocco, O. Byron, Computing translational diffusion coefficients: an evaluation of experimental data and programs, *Eur. Biophys. J.* 44 (2015) 417–431.
- [73] M. Rocco, O. Byron, Erratum to: computing translational diffusion coefficients: an evaluation of experimental data and programs, *Eur. Biophys. J.* 44 (2015) 433–436.
- [74] E. Brookes, M. Rocco, Recent advances in the UltraScan SOLUTION MOdeller (US-SOMO), *Eur. Biophys. J.* 47 (2018) 855–864.
- [75] J. García de la Torre, M.L. Huertas, B. Carrasco, HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations, *J. Magn. Reson.* 147 (2000) 138–146.
- [76] P. Bernadó, J. García de la Torre, M. Pons, Interpretation of ^{15}N NMR relaxation data of globular proteins using hydrodynamic calculations with HYDRONMR, *J. Biomol. NMR* 23 (2002) 139–150.
- [77] A. Ortega, J. García de la Torre, Efficient, accurate calculation of rotational diffusion and NMR relaxation of globular proteins from atomic-level structures and approximate hydrodynamic calculations, *J. Am. Chem. Soc.* 127 (2005) 12764–12765.
- [78] N.H. Pawley, S. Koide, L.K. Nicholson, Backbone dynamics and thermodynamics of borrelia outer surface protein A, *J. Mol. Biol.* 324 (2002) 991–1002.
- [79] J. García de la Torre, B. Carrasco, Hydrodynamic properties of rigid macromolecules composed of ellipsoidal and cylindrical subunits, *Biopolymers* 63 (2002) 163–167.
- [80] Y. Lu, E. Longman, K.G. Davis, A. Ortega, G.G. Grossmann, T.E. Michelsenand, J. García de la Torre, Crystalhydrodynamics of protein assemblies: combining sedimentation, viscometry and x-ray scattering, *Biophys. J.* 91 (2006) 1688–1697.
- [81] D. Amorós, A. Ortega, S.E. Harding, J. García de la Torre, Multi-scale calculation and global-fit analysis of hydrodynamic properties of biological macromolecules: determination of the overall conformation of antibody IgG molecules, *Eur. Biophys. J.* 39 (2010) 361–370.
- [82] J. García de la Torre, S.E. Harding, B. Carrasco, Calculation of NMR relaxation, covolume and scattering-related properties of bead models using the SOLPRO computer program, *Eur. Biophys. J.* 28 (1999) 119–132.
- [83] D. Svergun, C. Barberato, M.H.J. Koch, CRY SOL - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates, *J. Appl. Crystallogr.* 28 (1995) 768–773.
- [84] S.F. Sun, *Physical Chemistry of Macromolecules*, second ed., John Wiley and Sons, New Jersey, 2004.
- [85] P.C. Hiemenz, T.P. Lodge, *Polymer Chemistry*, second ed., CRC Press, Boca Raton, 2007.
- [86] P.J. Flory, T.G. Fox, Treatment of intrinsic viscosity, *J. Am. Chem. Soc.* 73 (1951) 1904–1908.
- [87] L. Mandelkern, P.J. Flory, The friction coefficient for flexible chain molecules in dilute solution, *J. Chem. Phys.* 20 (1952) 212–224.
- [88] O. Kratky, G. Porod, Röntgenuntersuchung geloster fagenmoleküle, *Rec. Trav. Chim.* 68 (1949) 1106–1122.
- [89] H. Benoit, P.M. Doty, Light scattering from non-Gaussian chains, *J. Phys. Chem.* 57 (1953) 958–963.
- [90] H. Yamakawa, M. Fujii, Translational friction coefficient of wormlike chains, *J. Chem. Phys.* 6 (1973) 408–414.
- [91] H. Yamakawa, M. Fujii, Intrinsic viscosity of wormlike chains. Determination of the shift factor, *Macromolecules* 7 (1974) 128–135.
- [92] D. Amorós, A. Ortega, J. García de la Torre, Hydrodynamic properties of wormlike macromolecules: Monte Carlo simulation and global analysis of experimental data, *Macromolecules* 44 (2011) 5788–5797.
- [93] B.H. Zimm, Chain molecule hydrodynamics by the Monte-Carlo method and the validity of the Kirkwood-Riseman approximation, *Macromolecules* 13 (1980) 592–602.
- [94] J. García de la Torre, A. Jiménez, J.J. Freire, Monte Carlo calculation of hydrodynamic properties of freely jointed, freely rotating and real polymethylene chains, *Macromolecules* 15 (1982) 148–154.
- [95] A. Iniesta, F.G. Díaz, J. García de la Torre, Transport properties of rigid bent-rod macromolecules and semiflexible broken rods in the rigid body approximation, *Biophys. J.* 54 (1988) 269–275.
- [96] J. García de la Torre, A. Ortega, H.E. Pérez Sánchez, J.G. Hernández Cifre, MULTIHIDRO, MONTEHYDRO, Conformational search and Monte Carlo calculation of solution properties of rigid and flexible macromolecular models, *Biophys. Chem.* 116 (2005) 121–128.
- [97] H. Le, W.L. Dean, R. Buscaglia, J.B. Chaires, J.O. Trent, An investigation of G-Quadruplex structural polymorphism in the human telomere using a combined approach of hydrodynamic bead modeling and molecular dynamics simulation, *J. Phys. Chem. B* 118 (2014) 5390–5405.
- [98] K.T. Walker, R. Nan, D.W. Wright, J. Gor, A.C. Bishop, G.I. Makhatadze, B. Brodsky, S.J. Perkins, Non-linearity of the collagen triple helix in solution and implications for collagen function, *Biochem. J.* 474 (2017) 2203–2217.
- [99] C. Möckel, J. Kubiak, O. Schillinger, R. Kühnemuth, D. Della Corte, G.F. Schröder, D. Willbold, B. Strodel, C.A.M. Seidel, P. Neudecker, Integrated NMR, fluorescence and molecular dynamics benchmark study of protein mechanics and hydrodynamics, *J. Phys. Chem. B* 123 (2019) 1453–1480.
- [100] D.L. Ermak, J.A. McCammon, Brownian dynamics with hydrodynamic interactions, *J. Chem. Phys.* 69 (1978) 1352–1360.
- [101] A. Iniesta, J. García de la Torre, A second-order algorithm for the simulation of the Brownian dynamics of macromolecular models, *J. Chem. Phys.* 92 (1990) 2015–2019.
- [102] J. García de la Torre, J.G. Hernández Cifre, A. Ortega, R. Rodríguez Schmidt, M.X. Fernandes, H.E. Pérez Sánchez, R. Pamies, SIMUFLEX: algorithms and tools for simulation of the conformation and dynamics of flexible molecules and nanoparticles in solution, *J. Chem. Theory Comput.* 5 (2009) 2606–2618.

- [103] T.T. Perkins, D.E. Smith, S. Chu, Single polymer dynamics in an elongational flow, *Science* 276 (1997) 2016–2021.
- [104] V.N. Uversky, What does it mean to be natively unfolded, *Eur. J. Biochem.* 269 (2002) 2–12.
- [105] V. Tozzini, Mapping all-atom models onto one-bead coarse-grained models: general properties and applications to a minimal polypeptide model, *J. Chem. Theory Comput.* 2 (2006) 667–673.
- [106] V. Tozzini, J. Trylska, C. Chang, J.A. McCammon, Flap opening dynamics in HIV-1 protease explored with a coarse-grained model, *J. Struct. Biol.* 157 (2007) 606–615.
- [107] D. Amorós, A. Ortega, J. García de la Torre, Prediction of hydrodynamic and other solution properties of partially disordered proteins with a simple, coarse-grained model, *J. Chem. Theory Comput.* 9 (2013) 1678–1685.
- [108] H. Taketomi, Y. Ueda, N. Go, Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effect of aminoacid sequence represented by specific inter-unit interactions, *Int. J. Pept. Protein Res.* 7 (1975) 445–459.
- [109] C. Clementi, H. Nyemeyer, J. Onuchic, Topological and energetic factors: what determines the structural details of the transition state assemble and en-route intermediates for protein folding? An investigation for small globular proteins, *J. Mol. Biol.* 298 (2000) 937–953.
- [110] G.J. Kleywegt, Validation of protein models from $C\alpha$ coordinates alone, *J. Mol. Biol.* 273 (1997) 371–376.
- [111] J.E. Kohn, I.S. Millett, J. Jacob, B. Zagrovic, T.M. Dillon, N. Cingel, R.S. Dothager, S. Seifert, P. Thiyagarajan, T.R. Sosnick, M.Z. Hasan, V.S. Pande, I. Ruczynski, S. Doniach, K.W. Plaxco, Random-coil behavior and the dimensions of chemically unfolded proteins, *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 12491–12496.
- [112] C. Tanford, K. Kawahara, S. Lapanje, Proteins as random coils. I. Intrinsic viscosities and sedimentation coefficients in concentrated guanidine hydrochloride, *J. Am. Chem. Soc.* 89 (1967) 729–736.
- [113] H. Zhou, Dimensions of denatured protein chains from hydrodynamic data, *J. Phys. Chem. B* 106 (2002) 5769–5775.
- [114] P. Bernadó, M. Blackledge, A self-consistent description of the conformational behaviour of chemically denatured proteins from NMR and small-angle scattering, *Biophys. J.* 97 (2009) 2839–2845.
- [115] G.C. Berry, The hydrodynamic and conformational properties of denatured proteins in dilute solutions, *Protein Sci.* 19 (2010) 94–98.
- [116] V. Sobolev, A. Sorokine, J. Prilusky, E.E. Abola, M. Edelman, Automated analysis of interatomic contacts in proteins, *Bioinformatics* 15 (1999) 327–332.
- [117] N. Koga, S. Takada, Roles of native topology and chain-length scaling in protein folding: a simulation study with a Go-like model, *J. Mol. Biol.* 313 (2001) 171–180.
- [118] T. Frembgen-Kesner, A.H. Elcock, Striking effects of hydrodynamic interactions on the simulated diffusion, folding of proteins, *J. Chem. Theory Comput.* 5 (2009) 242–256.
- [119] L. Blanchard, N. Tarbouriech, M. Blackledge, P. Timmins, W.P. Burmeister, R.W.H. Ruigrok, D. Marion, Structure and dynamics of the nucleocapsid-binding domain of the Sendai virus phosphoprotein in solution, *Virology* 319 (2004) 201–211.
- [120] P. Bernadó, L. Blanchard, P. Timmins, D. Marion, R.W.H. Ruigrok, M. Blackledge, A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering, *Proc. Natl. Acad. Sci. U. S. A* 102 (2005) 17002–17007.
- [121] T. Ohashi, C. Hale, P. de Boer, H. Erikson, Structural evidence that the P/Q domain of ZipA is an structured, flexible tether between the membrane and the C-terminal FtsZ-binding domain, *J. Bacteriol.* 184 (2002) 4313–4315.
- [122] M. Hammel, H.-P. Fierobe, M. Czjzek, V. Kurkal, J.C. Smith, E.A. Bayer, S. Finet, V. Reveceur-Bréchet, Structural basis of cellulosome efficiency explored by small angle X-ray scattering, *J. Biol. Chem.* 280 (2005) 38562–38568.
- [123] I. Bertini, V. Calderone, M. Fragai, R. Jaiswal, C. Luchinat, M. Melikian, E. Mylonas, D.I. Svergun, Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full-length MMP-12, *J. Am. Chem. Soc.* 130 (2008) 7011–7021.
- [124] J.A. Marquez, C.I. Smith, M.V. Petoukhov, P. Lo Surdo, P.T. Mattsson, M. Knekt, A. Westlund, K. Scheffzek, M. Saraste, D.I. Svergun, Conformation of full-length Bruton tyrosine kinase (BTK) from synchrotron X-ray solution scattering, *EMBO J.* 22 (2003) 4616–4624.
- [125] P. Bernadó, E. Mylonas, M.V. Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering, *J. Am. Chem. Soc.* 129 (2007) 5656–5664.
- [126] T. Pham, M. Bajaj, J.R. Prakash, Brownian dynamics simulation of polymer collapse in a poor solvent: influence of implicit hydrodynamic interaction, *Soft Matter* 4 (2008) 1196–1207.
- [127] J.G. Hernández Cifre, J. García de la Torre, Ionic strength effect in polyelectrolyte dilute solutions within the Debye-Hückel approximation: Monte Carlo and Brownian dynamics simulations, *Polym. Bull.* 71 (2014) 2269–2285.
- [128] A.H. Mao, S.L. Crick, A. Vitalis, C.L. Chicoine, R.V. Pappu, Net charge per residue modulates conformational ensembles of intrinsically disordered proteins, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 8138–8188.
- [129] R.K. Das, R.V. Pappu, Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues, *Proc. Natl. Acad. Sci. U.S.A.* 110 (2013) 13392–13397.