Published in final edited form as:
Methods. 2011 May ; 54(1): 101–114. doi:10.1016/j.ymeth.2010.10.005.

RECENT ADVANCES IN MACROMOLECULAR HYDRODYNAMIC MODELING

Sergio R. Aragon
Department of Chemistry, San Francisco State University

Abstract

The modern implementation of the boundary element method (S.R. Aragon, J. Comput. Chem. 25(2004)1191–12055) has ushered unprecedented accuracy and precision for the solution of the Stokes equations of hydrodynamics with stick boundary conditions. This article begins by reviewing computations with the program BEST of smooth surface objects such as ellipsoids, the dumbbell, and cylinders that demonstrate that the numerical solution of the integral equation formulation of hydrodynamics yields very high precision and accuracy. When BEST is used for macromolecular computations, the limiting factor becomes the definition of the molecular hydrodynamic surface and the implied effective solvation of the molecular surface. Studies on 49 different proteins, ranging in molecular weight from 9 to over 400 kDa, have shown that a model using a 1.1 A thick hydration layer describes all protein transport properties very well for the overwhelming majority of them. In addition, this data implies that the crystal structure is an excellent representation of the average solution structure for most of them. In order to investigate the origin of a handful of significant discrepancies in some multimeric proteins (over −20% observed in the intrinsic viscosity), the technique of Molecular Dynamics simulation (MD) has been incorporated into the research program. A preliminary study of dimeric α-chymotrypsin using approximate implicit water MD is presented. In addition I describe the successful validation of modern protein force fields, ff03 and ff99SB, for the accurate computation of solution structure in explicit water simulation by comparison of trajectory ensemble average computed transport properties with experimental measurements. This work includes small proteins such as lysozyme, ribonuclease and ubiquitin using trajectories around 10 ns duration. We have also studied a 150 kDa flexible monoclonal IgG antibody, trastuzumab, with multiple independent trajectories encompassing over 320 ns of simulation. The close agreement within experimental error of the computed and measured properties allows us to conclude that MD does produce structures typical of those in solution, and that flexible molecules can be properly described using the method of ensemble averaging over a trajectory. We review similar work on the study of a transfer RNA molecule and DNA oligomers that demonstrate that within 3% a simple uniform hydration model 1.1 A thick provides agreement with experiment for these nucleic acids. In the case of linear oligomers, the precision can be improved close to 1% by a non-uniform hydration model that hydrates mainly in the DNA grooves, in agreement with high resolution x-ray diffraction. We conclude with a vista on planned improvements for the BEST program to decrease its memory requirements and increase its speed without sacrificing accuracy.

1. Introduction

Hydrodynamic modeling plays an important role in the interpretation and study of molecular motion in liquids. There are many experimental measurements that depend on the transport properties because the diffusion coefficients are parameters in the diffusion equation – this equation governs the probability distribution of molecular positions and/or orientations in the fluid. For example, polarized dynamic light scattering determines the average translational diffusion coefficient and a combination with rotation if the objects are large compared to the wave length [1]; depolarized dynamic light scattering is sensitive to the combination of translation and rotation n[1]; transient electric birefringence is sensitive to the rotational diffusion tensor [2], as is fluorescence depolarization spectroscopy [3]. The translational diffusion coefficient can also be readily determined by NMR pulsed field gradient spin echo techniques [4], and the NMR T1 and T2 time constants are sensitive to a combination of local librations and overall molecular rotational diffusion [5]. An important method apart from these purely spectroscopic methods is ultra centrifugation. This technique induces the molecule to flow in the presence of a centrifugal field and its steady state drift is carefully measured to obtain the sedimentation coefficient [6]. The sedimentation coefficient is proportional to the average translational diffusion coefficient and includes a term that contains the specific volume of the molecule in question. Great advances have been made in ultracentrifugation in recent times allowing the deconvolution of mixtures of several molecules [7]. As experimental techniques advance in precision, a greater need in accuracy and precision in hydrodynamic modeling arises for the proper interpretation of experimental measurements that depend on hydrodynamic transport properties.

The field of hydrodynamic modeling started with ellipsoidal models of molecules because the triaxial ellipsoid (and its degenerate brethren such as a sphere) is one of the few finite shapes for which exact analytic solutions to the Stokes Creeping Flow equations exist (the toroid and the dumbbell complete the set) [8]. The next advance occurred mainly through the work of Bloomfield and co-workers [9,10] and Teller et. al. [11] using an assembly of beads to represent a molecule of any shape, at first as a coarse grained representation. In time, such bead modeling reached atomistic resolution but these methods use approximate hydrodynamic interaction tensors and have limitations for the case of overlapping beads of different sizes [12]. The hydrodynamic interaction of two spheres is given in general as an infinite series expansion in the distance between the spheres. When that distance between spheres exceeds the sum of the diameters, the tensor to first order in the bead size, is given by a variational expression first obtained by Rotne and Prager [13] for the case of equal diameter spheres, and generalized to two unequal bead sizes by Bloomfield and Garcia de la Torre [14]. However, when atomistic resolution is attempted, the spheres representing individual atoms will overlap with nearest neighbors and the bead methodology reaches an impasse: while there exists a limiting form of the Rotne-Prager tensor for overlapping beads of equal diameter, no such tensor has been obtained for unequal diameter spheres. This has led bead modelers to use a basic approximation: resize spheres to make them of equal diameter if they overlap, or even more coarsely, assign a single atomic effective radius (AER) to all the heavy atoms of a molecule in order to avoid this problem [15]. The methodology is successful to a certain degree – the results are generally not accurate enough to correctly interpret subtle effects of hydration or molecular conformation that other more accurate hydrodynamic treatments are able to handle. In this paper we emphasize a much more accurate methodology to solve the Stokes equations of hydrodynamics, the boundary element method

A full implementation of the boundary element method was first provided in hydrodynamics by Youngren and Acrivos [16] in 1975. These authors pointed out that the Stokes equations, ordinarily written as partial differential equations with specified boundary conditions, could

also be written down *exactly* as an integral equation for the velocity field outside an arbitrarily shaped body and implemented an algorithm for its solution. It is interesting to note that the integral equation formulation was known since the work of Odqvist [17] as early as 1930, but the practical implementation of the BE method had to await the emergence of inexpensive and fast computers in modern times. In addition, in integral equation form, it is a simple matter to treat stick, slip, or a mixture of the two boundary conditions because they are incorporated into the integral equation [16,18,19]. In bead methodology, a rigorous treatment of the slip boundary condition does not exist and only adhoc approximations have been attempted so far [20,21]. The introduction of a tensor bead frictional element in the work of Pastor and Zwanzig [21] has been shown to capture some important elements of slip hydrodynamics, however. Another fundamental advantage of the integral equation approach is that the focus is on the hydrodynamic surface of the body being modeled and this surface can be represented by a continuous tessellation of surface elements and the problem of overlapping beads never arises. The starting equation is exact, as emphasized much later by Wegener [22], while in the bead methodology the hydrodynamic interaction tensors are approximate.

The integral equation of hydrodynamics is a Fredholm integral equation of the first kind. After Youngren and Acrivos, Kim and Karilla [8] expounded at length in their modern microhydrodynamics treatise about the pitfalls of using this equation due to that fact that it is ill-conditioned. These authors developed a complex methodology in order to overcome this difficulty – the Completed Double Layer Boundary Integral Method, which has not found much favor so far. However, what was essentially missing is an effective method of regularization of the first kind integral equation. It turns out that the condition that the divergence of the Oseen Tensor is zero yields a zero eigenvalue for the hydrodynamic matrix that arises when the integral equation is discretized. This eigenvalue makes the matrix singular and not invertible. This accounts for the observation of early implementers of the BE method such as Allison [19], that as the number of surface elements was increased, the results of the BE method decreased in quality. Essentially, the round-off error in the matrix computation allowed it to be invertible for small sizes but as the matrix size increases, instability arises.

However, in 2004, Aragon [23] published a new implementation of the BE method in which a robust regularization method was incorporated in a program called BEST. This allowed the solution of the Stokes equations to unprecedented accuracy, as we will demonstrate below. Solutions to the large memory requirements of BEST are in the horizon, and they will be mentioned at the end of this article.

In addition to an accurate solution to the Stokes equations, recent work has also demonstrated that it is possible to treat flexible molecules with a high degree of accuracy when the structures are generated by state of the art molecular dynamics simulations. This work will also be reviewed below.

Lastly, it should be mentioned that even though this volume is mainly concerned with the sedimentation coefficient, we will discuss translation, rotation and intrinsic viscosity in our effort to demonstrate that a properly formulated hydrodynamic model must yield accurate results using the same parameters for all transport properties, not just translation.

2. The Integral Equation of Stokes Flow for stick boundary conditions

For macromolecules, consideration of the solvent as a continuum is an excellent approximation and the governing equations, in the limit of small Reynolds number appropriate for the diffusion process, are known as the Stokes or Creeping Flow equations [8]. Whereas bead methods aim to directly solve a **mobility** problem which cannot be

formulated exactly, an alternative method is to solve a **resistance** problem which can be formulated exactly as an integral equation. As is shown below, once one has precise friction tensors, it is straightforward to compute the mobility: diffusion tensors. For the case of macromolecules in aqueous solution, "stick" boundary conditions are appropriate. In this case, the velocity field of the flow, **v**(**y**) at position **y** in the fluid, can be written *exactly* as an integral over the particle surface (SP),

$$\mathbf{v}(\mathbf{y})=\mathbf{u}_o(\mathbf{y})+\int_{sp}\overleftrightarrow{\mathbf{T}}(\mathbf{x},\mathbf{y})\cdot\mathbf{f}(\mathbf{x})\mathrm{d}S_x \tag{1}$$

where $\mathbf{u}_o(\mathbf{y})$ is the flow velocity of the fluid if the particle was not there (which can be taken to be zero for diffusive motion), and $\mathbf{T}\leftrightarrow(\mathbf{x},\mathbf{y})$ is the Oseen hydrodynamic interaction tensor. The surface stress force, **f**(**x**) is the unknown quantity that we must obtain. Once this quantity is known, the transport properties of the macromolecule can be directly computed, as shown below. The Oseen Tensor, given by, [8,24]

$$\overleftrightarrow{\mathbf{T}}(\mathbf{x},\mathbf{y})=\frac{1}{8\pi\eta|\mathbf{x}-\mathbf{y}|}\left[\overleftrightarrow{\mathbf{I}}+\frac{(\mathbf{x}-\mathbf{y})(\mathbf{x}-\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^2}\right] \tag{2}$$

is an *exact* representation of the hydrodynamic interaction of the infinitesimal surface elements. Thus the starting expressions for the calculation, unlike the bead modeling case, are exact [16,22], moreover, the equation is applicable to bodies of arbitrary shape.

Since equation (1) is an integral equation, the solution requires an approximate numerical method. The method, however, can be iterated to obtain arbitrary precision. The first step is to discretize the surface by replacing it with a collection of N patches that smoothly tile the molecular surface. We can then write,

$$\mathrm{SP}=\sum_{j=1}^{N}\Delta_j \tag{3}$$

We place the coordinate $\mathbf{x}_j$ at the center of the small patch Δ_j and take the surface stress force **f**(**x**) to be a constant over the entire patch area. This is the basic approximation: it is clear that it will become a better and better approximation as the patch is made small. Thus, an extrapolation to zero size patch leads to a very precise value for the transport properties. With this approximation, eq. (1) becomes a set of 3N equations for 3N unknowns **f**(**x**),

$$\mathbf{v}(\mathbf{y}_k)=\sum_{j=1}^{N}\overleftrightarrow{\mathbf{G}}_{kj}\cdot\mathbf{f}_j \tag{4}$$

The centerpiece of this set of equations is a set of N completely known 3×3 matrices of coefficients that contain all geometric information, the integrals of the Oseen Tensor over a surface patch,

$$\overleftrightarrow{\mathbf{G}}_{kj}=\int_{\Delta_j}\overleftrightarrow{\mathbf{T}}(\mathbf{x},\mathbf{y}_k)\,dS_x \tag{5}$$

The regularization method perturbs the hydrodynamic super matrix by different factors along the diagonal than for the off-diagonal elements. The perturbation factors tend to one as the number of elements increases, making the perturbation disappear during the extrapolation. In addition to the introduction of a robust regularization method, the other significant advance made in our work is the essentially exact integration of the Oseen Tensor in the above expression [23]. The set of 3N equations can be written all at once,

$$\begin{bmatrix}\mathbf{v}_1\\ \cdot\cdot\\ \cdot\cdot\\ \mathbf{v}_N\end{bmatrix}_{3Nx1}=\begin{bmatrix}\overleftrightarrow{\mathbf{G}}_{11} & \cdot\cdot & \cdot\cdot & \overleftrightarrow{\mathbf{G}}_{1N}\\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot\\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot\\ \overleftrightarrow{\mathbf{G}}_{N1} & \cdot\cdot & \cdot\cdot & \overleftrightarrow{\mathbf{G}}_{NN}\end{bmatrix}_{3Nx3N}\begin{bmatrix}\mathbf{f}_1\\ \cdot\cdot\\ \cdot\cdot\\ \mathbf{f}_N\end{bmatrix}_{3Nx1} \tag{6}$$

from which the unknown surface stress forces can be readily obtained by matrix inversion of the 3N×3N super matrix $\mathbf{G}\leftrightarrow$ (or more efficiently by direct LAPACK [25] solution of the linear system),

$$[\mathbf{f}]_{3Nx1}=[\overleftrightarrow{\mathbf{G}}]^{-1}_{3Nx3N}[\mathbf{v}]_{3Nx1} \tag{7}$$

The total force and torque on the body can be computed from the surface stress forces and these are directly related to the friction tensors $\mathbf{K}\leftrightarrow$ of the body,

$$\mathbf{F}=\sum_{j=1}^{N}\mathbf{f}_j(\mathbf{x})\Delta_j=-\overleftrightarrow{\mathbf{K}}_{tt}\cdot\mathbf{v}_p-\overleftrightarrow{\mathbf{K}}_{tr}\cdot\vec{\omega}_p \tag{8}$$

$$\mathbf{T}=\sum_{j=1}^{N}\mathbf{x}_p\times\mathbf{f}_j(\mathbf{x})\Delta_j=-\overleftrightarrow{\mathbf{K}}_{rt}\cdot\mathbf{v}_p-\overleftrightarrow{\mathbf{K}}_{rr}\cdot\vec{\omega}_p \tag{9}$$

The body can be assumed to have specific translation velocity $\mathbf{v}_p$ and angular velocity ω_p (for example $\omega_p = 0$ and $\mathbf{v}_p = (v_x,0,0)$) to solve the above equations. Thus, 6 calculations suffice to determine all components of the friction tensors. The friction tensors form part of a larger 6×6 tensor that contains information about the pure translational friction (tt), the pure rotational friction (rr) and the coupling that may exist between these (rt and tr). There are actually only 3 independent 3×3 friction tensors because the $\mathbf{K}\leftrightarrow_{tr}$ tensor is the transpose of the $\mathbf{K}\leftrightarrow_{rt}$ tensor. This coupling is small unless the body has a screw-like axis of symmetry [26]. The 6×6 translation-rotation diffusion tensor is given exactly as the inverse of the 6×6 complete friction tensor whose 4 3×3 blocks are the $\mathbf{K}\leftrightarrow$ mentioned above. It is straightforward to show that the 3×3 diagonal blocks of the complete diffusion tensor can be obtained from the friction tensors by an easy 3×3 matrix inversion,

$$\overleftrightarrow{\mathbf{D}}_{tt}=kT\left[\overleftrightarrow{\mathbf{K}}_{tt}-\overleftrightarrow{\mathbf{K}}_{tr}\cdot\overleftrightarrow{\mathbf{K}}_{rr}^{-1}\cdot\overleftrightarrow{\mathbf{K}}_{rt}\right]^{-1} \tag{10}$$

$$\overleftrightarrow{\mathbf{D}}_{rr}=kT\left[\overleftrightarrow{\mathbf{K}}_{rr}-\overleftrightarrow{\mathbf{K}}_{tr}\cdot\overleftrightarrow{\mathbf{K}}_{tt}^{-1}\cdot\overleftrightarrow{\mathbf{K}}_{rt}\right]^{-1} \tag{11}$$

Note that the above expressions show that unless the rotation translation coupling is strictly zero (as it is for spheres and symmetric tops), it is not correct to simply invert the friction tensors to obtain the diffusion tensors – other authors have glossed over this fact [27]. A completely correct treatment for the bead case had been presented earlier by Goldstein [28], who showed that the problems addressed by the “volume” and other corrections used in the popular bead codes arise because of the disregard of the rotation-translation coupling which is always present in the unconstrained general friction matrix, regardless of the body symmetry.

BEST computes diffusion tensors in the Center of Diffusion and the friction tensors in the Center of Resistance. Details are presented in Aragon [23]. Furthermore, the more complex expressions for the computation of the intrinsic viscosity are available in Allison [19] and Hahn and Aragon [29]. In the paper by Hahn and Aragon it is also shown that the center of viscosity is not equivalent to the center of diffusion and that a full matrix inversion is required to calculate the viscosity factor in the center of viscosity. These authors also found that the viscosity factor calculated at the body centroid is an excellent approximation to the true value for globular proteins. In centro-symmetric particles, all of these “centers” coincide.

3. Modeling of bodies with a smooth surface

The intrinsic accuracy and precision of the BE method implemented in BEST can be demonstrated by modeling ellipsoids, cylinders and a dumbbell. The smooth surface of these objects implies that there is no error in defining the hydrodynamic surface compared to the molecular case where some detail is necessarily lost due to surface roughness. High precision computations of the transport properties of polygons have also been published [29] but are omitted here for brevity. In Table I we show the computations for the viscosity factor, and the parallel and perpendicular components of the translational and rotational diffusion tensors for symmetric ellipsoids as a function of axial ratio, p. The accuracy shows that these calculations are essentially numerically exact. The viscosity factor has been calculated to 6 digits of precision, while the diffusion tensors have been calculated to 5 digits of precision. For smooth surfaces, the precision is limited ultimately by the number digits in the triangle coordinates, and the smoothness with which the tessellation extrapolates to the exact surface as the number of triangles is increased.

In the case of cylinders there are no analytic formulas to compare with and the precision of the calculation is equal to the accuracy because there is little systematic error in the calculation when suitable tessellations are designed (example shown in Fig. 4). Aragon and Flamik [30] have presented improved numerical computations of the transport properties of cylindrical shapes (rectangular, hemispherical, and open cylinders) for axial ratios between 1 and 100. An example of the precision of the extrapolations as a function of the number of surface elements is shown in Figure 1 for the case of the intrinsic viscosity. The graph is a plot of the transport property as a function of the inverse of the number of triangles used to

represent the surface – the precise value of the property is the intercept of this graph. Note that the Mathematica (Wolfram Research) extrapolation uses a quadratic to represent the curve and the statistical properties of the fit validate all the parameters used. The precision of the intercept extends to 6 significant figures as automatically determined by Mathematica.

The actual data and statistics of the results are given in Table II. Note that the standard error (SE) in the intercept, the viscosity factor at infinite number of triangles, is 0.001, validating the 6^{th} digit in the extrapolation, and the number of digits provided by BEST. Thus the value of the intercept is 104.579 in this example.

In the work of Aragon and Flamik, calculations of this type were used to produce interpolation formulas for all the transport properties of rectangular, open and spherical cylinders to near exact precision valid in the range, $1 \leq p \infty$. Careful attention was given to the mathematical form of the interpolation formulas to yield expressions that are correct in the asymptotic limit. It is worth mentioning that a new mathematical form for the intrinsic viscosity was proposed that is accurate in the entire range of p. The formula is:

$$[\eta]=\frac{N_o V \chi(\mathrm{p})}{M} \tag{12}$$

where the new expression for the dimensionless viscosity factor as a function of axial ratio p is:

$$\chi(p)=\frac{8p^2}{45}\left(Ln(p)+\frac{8}{45*\chi(1)}\right)^{-1}[1+X_\eta(p)] \tag{13}$$

In eq. (12), V is the volume of the body, N_o is Avogadro's number, and M is the molecular weight, giving the intrinsic viscosity the normal units of cm^3/g. Eq. (13) depends applies to rod-like bodies with specific values of the viscosity factor at unit axial ratio $\chi(1)$, and a short polynomial in inverse powers of p characteristic of each shape, $X_\eta(p)$. Details are given in ref 30.

In Figures 2 and 3 we compare the precise calculations of Aragon and Flamik to others available in the literature for the case of the rectangular cylinder as an example.

The figures show that only the path integral method of Mansfield and Douglas [31] has comparable accuracy to the BE method and properly satisfies the asymptotic behavior of the properties. Similar results are obtained for the rotational diffusion tensor for all the cylinder types modeled. Much more detail is available in reference 30. These results show that the regularization method implemented in BEST works extremely well and yields a small dependence of the property as a function of inverse triangle number. The extrapolations make small corrections yielding very high precision.

A final comparison in the case of the dumbbell composed of two spheres of radius "a" gives an example where high precision calculations can help evaluate published data. The results from a BEST computation can be expressed succinctly as follows. The translational diffusion tensor is $\mathbf{D}_{tt} = \{0.919999, 1.03337\}\ kT/(8\pi\eta a)$, the rotational diffusion tensor is $\mathbf{D}_{rr} = \{0.919999, 1.03337\}\ kT/(8\pi\eta a^3)$, and the dimensionless viscosity factor is $\chi = 3.44923$. The quantities in curly brackets represent the perpendicular and the parallel component values of the tensor, in that order. The analytic computation of the viscosity factor of the dumbbell, whose tessellation is seen in Figure 4, can be found published in two theoretical papers. Wakiya [35] gives a value of 3.45, while Brenner [36] gives 3.58. BEST

yields a precise value of 3.44923, clearly indicating that the Wakiya value is the correct one. As a further example, the value of the parallel component of the translational friction tensor, $6\pi\eta a \times 1.29028 = 8\pi\eta a/1.03337$, agrees to all 6 significant figures analytically computed by Goldman et. al. [37]. Unlike the bead methodology [12,27], BEST produces essentially exact results for all transport properties during the same computation.

4. Studies of Globular Proteins

Connolly's program MSROLL [38–40] provides a convenient method to define the hydrodynamic surface of a rigid molecule. Given a pdb structure file that contains atomic coordinates, produced by either crystallography or molecular modeling, a probe sphere of solvent size rolls around the atomic arrangement defining the Molecular Surface. The probe size is obtained by using the same procedure on a water molecule, measuring the surface area obtained for a probe sphere of zero size, and extracting the radius of the equivalent sphere with the same area. The value obtained by this procedure is the default 1.5 A used in MSROLL. MSROLL enables one to define atomic radii but it comes with a default set of atomic Van der Waals radii that have been enlarged slightly due to bonding to hydrogen atoms typically absent in a crystallographic file. In the work of Aragon and Hahn [29,41], where 49 proteins ranging from 9 to over 465 kDa where studied, the default Connolly radii were used. This radius set gives different values to each atom type and to different hydrogen contents of the same atom type. For example NH = 1.65 A, while NH2 = 1.70 A. However, if one makes a calculation of transport properties using a triangulation of the naked molecular surface, the values are systematically lower than experiment. Thus, the hydrodynamic surface must contain a certain amount of hydration in aqueous systems and this amount must be determined by comparison to experiment. A simple way to generate the hydrated surface is to "inflate" all the atoms by uniformly increasing the radius of all atoms in the radius file. This new surface is probed and triangulated by MSROLL to produce input files for BEST. Thus, the atomic radii are only used to define the hydrodynamic surface to be triangulated. The fine triangulations produced by MSROLL are further processed by COALESCE [23], a program that can generate sub-triangulations of a given one, preserving the topological properties of the surface. A sequence of such sub-triangulations with increasing numbers of triangles are analyzed by BEST to produce accurate transport properties via extrapolation, as shown in Fig. 1.

Several authors have tackled the hydration assignment using different methodologies. Zhou [42] used a hydrated surface composed of overlapping atomic spheres for ease of triangulation in an approximate integral equation approach by analogy with electrostatics and obtained a value of 0.9 A, while Garcia de la Torre and co-workers [12], utilizing the traditional hydrodynamically interacting beads methodology obtained a value of 1.2 A. In these approaches it was not generally possible to obtain accurate values of all transport properties with the same parameter, however. Aragon and Hahn [29,41] demonstrated that it was possible to assign a value of 1.1 A to the hydration thickness and simultaneously obtain intrinsic viscosity, translational diffusion, and rotational diffusion tensors in agreement with experiment using the program BEST for a large set of proteins, starting from crystallographic coordinates. The value of the hydration thickness was assigned by simply matching the measured translational diffusion coefficient of a set of four well characterized small proteins (ribonuclease, myoglobin, lysozyme, and chymotrypsinogen) with the uniform increase in atomic size required for the computation to agree. Thereafter ALL proteins, large or small, were treated with the same value of the hydration parameter for all the transport properties. Recently, Venable et. al. [43] have presented a hydrodynamic treatment using explicit waters distributed around a solute molecule using energy criteria instead of an average hydration model as used in the present work.

Note that the hydration parameter is smaller than the radius of a water molecule, indicating the uniform hydration model is an approximation to the average distribution of water molecules on the surface of proteins. Nevertheless, the amount of hydration water deduced from the increase in volume has reasonably good agreement with that measured by other techniques [41]. The typical experimental errors of 3% for translation and 5% for rotation and the intrinsic viscosity limit our ability to distinguish this simple hydration model from more elaborate models in which nonuniform hydration is considered for proteins. An exception to this case is reviewed below in the nucleic acid section of this paper.

The few significant discrepancies with experiment for proteins found by Hahn and Aragon [29] are worth mentioning in more detail. Whereas the computed transport properties for 18 monomeric proteins treated as rigid objects generally agreed within experimental error (and the discrepancies were randomly distributed), there was a subset of 4 out of 13 multimeric proteins (α-chymotrypsin, citrate synthase, inorganic pyrophosphates, catalase) that showed large negative systematic deviations in the intrinsic viscosity exceeding −20%. A portion of the data of Aragon and Hahn is shown in Table III. Whereas the translational diffusion coefficient is a functional of shape divided by a characteristic length, and the rotational diffusion tensor components are functionals of shape divided by a volume, the intrinsic viscosity is exclusively a functional of shape and is thus the most sensitive of the measurements to changes in molecular shape. The results of our protein study indicated that for monomeric proteins, and most multimeric proteins, the crystal structure was a good representation of the average structure in solution. Given that there were only 4 very deviant cases out of 13 in the multimeric protein set, the most reasonable conclusion is that the crystal structure and the average solution structure are significantly different for these proteins. This is a good example of a case where the extra precision available in the BE method enables one to catch a significant molecular behavior that other researchers have missed. In addition, this observation prompted this laboratory to combine the technique of molecular dynamics simulation with hydrodynamic computations and our new results from this combination will be described in the next section. Our preliminary results validate the above conclusion.

Before continuing, we can also ask what other information can be obtained from a precise solution to the integral equation (1). Another great advantage of the integral equation formulation is that the equation references the fluid velocity flow field. Once the surface stress forces have been obtained for a specific rigid motion of the body under study, eq. (1) becomes a tool for the direct computation of the velocity field at any point outside the hydrodynamic surface. Aragon and Hahn [44] used this method to compute the velocity field in the pockets of several small proteins (lysozyme, myoglobin and albumin) and demonstrated that there is significant hydrodynamic stagnation of fluid in such pockets. The fluid essentially moved with the body in this pocket and the velocity magnitude does not decay significantly in the pocket. The calculations were accurate enough to show traces of small eddies in protein pockets. Some typical results are shown in figure 5.

5. Combination of molecular dynamics simulation and hydrodynamic modeling

In the previous section we described work in which proteins were assumed to be rigid objects with the crystal structure representing the average solution structure. This picture works very well for most proteins, however, we would like to know what effect the structural fluctuations present in solution have on the measured transport properties of globular proteins and also how to describe proteins that are flexible or have flexible subdomains. The technique of Molecular Dynamics (MD) simulation is well suited for this task. Note, however, that we use MD to generate a set of conformations from which

transport properties are computed via hydrodynamics, and we do not attempt to generate the diffusion coefficients from the MD dynamics trajectory itself – that method requires very long trajectories which are not practical for proteins in the size range of our interest. Modern day computers do enable us to generate solution conformations for even moderately sized proteins with an explicit solvent simulation. MD work on the large multimeric proteins is ongoing in this laboratory to test the conclusion that some have different structure in solution than in the crystal. We have made a good start in that direction.

5.1 Implicit Water MD of α-chymotrypsin

One of the multimeric proteins that may have a significantly different structure in solution compared to the crystal is α-chymotrypsin [29]. This protein has a dimer-monomer equilibrium that is pH dependent [45] and may be treated with AMBER's sander module (Version 9) at constant pH [46] using an implicit solvent model (explicit solvent simulations are not presently available for constant pH). In an implicit solvent model, water is approximately represented by a continuum fluid with no viscosity, thus the dynamics occur much faster than in a real molecular system, allowing a short simulation to display significant changes. The protocol is similar to that of explicit water simulations described below.

Simulations done at pH 7, where the protein exists as a monomer in solution, do indeed demonstrate that the initial crystal structure falls apart, and, the two pieces separate in time (not shown). At pH 3, however, where the protein is a dimer in solution, the simulation keeps the protein together and deforms its shape, elongating somewhat as the simulation proceeds over 3 ns. The initial and a sample deformed shapes are shown in Fig. 6. The trajectory graphs for the translational diffusion coefficient and for the intrinsic viscosity are shown in figure 7. The graphs clearly show the deformation of the structure as the simulation proceeds because relaxation of the values occurs within the first 1–2 ns of trajectory. The transport properties computed as an average over the last 1 ns of simulation agree much better with experiment [44] than those of the crystal structure, but some discrepancy remains. The hydrodynamic analysis is shown in Table 4, where data for an additional monomeric protein, β-lactoglobulin [47,48] is shown as a control. The β-lactoglobulin MD results are only slightly improved from the crystal structure results, indicating that the force field is sufficiently accurate to model the system well. This is a result in the right direction but the implicit solvent model is a coarse representation of the aqueous medium. What can we learn from a more realistic solvent model?

5.2 Explicit water MD of small proteins

The explicit water MD simulations of the four proteins with significant discrepancies mentioned above are more computationally intensive and the results will be presented in a separate publication. Here we report on MD simulations of several small proteins, some with flexible subdomains. We have used the AMBER (Version 9,10) suite of programs [49], and in particular the parallel program pmemd, to perform explicit water simulations with a TIP3P water model in an octahedral box with periodic boundary conditions. A typical simulation protocol consists of four steps: 1) Energy minimization of the solvated system at constant volume and fixed protein coordinates to relax close contacts with solvent, 2) Energy minimization of the entire system at constant volume with no restraints on the protein atoms, 3) 20 ps of MD simulation at constant volume with temperature increasing from 0 to 300 K with mild restraints on protein atoms, and 4) Production run of MD simulation at constant pressure of 1 atm, and temperature of 300K with no restraints.

The first issue we must confront is the accuracy of the force fields that will yield the computed structures during the simulation. We have used two modern force fields that have

been developed for accurate modeling of proteins: ff03 and ff99SB. These are compared in detail by Hornak et. al. [50] who show that ff03 performs slightly better for small systems such as ubiquitin, while ff99SB performs better for larger systems in the prediction of NMR order parameters which are sensitive to detailed local conformational structure. In order to validate these force fields for whole molecule scale structure probed by hydrodynamics we can perform computations of small monomeric proteins whose crystal structure is a good predictor of solution structure and see if this agreement is maintained during MD simulation. If the force field and simulation process are good, the agreement with experiment will be maintained or only slightly improved.

Aragon and Hong [51] have studied several small proteins with explicit water MD simulation (lysozyme, ribonuclease, bpti, human and mouse ubiqutin) using the Amber pmemd parallel program with a protocol as described previously and an electrostatic cutoff that varied between 15 and 12 A, depending on the size of the octahedral solvent box. The solvent contained only as many ions to make the system neutral, but no added salt. The typical buffer used in experiments has a viscosity about 1% higher than pure water and around 0.1 mM salt which serves to screen electrostatics. In addition, this work included a comparison with implicit water MD (not shown) on the same proteins and found a systematic discrepancy of about 15% compared to explicit water simulations. The more salient points of the data obtained in this study will be reviewed here. Some of the transport properties of lysozyme are shown in Fig. 8. Note that, unlike the MD trajectory observed in Fig. 7 for α-chymotrypsin, the graph of the transport properties for lysozyme along the trajectory do not show a relaxation at small times. The graph fluctuates about the average from the initial points in the trajectory, indicating that the crystal structure is already close to the minimum energy in solution and the structure shows only thermal fluctuations, not a deformation. This result is typical of all the small proteins in this MD study. Average transport properties of lysozyme, ribonucleuase and human ubiquitin are shown in Table V. The first two molecules belong to the initial parametrization set for the determination of the hydration thickness of proteins from the translational diffusion coefficient, so the discrepancy between experiment and the crystal structure is much less than 1%. It is noteworthy, however, that the MD simulation value for Dt also agrees to this level of precision, indicating that the ff03 force field is an excellent descriptor of the structure in solution. The agreement is less satisfactory for the intrinsic viscosity, but the experimental error in these determinations can vary between 5–10%, making both the crystal structure values and the MD simulation values statistically equivalent.

The human ubiquitin molecule has a 6 residue end chain whose last 4 residues are quite flexible, compared to the fairly rigid structures of the other two proteins. However, despite this flexibility, the crystal structure is quite a good representative of the translational diffusion coefficient. In the crystal structure, the conformation of the chain sticks straight out of the molecule, while in the molecular dynamics structures it is generally folded inwards. The MD average intrinsic viscosity has a substantial difference with that from the crystal but unfortunately we are not aware of an experimental measurement to make a fruitful comparison. This example shows that the translational diffusion coefficient is not very sensitive to small conformational changes in solution. The effects of shape can be offset by a change in size, leaving the value of Dt relatively unchanged. The intrinsic viscosity is sensitive only to shape and is a much better discriminator – the MD trajectory structures of ubiqutin show that only the last 4 residues, comprising about 5% of the molecule are actually flexible. In the case of ubiquitin, the table also shows that making the water model more realistic by using a 4 point model yields insignificant change in the computed transport properties. Thus, we can conclude that a TIP3P water model yields an excellent descriptor of the conformations in solution even though the diffusion coefficient of water is more than twice the experimental value [52]. The timing of the dynamics is faster

than in a real solution (allowing useful data to be obtained from shorter trajectories), but the range of structures thermally sampled is unaffected. A more critical test of this conclusion would replace the water model by a more realistic models such as the SCP/E or polarizable water models – this is being pursued in our laboratory.

The experimental rotational diffusion data for lysozyme in Table V tell an interesting story. At first glance it may seem that the two values imply a range of experimental error but theoretically the value measured by fluorescence [53], which samples all the eigenvalues of the diffusion tensor, may be different from the depolarized dynamic light scattering value [54]. If the rotational diffusion tensor is diagonalized in the same principal axes system as the polarizability of lysozyme, then the birefringence value will not depend on the faster "axial" eigenvalue, called Dr_2 in Table V.. Aragon [55] has also implemented a very accurate BE method (POL) for the solution of the electrostatics equations for the determination of classical polarizabilities. Using the program POL, with the identical triangulation input file used for the hydrodynamics, it can be shown that both the polarizability and rotational diffusion tensor are diagonalized in essentially the same principal axes – despite its irregular shape, lysozyme is optically a symmetric top! Thus, the depolarized light scattering value should be compared to the average of the two smaller eigenvalues shown in Table V as Dr_1. The MD value of $Dr_1 = 1.79\ 10^7\ s^{-1}$ is in good agreement with the depolarized dynamic light scattering value of $1.67\ 10^7\ s^{-1}$ [54]. The fluorescence value samples all the eigenvalues because the transition moment is unlikely to be oriented along the principal axes of the rotational diffusion tensor. The fluorescence value [53] $Dr = 2.0\ 10^7\ s^{-1}$ compares very well with the average of the MD (2.03) or crystal structure (2.07) eigenvalues of Dr.

The MD simulations in explicit water appear to provide a very good description of the solution structure of small proteins as measured by hydrodynamic transport properties. Thus, in combination with the data from local structure provided by NMR order parameters, both the whole molecule scale structure and the local structure are well described by the ff03 force field. In the next section we describe similar results for a large flexible protein.

5.3 Explicit water MD simulations of Trastuzumab

Brandt et. al. [56] have carried out explicit water MD simulations of a medium sized flexible protein, Trastuzumab, a monoclonal humanized IgG antibody produced by Genentech which is used in the treatment of breast cancer. This study used the ff99SB force field of Simmerling and co-workers [50] for its enhancement of the description of alpha helix secondary structure in proteins. The antibody is a larger flexible system (150 kDa) whose range of motion is very dependent on an accurate representation of the forces between atoms – the flexibility is due to a small hinge length of protein helix in the middle of the molecule. The simulation of trastuzumab required the construction of a model from pieces that could be crystallized because flexibility has impeded the determination of the structure of the entire antibody by X-ray crystallography. The construction procedure relied on an approximate structure for the hinge postulated by Padlan, and the in-silico mutation of residues to make the model identical in atomic composition to trastuzumab. This initial construct was subject to energy minimization with the ff99 force field to eliminate construction artifacts, and subsequent 20 ns MD simulation with the TIP3P water model, using a protocol as described above for the small proteins. The final structure produced by that simulation was subsequently used in 8 independent 40 ns (TIP3P, 300K, 1 atm, 2 fs time step, SHAKE algorithm applied) simulations with ff99SB and Glycam04 force fields carried out in parallel in Genentech computer clusters. Given that the simulation system contained 318,064 atoms, the production work took several months to finish. A snapshot of the trastuzumab structure from one of the independent simulations is shown in Fig. 9. The 0.34 μs piecewise trajectory was analyzed by computing the transport properties with the

BEST suite, using a 1.1 A uniform hydration model and compared to experiment. The transport properties were averaged over 6000 structures from the simulation. The translational diffusion coefficient of trastuzumab was measured by dynamic laser light scattering and the intrinsic viscosity was measured by a rolling ball viscometer. Both measurements were carefully extrapolated as a function of concentration.

The results of this study are presented in Table VI. The values of the transport properties for each subsimulation are shown, along with the overall average. It is immediately apparent that the experimental data and the simulation ensemble averages agree extremely well. The MD simulation is able to determine the translational diffusion coefficient with a precision of 1.7% and it agrees with the experimental measurement to 0.25%. The rotational correlation time was determined by MD to within 6.3% and agreed with literature values for other IgG's to better than 5%. Finally the intrinsic viscosity was determined to within 4.8% and agreed with the measurement within 2%, well within the measurement uncertainty of 3%. The high precision of the experimental measurements, combined with the high precision of the hydrodynamic computations are key components of the extremely good agreement observed in this study. The only other published MD study of a complete antibody in solution used much smaller length trajectories and did not make comparisons with experiment [58,59].

This study demonstrates that the force fields used generate an excellent representation of the solution structure of the antibody. The original paper by Brandt et. al. contains a movie of the complete simulation trajectory in the published supplementary data [56], along with several figures showing the transport properties along the multiple MD trajectories.

6. Nucleic Acids

We have studied nucleic acids such as RNA and DNA oligomers to determine what degree and distribution of hydration is appropriate for these types of molecules. Perhaps not surprisingly, we find that a simple uniform hydration model identical to that of proteins will yield a quite adequate representation of the hydrodynamics of nucleic acids.

Aragon et. al. [60] have published experimental work on the Optical Kerr effect (OKE) on a system of DNA oligomers and yeast tRNA. As part of this study we computed the diffusion tensors of the t-RNA and compared them to measured values by OKE and fluorescence. We used the same simple hydration model for the protein case with a 1.1 A thickness and a crystallographic structure to obtain with BEST the 5 different relaxation times (for the J=2 rotational diffusion case): 16.93, 17.08, 21.10, 23.89, 25.35 ns. The tRNA is a very asymmetric structure, yielding a set of closely spaced relaxation times that are not resolvable experimentally. The average of the 5 relaxations is 21 ns, in excellent agreement with Fluoresence Polarization Anisotropy experimental values of Schurr et. al. [61] and our own OKE measurements [60]. Fig. 10 presents a triangulation prepared by our program Coalesce that shows the solvent accessible surface of the tRNA. These values were computed from the crystal structure of yeast $tRNA^{phe}$ (1EVV). This work demonstrates that the solution native structure is well represented by the crystal structure in the presence of Mg^{++}, and that a uniform hydration model is a reasonable approximation for amorphous nucleic acids. Since we found for proteins that the hydration picture is not at all affected by flexibility, we would expect this uniform level of hydration to be valid also for flexible RNA molecules.

Furthermore, my graduate student T. Takeda performed a study on the solvation of DNA oligomers for her SFSU M.S. thesis [62]. Short oligomers are effectively rod shaped, and the shape can more sensitively portray the actual distribution of water of hydration in these systems. It has long been thought that the water of hydration is more concentrated in the grooves [63], and x-ray crystallographic data [64] of sufficient resolution that shows the distribution of water and other co-crystallyzing molecules and ions in a DNA dodecamer

corroborates this picture. In the case of DNA oligomers, there is the high quality study of Eimer and Pecora [65] in which both translational and rotational diffusion coefficients of three oligomers (8, 12, 20 mer) have been measured by a combination of polarized and depolarized dynamic light scattering. In our study we computed both diffusion tensors and we compared a uniform hydration model, as found appropriate for proteins, to models in which the solvation (which includes ions) was distributed on the backbone, or in the grooves. We find that the uniform hydration model requires different amount of water for each oligomer in order to match experiment and a discrepancy with experiment of 3%, while solvating the grooves by 3.6 +− 0.3 A thick layer (just slightly thicker than one molecule of water, and independent of oligomer length, and sequence) produces an agreement with experiment better than 1%. Solvating the backbone produced a discrepancy 3 times larger. Fig. 11 shows tessellations of a DNA oligomer with three different hydration models. Fig. 12 shows a typical data set in which the discrepancy with experiment for both translation and rotation are plotted as a function of nitrogen inflation (the extra radius added to the nitrogen atom). Two simultaneously measured hydrodynamic properties are required to pinpoint a hydration parameter.

These preliminary results are very interesting, however, they are dependent on two assumptions that need to be mentioned. First, the uncharged molecule hydrodynamic computation was applied. Since the ionic strength of the solutions was around 0.1 M, the effect of phosphate charges is not expected to be large, but this should be tested. Second, the DNA coordinates for the computations where produced by Spartan (Wavefunction, Inc.), a molecular modeling program that makes DNA of uniform helical diameter, irrespective of sequence. DNA oligomer crystallographic data show differences in the GC and AT phosphorous distances bringing into question the suitability of using molecular modeling programs to construct DNA when we are pursuing questions of fine detail. We did, however, compare the computed transport values from the Williams [63] dodecamer crystal structure to those obtained from Spartan and found an insignificant difference. In addition, ignoring the difference in sequence between the Williams dodecamer and a pure GC tract dodecamer studied by Eimer and Pecora [65] yields the same solvation picture as before, indicating that our assumptions are reasonable. This study demonstrates that one can improve the hydration model for DNA by making the hydration non-uniform, a result that depends on the high accuracy of the computations produced by BEST. The penalty for omitting this level of detail is not large – a uniform hydration model can still predict transport properties with an error no larger than 3%. Other authors have arrived at similar conclusions using a uniform hydration model for nucleic acids but to our knowledge, no one else has studied non-uniform hydration hydrodynamic models for DNA.

7. Future improvement of the BEST hydrodynamic program

We are presently developing a more memory efficient BE method for performing hydrodynamics computations, the Gradient Corrected Boundary Element Method (GCBE). The only significant numerical approximation in our implementation of BE hydrodynamics is taking the surface stress force to be constant over a triangular surface element. This discretization error is eliminated by an extrapolation to an infinite number of triangles. However, fairly large numbers of triangles are required for high accuracy and the memory grows as Mem = 0.5+ 0.06705 N^2 Gbytes of ram per N thousand triangles (the constant allows for the OS requirements). For 15,000 triangles, N = 15 and Mem = 15.6 Gbytes. This large memory requirement has several disadvantages. First, in a multiprocessor machine, one scalar job uses all the available memory, forcing the other processors to be idle. Second, such large memory machines require a 64 bit OS and are not common, inhibiting other research laboratories from using our code. This problem can be partially ameliorated by performing an approximate extrapolation from data obtained at only around 3000 triangles

with a consequent small reduction in precision. Lastly, using a very large number of triangles slows the computation considerably because the solution of the linear system scales as 27 N^3. The GCBE is an alternative methodology that preserves the full accuracy of the calculation but uses about half the number of triangles required presently.

The basic idea behind the GCBE is to represent the unknown function in the integral equation as a Taylor series expansion as shown in eq. 12 below, and use information about the gradient of the surface stress to correct a computation at a smaller number of triangles. Thus, one produces the hydrodynamic super matrix once, but solves the linear system twice, once with a constant surface stress force over a triangle, and a second time with a representation of the surface stress force containing up to quadratic terms in the Taylor expansion. The gradient is estimated by making a general fit to a quadratic basis set with the value of $\vec{f}_k$ from several neighboring triangles. This method also has a greater potential for parallel processing.

$$\vec{f}_k(\vec{x}) = \vec{f}_k(\vec{x}_k) + (\vec{x} - \vec{x}_k) \cdot \vec{\nabla}\vec{f}_k(\vec{x})) \tag{14}$$

The second area of improvement of BEST is the parallelization of the code to make it much faster. Processing hundreds to thousands of structures that result from an MD trajectory can take more than a day of CPU time. The recent introduction of the CUDA Software Development Kit by Nvidia (www.nvidia.com/object/cuda_home_new.html) for the use of advanced graphics cards (GPU) as a general purpose computational engine provides an interesting opportunity to parallelize BEST. The GPU has hundreds of processing units with memories up to 4 GB in the present Tesla cards. The generation of the hydrodynamic matrix is a trivially parallelizable. The solution of the linear system can be parallelized by the utilization of cuBLAS, provided by Nvidia, to generate a LAPACK system with high throughput. This has already been achieved by a commercial group called CULA (www.culatools.com/html_guide), but our own implementation will be license free. The utility of the CUDA SDK has already been demonstrated for scientific applications by the porting of AMBER 11 to the CUDA architecture. We have obtained speedups of a factor of 8 for small proteins using this system.

8. Conclusions

The high precision implemented via the BE method in BEST has allowed us to generate a general model to numerically treat the transport properties of proteins with a single hydration parameter for all proteins regardless of size or flexibility. The hydration thickness of 1.1 A represents the average hydration over the protein surface. Experimental data are presently not sufficiently precise to attempt non uniform hydration models for proteins. The hydration model we have utilized allows for atomic size variation, unlike the approximate models of other authors [15] who have proposed a single atomic equivalent radius (AER) for all heavy atoms. A similar picture is obtained for nucleic acids. The number of nucleic acid systems we have studied is smaller than the protein data set, but our data yield a simple hydration picture for RNA and DNA, similar to that of proteins, capable of achieving experimental agreement within 3% or better with a uniform hydration model of 1.1 A thick. For DNA oligomers, it is possible to generate a statistically more accurate model, with half the uncertainty, by concentrating the hydration in the grooves, in agreement with X-ray crystallography.

Our studies of proteins led us to propose that some multimeric proteins have a conformational rearrangement upon going into solution from the crystal. Our preliminary

work using MD simulation appears to bear this out. In order to validate that the structures generated by MD are actually representative of solution structure, we have performed simulations on a number of small proteins, rigid and flexible, and one medium sized flexible protein. The good agreement we obtain with experiment demonstrates that we have validated both the force fields and the hydrodynamic hydration model for proteins. Furthermore, we do not find any evidence, as claimed by Wright and co-workers [66], that there is any need to implement a special screened hydrodynamic interaction to treat flexible molecules, but this topic merits further work. Our application with the precise hydrodynamics in BEST in combination with the trajectory ensemble average method yields very good agreement with experiment for both small and large proteins, flexible or not.

These studies have opened the door to the study of proteins whose sequence is known, but whose structure is unknown but expected to be similar to other known structures, by a combination of hydrodynamic measurements, accurate hydrodynamic modeling, and molecular dynamics. Careful measurements of the intrinsic viscosity, which do not require very expensive equipment, can provide data with sufficient discriminating power to test different structure hypothesis. This methodology, a generalization of the construction work done on the antibody, would be particularly useful for large proteins not amenable to study by solution NMR methods.

Acknowledgments

The author thanks David Hahn for performing implicit water MD for the α-chymotripsyn and β-lactoglobulin molecules. Financial support for some of the small protein and trastuzumab work was provided by Genentech. This work was funded in part by NIH grant GM52588 to S. Aragon.

References

1. Berne, B.; Pecora, R. Dynamic Light Scattering: with applications to Chemistry, Biology and Physics. Wiley-Interscience; New York: 1976.
2. Eden, D.; Elias, JG. Transient Electric Birefringence of DNA restriction fragments and the filamentous virus Pf3. In: Dahneke, B., editor. Measurement of Suspended Particles by Quasi-Elastic Light Scattering. Wiley-Interscience; New York: 1983.
3. Stryer L. Fluorescence spectroscopy of proteins. Science. 1968; 162:526–533. [PubMed: 5706935]
4. Stallmach F, Galvosas P. Spin Echo NMR Diffusion Studies. Annual Reports on NMR Spectroscopy. 2007; 61:51–131.
5. Sanders, JKM.; Hunter, BK. Modern NMR Spectroscopy. Oxford University Press; New York: 1987.
6. Richards, EG. An introduction to physical properties of large molecules in solution. Cambridge University Press; London: 1980.
7. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J. 2000; 78:1606–1619. [PubMed: 10692345]
8. Kim, S.; Karilla, SJ. Microhydrodynamics. Butterworth-Heinemann; New York: 1991.
9. Bloomfield VA, Dalton WO, van Holde KE. Frictional coefficients of multisubunit structures. I. Theory. Biopolymers. 1967; 5:135–148. [PubMed: 6040712] Frictional coefficients of multisubunit structures. II. Application to proteins and viruses. Biopolymers. 1967; 5:149–159. Ibid. [PubMed: 6040713]
10. Garcia de la Torre J, Bloomfield VA. Hydrodynamic properties of macromolecular complexes III. Bacterial viruses. Biopolymers. 1977; 16:1779–1793. [PubMed: 890069]
11. Teller DC, Swanson E, de Haen C. The translational friction coefficients of proteins. Methods Enzymol. 1979; 61:103–124. [PubMed: 481223]
12. Garcia de la Torre J, Huertas ML, Carrasco B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J. 2000; 78:719–730. [PubMed: 10653785]

13. Rotne J, Prager S. Variational treatment of hydrodynamic interaction in polymers. J Chem Phys. 1969; 50:4831–4837.
14. Garcıa de la Torre J, Bloomfield VA. Hydrodynamic properties of macromolecular complexes. I. Translation. Biopolymers. 1977; 16:1747–1763.
15. Garcia de la Torre J, Huertas ML, Carrasco B. HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. J Magn Reson. 2000; 147:138–146. [PubMed: 11042057]
16. Youngren GK, Acrivos A. Stokes flow past a particle of arbitrary shape: a numerical method of solution. J Fluid Mech. 1975; 69:377–402.
17. Odqvist FGK. On the boundary value problems in hydrodynamics of viscous fluids (German). Math Z. 1930; 32:329–375.
18. Hu CM, Zwanzig R. Rotational friction coefficients for spheroids with the slipping boundary condition. J Chem Phys. 1974; 60:4354–4357.
19. Allison SA. Low Reynolds number transport properties of axisymmetric particles employing stick and slip boundary conditions. Macromolecules. 1999; 32:5304–5312.
20. Venable RM, Pastor RW. Frictional models for stochastic simulations of proteins. Biopolymers. 1988; 27:1001–1014. [PubMed: 3401553]
21. Pastor RW, Zwanzig R. Anisotropic bead models for molecular hydrodynamics. J Chem Phys. 1989; 90:5729–5734.
22. Wegener WA. On an exact starting expression for macromolecular hydrodynamic models. Biopolymers. 1986; 25:627–637. [PubMed: 3708107]
23. Aragon SR. A precise boundary element method for macromolecular transport properties. J Comput Chem. 2004; 25:1191–1205. [PubMed: 15116362]
24. Oseen, CW. Hydrodynamik. Academiches Verlag; Leipzig: 1927.
25. Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Du Croz, J.; Greenbaum, A.; Hammarling, S.; McKenney, A.; Sorensen, D. LAPACK User's Guide. 3. SIAM; Philadelphia: 1999.
26. Brenner H. Coupling between the translational and rotational Brownian motions of rigid particles of arbitrary shape. II. General theory. J Colloid Interface Sci. 1967; 23:407–436.
27. Carrasco B, Garcia de la Torre J. Hydrodynamic Properties of Rigid Particles: Comparison of Different Modeling and Computational Procedures. Biophysical J. 1999; 75:3044–3057.
28. Goldstein RF. Macromolecular diffusion constants: a calculational strategy. J Chem Phys. 1985; 83:2390–2397.
29. Hahn DK, Aragon SR. Intrinsic viscosity of proteins and platonic solids by boundary element methods. J Chem Theory Comput. 2006; 2:1416–1428.
30. Aragon SR, Flamik D. Precise computation of transport properties of cylinders by the boundary element method. Macromolecules. 2009; 42:6290–6299.
31. Mansfield ML, Douglas JF. Transport properties of rod-like particles. Macromolecules. 2008; 41:5422–32.
32. Ortega A, Garcia de la Torre J. Hydrodynamic properties of rodlike and disklike particles in dilute solution. J Chem Phys. 2003; 119:9914–9919.
33. Tirado MM, Garcia de la Torre J. Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders. J Chem Phys. 1979; 71:2581–2587.
34. Broersma SJ. Viscous Force and Torque Constants for a Cylinder. J Chem Phys. 1981; 74:6989–6990.
35. Wakiya S. Slow motion in shear flow of a doublet of two spheres in contact. J Phys Soc Japan. 1971; 31:1581–1587.
36. Brenner H. Rheology of a dilute suspension of axisymmetric Brownian particles. Int J Multiphase Flow. 1974; 1:195–341.
37. Goldman AJ, Cox RG, Brenner H. The slow motion of two identical arbitrarily oriented spheres through a viscous fluid. Chem Eng Sci. 1966; 21:1151–1170.
38. Connolly ML. Molecular surface program. QCPE Bull. 1981; 1:75–83.
39. Connolly ML. The molecular surface package. J Mol Graphics. 1993; 11:139–141.

40. Connolly ML. Analytical molecular surface calculation. J Appl Crystallogr. 1983; 16:548–558.
41. Aragon SR, Hahn DK. Precise Boundary Element Computation of Protein Transport Properties: Diffusion Tensors, Specific Volume, and Hydration. Biophys J. 2006; 91:1591–1603. [PubMed: 16714342]
42. Zhou H. Boundary element solution of macromolecular electrostatics: Interaction energy between two proteins. Biophys J. 1993; 65:955–963. [PubMed: 8218918]
43. Venable RM, Hatcher E, Guvench O, MacKerell AD Jr, Pastor RW. Comparing Simulated and Experimental Translation and Rotation Constants: Range of Validity for Viscosity Scaling. J Phys Chem B. 2010; 114:12501–12507. [PubMed: 20831149]
44. Aragon SR, Hahn DK. Stagnation of flow in protein cavities by boundary element microhydrodynamics. J Phys Chem B. 2009; 113:2657–2663. [PubMed: 19708204]
45. Schwert GW, Kaufman S. The molecular size and shape of the pancreatic proteases. III. α-Chymotrypsin. J Biol Chem. 1951; 190:807–816. [PubMed: 14841234]
46. Mongan J, Case DA, McCammon JA. Constant pH molecular dynamics in generalized Born implicit solvent. J Comput Chem. 2004; 25:2038–2048. [PubMed: 15481090]
47. McKenzie HA, Sawyer WH. Effect of pH on β-lactoglobulin. Nature. 1967; 214:1101–1104. [PubMed: 6053067]
48. Ogston AG. The Guoy diffusiometer; further calibration. Proc Roy Soc (London). 1949; 196:272–285.
49. Perlman DA, Case DA, Caldwell JW, Ross WS, Cheathamm TE III, Debold S, Ferguson D, Seibel G, Kollman P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Computer Phys Commun. 1995; 91:1–41.
50. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins. 2006; 65:712–725. [PubMed: 16981200]
51. Aragon, SR.; Hong, Q. J Phys, Chem (to be submitted); Q Hong, MD of small proteins in solution, MS Thesis, SFSU. 2009. Hydrodynamics and MD of small proteins in solution.
52. Mark P, Nilsson L. Structure and Dynamics of the TIP3P, SPC, and SPC/E water models at 298K. J Phys Chem A. 2001; 105:9954–9960.
53. Cross AJ, Fleming GR. Influence of inhibitor binding on the internal motions of lysozyme. Biophys J. 1986; 50:507–512. [PubMed: 3756301]
54. Dubin SB, Clar NA, Benedek GB. Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: configuration of lysozyme in solution. J Chem Phys. 1971; 54:5158–5164.
55. Aragon SR, Hahn DK. Polarizability and Kerr constant of Proteins by boundary element methods. Colloids and Surfaces B: Biointerfaces. 2007; 56:19–25.
56. Brandt JP, Patapoff TA, Aragon SR. Construction, MD simulation, and hydrodynamic validation of an all-atom model of a monoclonal IgG antibody. Biophysical J. 2010; 99:905–913.
57. Brandt, P. Construction, Molecular Dynamics Simulation, and Hydrodynamic Validation of an All-Atom Model of a Monoclonal Antibody, MS Thesis, SFSU. 2010.
58. Chennamsetty N, Helk B, Voynov V, Kayser V, Trout BL. Aggregation-prone motifs in human immunoglobulin. G J Mol Biol. 2009; 391:404–413.
59. Chennamsetty N, Voynov V, Helk B, Kayser V, Trout BL. Design of therapeutic proteins with enhanced stability. Proc Natl Acad Sci USA. 2009; 106:1937–11942.
60. Aragon, SR.; Perez, M.; Ng, K.; Eden, D. Optical Kerr effect of Dna oligomers and Rna. In: Stoylov, SP.; Stoimenova, MV., editors. Molecular and Colloidal Electro-optics. Taylor & Francis; New York: 2006. p. 265-284.
61. Thomas JC, Schurr JM, Reid BR, Riberiro NS, Hare DR. Effect of magnesium(2+) on the solution conformation of two different transfer ribonucleic acids. Biochem. 1984; 23:5414–5420. [PubMed: 6210103]
62. Takeda, T. MS Thesis, SFSU. August. 2005 Determination of Water Hydration Location in DNA Oligomers through Diffusion Coefficient Calculation.

63. Drew HR, Dickerson RE. Structure of a B-DNA dodecamer. III. Geometry of hydration. J Mol Biol. 1981; 15:535–556. [PubMed: 7338904]

64. Shui X, McFail-Isom L, Hu G, Williams LD. The B-DNA Dodecamer at high Resolution Reveals a Spine of Water on Sodium. Biochem. 1998; 37:8341–8355. [PubMed: 9622486]

65. Eimer W, Pecora R. Rotational and translational diffusion of short rod-like molecule in solution: Oligonucleotides. J Chem Phys. 1991; 94:2324–2329.

66. Bae SH, Dyson HJ, Wright PE. Prediction of the Rotational Tumbling Time for Proteins with Disordered Segments. JACS. 2009; 131:6814–6821.

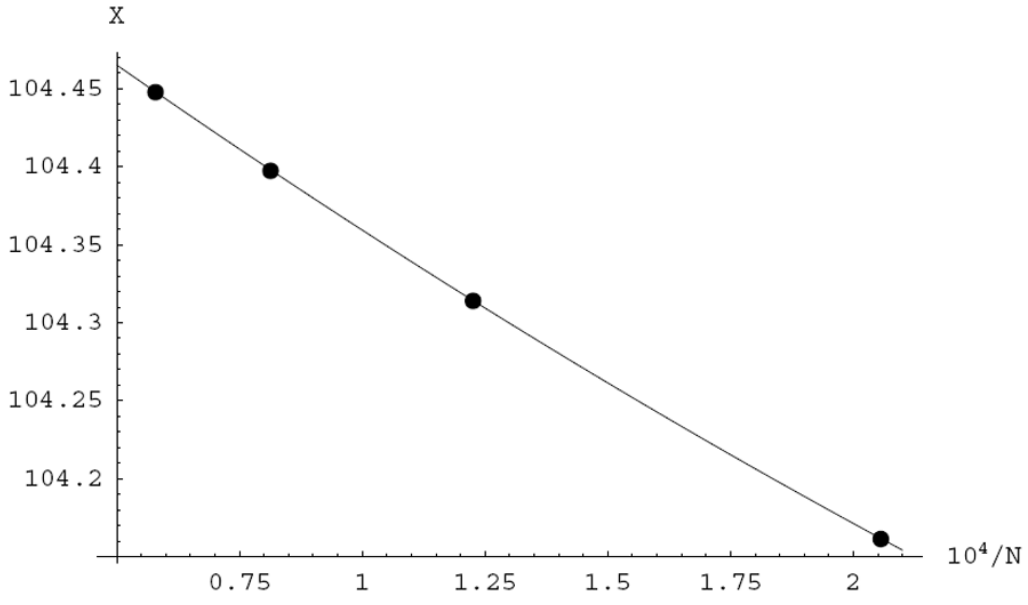

Fig. 1.
The least squares fit line for the viscosity factor extrapolation vs. 1/N of a rectangular cylinder of axial ratio = 40. Taken from reference 29.

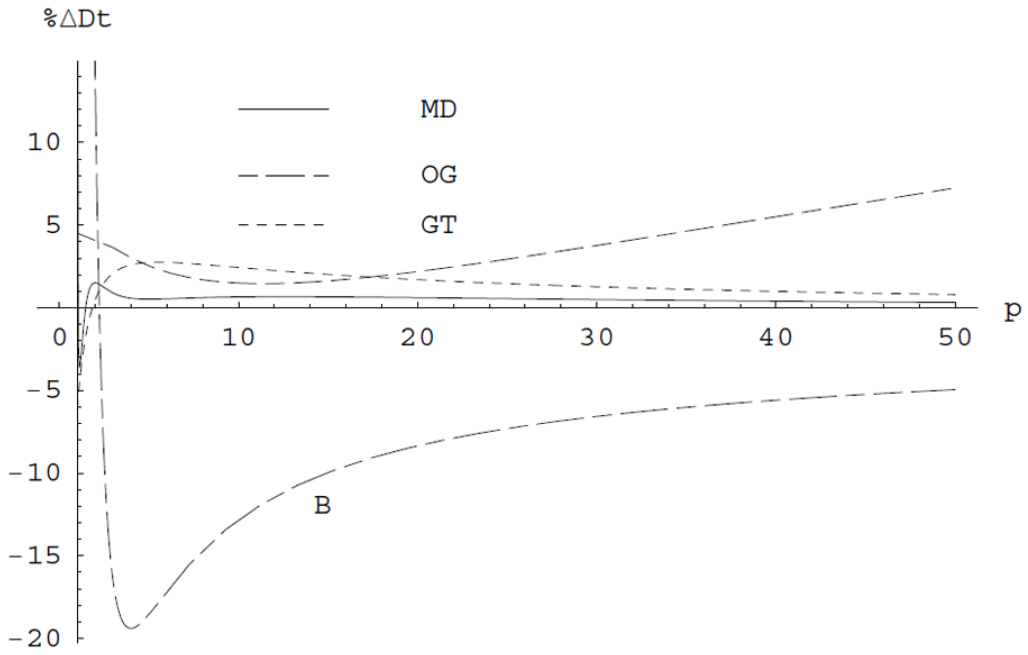

Fig. 2.

The percent difference for the average translational diffusion tensor of literature values to our accurate formulas as a function of axial ratio for the rectangular cylinder is shown. The symbols represent the works of Mansfield and Douglas [30] (MD), Ortega & Garcia de la Torre [31] (OG), Tirado and Garcia de la Torre [32] (TG), and Broersma [33] (B). Taken from reference 29.

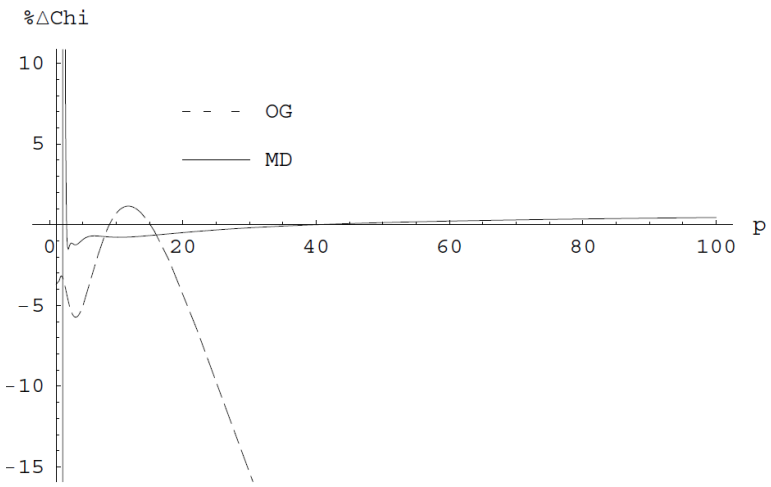

Fig. 3.

The % difference for the viscosity factor between our results and the literature for the rectangular cylinder as a function of axial ratio is shown. The symbols represent the referenced works of Ortega and Garcia de la Torre [31] (OG) and Mansfield and Douglas [30] (MD). Taken from reference 29.

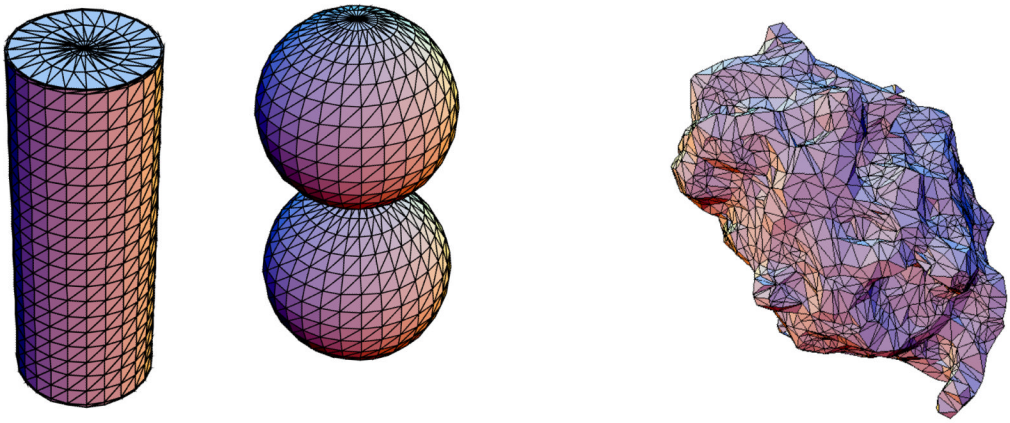

Figure 4.
Triangular tessellations of a rectangular cylinder, the dumbbell, and lysozyme. Note that the sharp edge of the cylinder is well defined by placing small triangles on both edge surfaces, and the polar tessellation of a dumbbell places small triangles where the beads touch

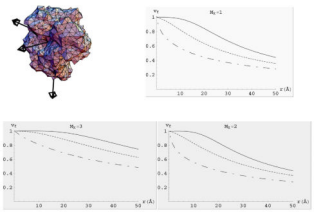

Fig. 5.
Fluid stagnation in Albumin (1AO6). The solid line represents a deep pocket in the middle of the protein, the dotted line represents a medium pocket at top left of the protein, while the dashed line represents a triangle at the bottom of the protein, where there is no pocket. Mx =3 represents protein motion parallel to the vector, while Mx =1,2 constitutes motion perpendicular to the vector. Taken from reference 42.

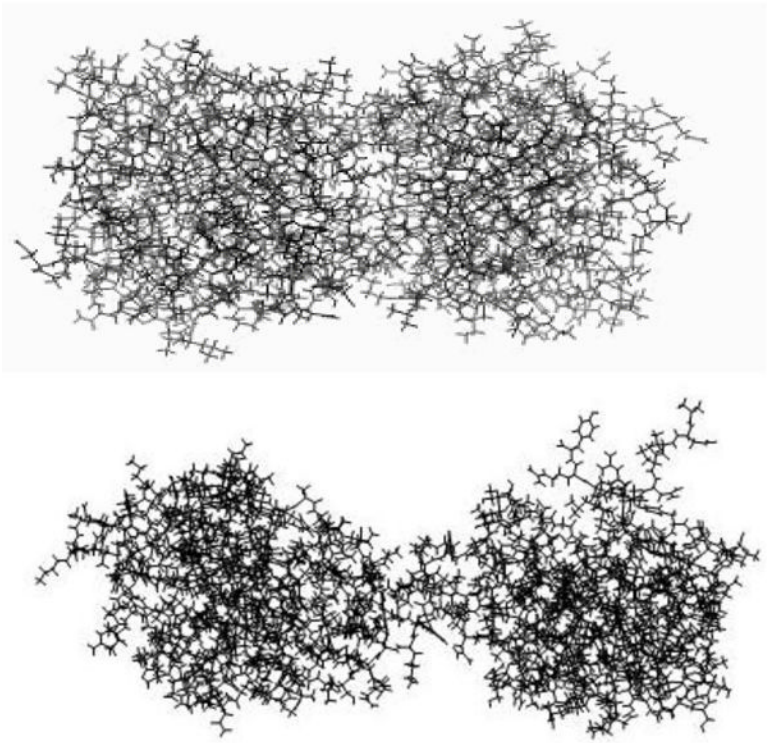

Figure 6.
α-chymotrypsin structures. Top Panel: crystal structure (4cha.pdb). Bottom Panel: Amber 9 typical geometry after 1 ns molecular dynamics with implicit solvent.

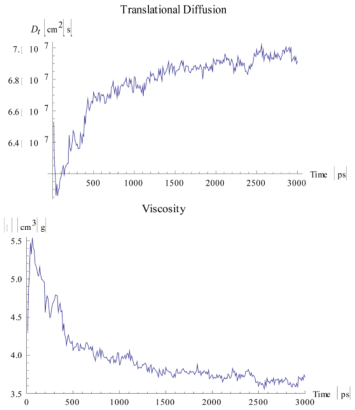

Fig. 7.

The translational diffusion coefficient and the intrinsic viscosity of α-chymotrypsin from an MD trajectory with implicit water at pH 3.0. As the molecule shape deforms from the initial crystal structure, the transport properties evolve and settle down after 2 ns.

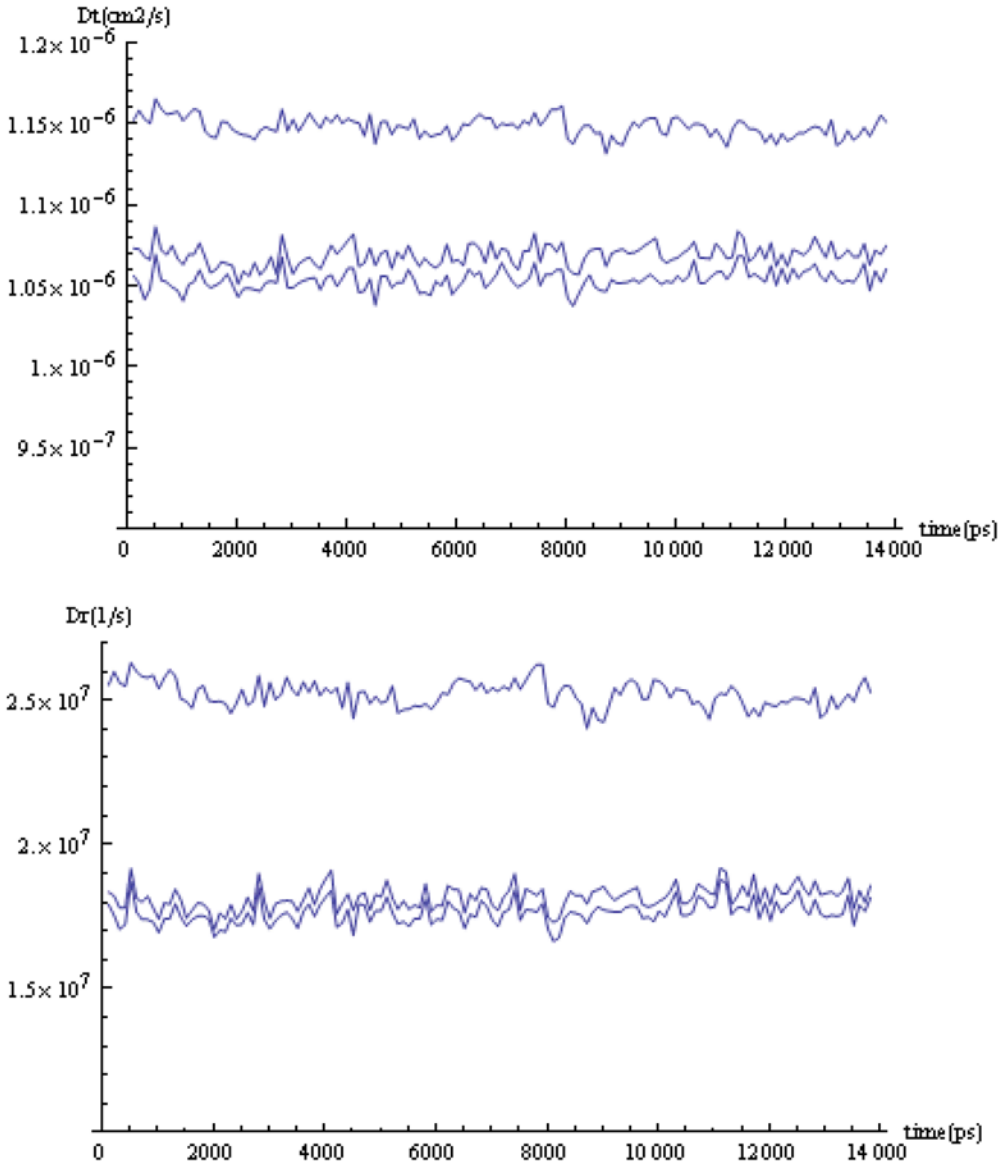

Fig. 8.

Top panel: The translational diffusion tensor eigenvalues along the MD trajectory for lysozyme (6LYZ). Note the small difference between the eigenvalues, justyfing the use of the average. Bottom Panel: The rotational diffusion tensor eigenvalues along the MD trajectory. Note the symmetric top appearance of the eigenvalues. In both cases, the data shows only small thermal fluctuations characteristic of a globular protein.

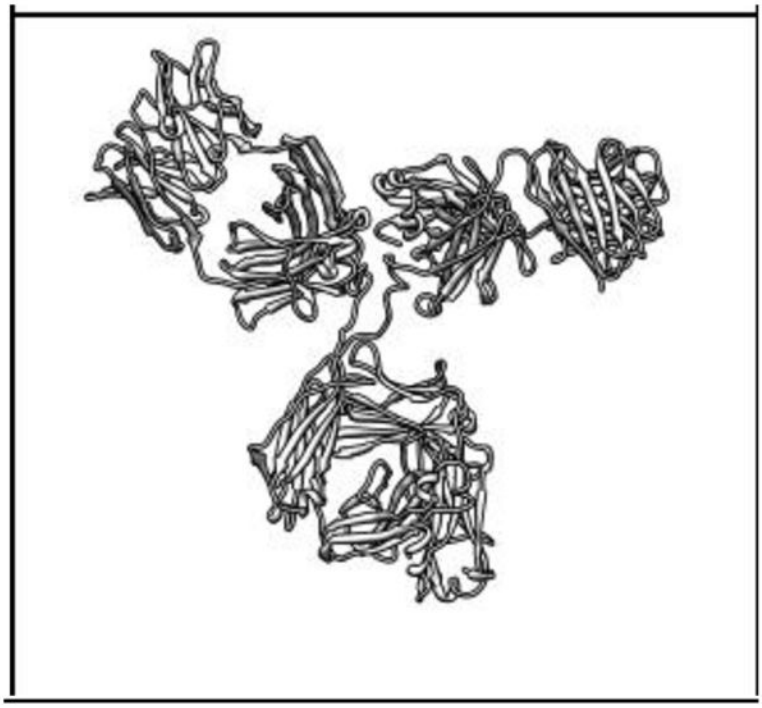

Fig. 9.

Ribbon structure of trastuzumab taken from one of the multiple MD trajectories [64]. The carbohydrate group has been removed from the lower FC region for clarity. The structural fluctuation in this flexible antibody can be appreciated in the movie mentioned in the text.

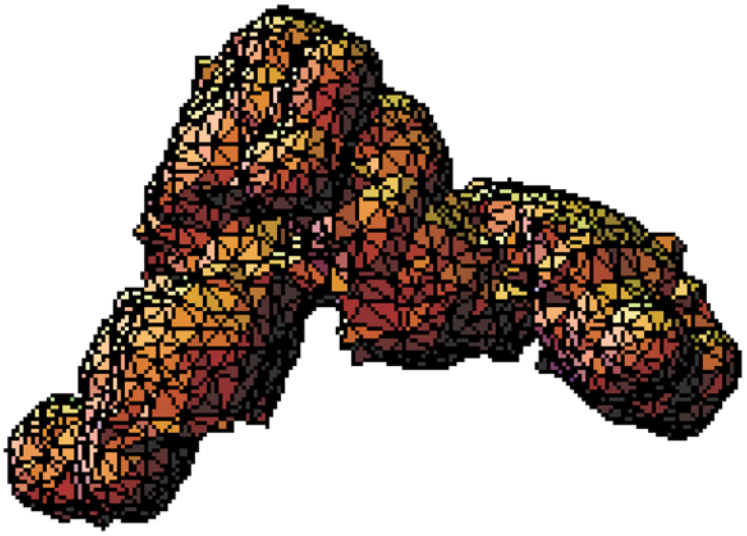

Fig. 10.
Phenyl-t-RNA (1EVV) surface with 2908 triangles.

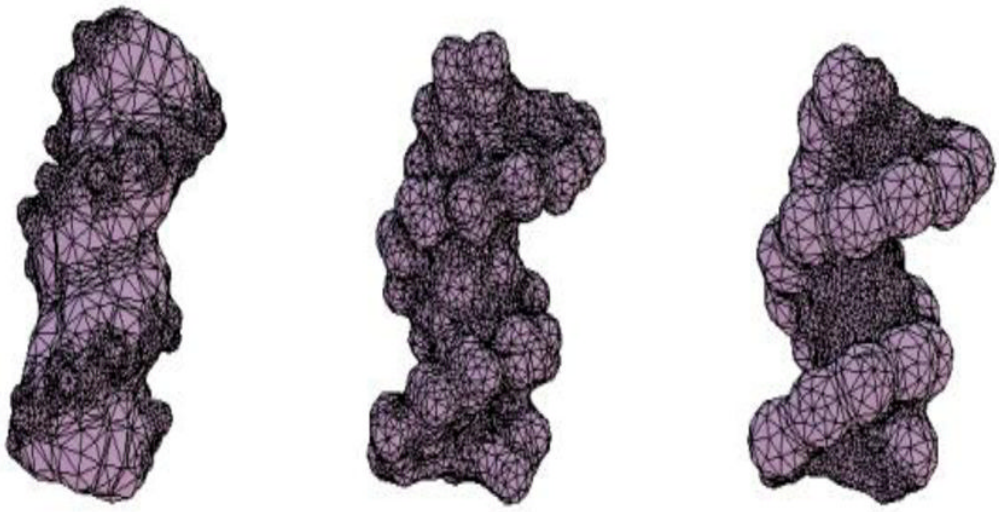

Fig. 11.

Left Panel: Groove hydration produced by inflated the nitrogen atoms by 3.6 A. Middle Panel: Mixed groove and backbone hydration produced by inflating the oxygen atoms. Right Panel: Pure backbone hydration produced by inflation of phosphorous atoms.

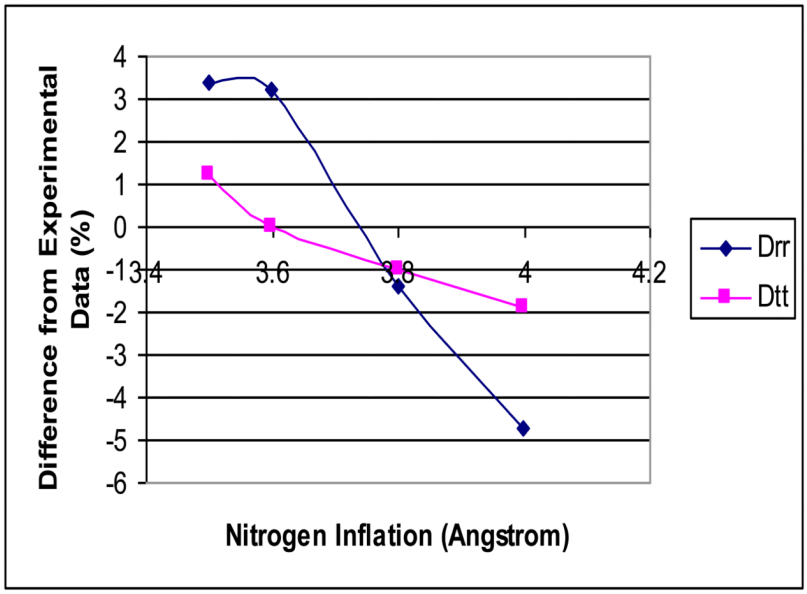

Fig. 12.

The discrepancy of the BEST computation for both translation (Dtt) and tumbling (Drr) as a function of the added radius of the nitrogen atom which is present only in the DNA bases. The best value is that which splits the error across zero between each measurement. In this case the value occurs at 3.65 A for the DNA 12mer.

Table I

The transport properties of ellipsoids of revolution as a function of axial ratio, p: the viscosity factor, ξ, and the perpendicular and parallel components of the translational and rotational diffusion tensor. The % columns indicate the error by comparison with the analytic formulas. Oblate has p > 1, prolate has p < 1. Dtt has units of 1/A and kT/8πηa has been factored out, while Drr has units of $1/A^3$ and $kT/8\pi\eta a^3$ has been factored out – "a" is the radius of the circular circumference of the ellipsoid.

p	ξ	%	Dtt⊥	%	Dtt$^{\|\|}$	%	Drr⊥	%	Drr$^{\|\|}$	%
1/8	10.1023	−0.003	0.40954	0.0007	0.57596	0.04	0.013296	0.008	0.18224	0.04
1/4	4.66306	−0.006	0.64839	0.006	0.83443	0.003	0.073631	0.01	0.34671	0.003
1/2	2.90751	−0.004	0.96702	0.007	1.10782	0.03	0.3323	0.02	0.61999	0.03
1	2.50004	0.002	1.33356	0.02	1.33356	0.02	1.0003	0.03	1.0003	0.03
2	2.85443	0.002	0.84104	0.004	0.73655	0.02	0.22105	0.06	0.17728	0.01
4	4.05940	0.002	0.48853	0.004	0.38441	0.02	0.033933	0.07	0.027774	0.03
8	6.70083	0.009	0.26664	0.01	0.19514	0.02	$4.5054\ 10^{-3}$	0.07	$3.9623\ 10^{-3}$	0.04

Table II

Analysis of viscosity raw data

Parameter	Value	Standard Error	TStat
1	104.579	0.00111243	94009.5
x	−2353.82	19.5104	−120.64
x^2	1.57524×10^6	71695.9	21.97

RSquared = 1.0 and Estimated Variance = $1.01\ 10^{-7}$

Table III

The intrinsic viscosity and translational diffusion coefficient of multimeric proteins from the crystal structure. References for experimental work are available in the original paper [28].

Protein	s[a]	Mass (kDa)	[η] (cm^3/g)			$D_t(10^{-7}cm^2/s)$		
			Calc.	Exp.	Δ[b]	Calc.	Exp.	Δ[b]
Superoxide Dismu. (2SOD)	2	32.5	3.57	3.3	9	8.10	8.27	−2
β-Lactoglobulin (1BEB)[c]	2	36.7	3.65	3.4 – 4.2	−5	7.74	7.3	5
α-Chymotrypsin (4CHA)	2	50.4	3.31	4.1, 4.25	−21	7.17	7.1, 7.40	−1
Concanavalin (1GKB)	2	51.4	3.95	4.1	−2	6.72	6.2	8
Triosephos. Isom. (8TIM)	2	53.2	3.59	3.75	−4	6.88	6.76	2
Ricin (2AAI)	2	61.5	3.33	2.96	13	6.61	6.0	10
Oxyhemoglobin A (1HHO)	4	63.2	2.89	2.77	4	7.03	6.78	4
Alkaline Phosphat. (1ALK)	2	94.6	3.09	3.4	−7	5.96	5.7	4
Citrate Synthase (1CTS)	2	98.0	3.20	3.95	−20	5.82	5.8	0
Inorganic Pyrophos. (1FAJ)	6	117.3	2.93	4.0	−28	5.62	5.7	−2
Aldolase (1ADO)	4	157.1	3.84	3.4, 4.0, 4.04	0	4.66	4.29 – 4.8	4
Catalase (4BLC)	4	235.7	3.08	3.9	−21	4.49	4.1	10
β-Galactosidase (1BGL)	4	465.8	3.84	3.78	2	3.26	3.13	4

Table IV

The Intrinsic Viscosity and Translational Diffusion Coefficient of α-Chymotrypsin (4cha) and β-lactoglobulin (1beb) from implicit water MD.

Geometry	n	Mass (kDa)	[η] (cm^3/g)			$D_t(10^{-7}cm^2/s)$		
			Calc.	Exp.	Δ	Calc.	Exp.	Δ
1beb.pdb	1	36.7	3.7	4.1[43]	−7.5	7.7	7.3[44]	5.5
Sander			**4.0**		−2.4	**7.5**		2.7
4cha.pdb	2	49.7	3.3	4.1[45]	−19	7.2	7.1[45]	1
Sander			**3.7**		−10	**6.9**		−3

Table V

Transport properties for lysozyme, ribonuclease, and ubiquitin. References to experimental work are available in reference [46]

Protein			Translational diffusion	Rotational diffusion tensor		
		Intrinsic Viscosity (cm3/g)	Dt (10^{-6} cm^2/s)	Dr_1 (10^7 1/s)	Dr_2 (10^7 1/s)	Ave. Dr (10^7 1/s)
Lysozyme (6lyz) 14.3kDa	MD TIP3P	3.33 ± 0.01	1.09 ± 0.001	1.79 ± 0.01	2.52 ± 0.01	2.03 ± 0.01
	Crystal	3.21	1.10	1.82	2.57	2.07
	Exp.	2.98~3.00	1.11±0.05 1.12±0.02* 1.06±0.01*	------	------	1.67±0.08 2.0
Ribonuclease (7rsa) 13.7kDa	MD TIP3P	3.59 ± 0.01	1.077 ± 0.002	1.74 ± 0.01	2.34 ± 0.02	1.94 ± 0.01
	Crystal	3.39	1.10	1.83	2.50	2.05
	Exp.	3.30–3.50	1.068	------	------	2.01
Ubiquitin (1ubq) 6kDa	MD TIP3P	3.54 ± 0.02	1.267 ± 0.003	2.75 ± 0.03	3.97 ± 0.02	3.16 ± 0.02
	MD TIP4P	3.57 ± 0.02	1.263 ± 0.002	2.73 ± 0.02	3.90 ± 0.02	3.16 ± 0.02
	Crystal	3.31	1.270	2.80	3.98	3.19
	Exp.		1.30 ± 0.01	3.17 ± 0.17	3.70 ± 0.14	3.34 ± 0.16

Table VI

Summary of hydrodynamic analysis of trastuzumab MD simulation data, experimental hydrodynamic results, and results from literature; all values correspond to a temperature of 20°C in pure water

Trajectory	D_t ($\times 10^{-7}$ cm^2/s)*†	T_r (ns)‡†	[η] (cm^3/g)†
1	4.00 (±0.04)	184 (± 9)	6.6 (±0.2)
2	4.12 (±0.05)	168 (± 8)	6.0 (±0.3)
3	4.15 (±0.06)	158 (± 10)	6.1 (±0.3)
4	3.97 (±0.03)	190 (± 5)	6.7 (±0.1)
5	4.05 (±0.03)	179 (± 5)	6.4 (±0.2)
6	4.11 (±0.03)	168 (± 5)	6.0 (±0.1)
7	4.17 (±0.07)	164 (± 10)	5.9 (±0.4)
8	4.09 (±0.04)	172 (± 8)	6.2 (±0.2)
AVERAGE§	4.08 (± 0.07)	173 (± 11)	6.24 (± 0.3)
EXPERIMENT¶	4.09 (± 0.01)		6.37 (± 0.2)
LITERATURE‖		168, 180	6.20 (±0.5)

*for computational results: $D_t = 1/3\ Tr(\mathbf{D}_{tt})$

†for individual trajectories, value quoted is the trajectory average and its standard deviation

‡for computational results: $T_r = (6D_r)^{-1}$, where $D_r = 1/3\ Tr(\mathbf{D}_{rr})$

§values quoted are the average of each trajectory's average and their standard deviation

¶see methods section; uncertainties quoted are the standard error of extrapolations to c=0

‖See reference 53 for references to experimental values. Experimental $\mathbf{T_r}$ values are for rabbit IgG and bovine IgG; Intrinsic viscosity values values are for human IgG1.