Biophysical Journal
Article

Biophysical Society

HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties

Patrick J. Fleming[1] and Karen G. Fleming[1,*]
[1]T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, Maryland

ABSTRACT Hydrodynamic properties are useful parameters for estimating the size and shape of proteins and nucleic acids in solution. The calculation of such properties from structural models informs on the solution properties of these molecules and complements corresponding structural studies. Here we report, to our knowledge, a new method to accurately predict the hydrodynamic properties of molecular structures. This method uses a convex hull model to estimate the hydrodynamic volume of the molecule and is orders of magnitude faster than common methods. It works well for both folded proteins and ensembles of conformationally heterogeneous proteins and for nucleic acids. Because of its simplicity and speed, the method should be useful for the modification of computer-generated, intrinsically disordered protein ensembles and ensembles of flexible, but folded, molecules in which rapid calculation of experimental parameters is needed. The convex hull method is implemented in a Python script called HullRad. The use of the method is facilitated by a web server and the code is freely available for batch applications.

INTRODUCTION

Hydrodynamic properties such as sedimentation and diffusion coefficients and their related hydrodynamic radii are useful parameters to estimate the size and shape of biological molecules in solution (1). Experimental methods such as analytical ultracentrifugation (2), small-angle x-ray scattering (3), dynamic light scattering (4), fluorescence polarization (5), size exclusion chromatography (6), and NMR (7) have all been used to measure these hydrodynamic properties experimentally. Hydrodynamic properties may also be predicted for three-dimensional structural models of proteins and nucleic acids. Comparisons of these predicted and experimentally determined hydrodynamic properties are useful to probe the solution properties of crystalline structures.

There are several methods available to predict the hydrodynamic properties of biological molecules from three-dimensional models: a boundary element method (BEST) that uses triangulation of the solvent accessible surface (8); bead-modeling methods (HYDROPRO, SOMO, AtoB), as reviewed in (9,10); methods based on the construction of an ellipsoidal shell (ELLIPS, ELM) (11,12); and a numerical path integration method (ZENO) (13). The major disadvantage of the above methods is slow speed. The HYDROPRO program that implements a bead modeling algorithm has been most often used to calculate hydrodynamic properties from protein structures (14). Several of these methods have been integrated into the UltraScan analytical ultracentrifugation analysis software suite (15).

The method described here for predicting hydrodynamic properties is orders of magnitude faster than other methods. The method has comparable accuracy with commonly used methods, it is suitable for both translational and rotational diffusion, it does not require runtime configuration files, it works for both folded and disordered proteins and for nucleic acids, and it accepts the largest known molecular structures without modification of input parameters. The new algorithm uses a convex hull to model the hydrodynamic volume of a molecular structure. The convex hull is a three-dimensional mathematical construct related to Voronoi diagrams. It is the smallest convex envelope that contains a set of points. It has been applied to areas such as cluster analysis, collision detection for robotic systems, and crystal growth, as well as many other mathematical problems (16). It also has been used previously to estimate

Submitted August 29, 2017, and accepted for publication January 2, 2018.
*Correspondence: karen.fleming@jhu.edu
Editor: David Eliezer.
https://doi.org/10.1016/j.bpj.2018.01.002

the particle size and surface roughness of aggregates and nanomaterials (17,18). The convex hull of a three-dimensional molecular model encapsulates both the overall volume of the molecule as well as its shape asymmetry and surface roughness; all characteristics that influence hydrodynamic diffusion.

The convex hull method is implemented in a simple Python script called HullRad. The program code is made freely available and the method may also be accessed through a web server.

METHODS

Convex hull transport model

A convex hull is mathematically defined as the smallest convex envelope that contains a set of points. We implemented calculation of the convex hull to model the hydrodynamic behavior of a molecular structure in a Python script entitled HullRad. HullRad reads coordinates of a standard Protein Data Bank (PDB) file as input; makes a reduced atom model of the structure (detailed below); calls an open source C-code version of Qhull (19), which returns the convex hull volume and surface area from the atom center point set; expands the initial convex hull volume to account for hydration (discussed below); calculates a shape factor correction based on an ellipsoid of revolution; and finally calculates the following hydrodynamic parameters: R_0, R_T, R_R, s, D_T, D_R, f/f_0, and $[\eta]$, all defined as follows.

R_0 is the radius in Å of an equivalent sphere corresponding to the volume of the anhydrous molecule (Eq. 1),

$$R_0 = R_{0,\mathrm{calc}} = \sqrt[3]{\frac{3V_0}{4\pi}}, \tag{1}$$

where V_0 equals the anhydrous molecular volume in Å^3 defined as

$$V_0 = \frac{10^{24} M\bar{v}}{\mathrm{N_A}}, \tag{2}$$

where M is the molar mass (g mol^{-1}), $\mathrm{N_A}$ is Avogadro's number, and $\bar{v}$ is the partial specific volume (ml g^{-1}). The anhydrous partial specific volumes of amino acids were taken from those reported by Cohn and Edsall (20) as corrected by Perkins (21); the volumes of nucleic acids are as reported in (22) and (23).

$R_{T,calc}$ and $R_{R,calc}$ (Å) are the translational and rotational hydrodynamic radii calculated from a sphere with the equivalent volume (V_{TH} or V_{RH}) as the expanded convex hull for a particular molecule (described below) and corrected by the translational or rotational shape factors F_T and F_R, respectively.

$$R_{T,calc} = F_T \times \sqrt[3]{\frac{3V_{\mathrm{TH}}}{4\pi}}, \text{ and} \tag{3}$$

$$R_{R,calc} = F_R \times \sqrt[3]{\frac{3V_{\mathrm{RH}}}{4\pi}}. \tag{4}$$

The hull volume expansion to account for hydration is different for translational diffusion and rotational diffusion as described below. The frictional ratio, f/f_0, is defined as $R_{T,calc}/R_0$.

A translational diffusion shape factor, F_T, is calculated as follows: the major axis (a) of a prolate ellipsoid of revolution is estimated from the maximum dimension of the initial convex hull of a molecule. The minor axis is calculated from Eq. 5, assuming the volume of a prolate ellipsoid is $4\pi ab^2/3$.

$$b = \sqrt[2]{\frac{3V_H}{4\pi a}}, \tag{5}$$

where V_H is the equivalent volume of the initial convex hull for a particular molecule. From the axial ratio, a/b, a translational shape factor based on Perrin's equation (24) is obtained. A prolate ellipsoid of revolution is used to model the nonsphericity of all molecules. For molecules with axial ratios less than 5:1, the translational shape factors for prolate and oblate ellipsoids of revolution are similar within 2%. Rotational shape factors are more complicated because the effect of shape depends on the axis of revolution for an ellipsoid and the overall tumbling is characterized by a rotational diffusion tensor (25). In addition, rotational shape factors are significantly larger than translational shape factors (1). However, as described below and in the Supporting Material, we empirically derived a rotational shape factor $F_R = F_T{}^4$ that accurately predicts the overall rotational hydrodynamic properties of elongated molecules up to an axial ratio of 2.62, the most nonspherical molecule in our data set.

The sedimentation coefficient, s, (in units of seconds) is calculated at 20°C in water. HullRad predicts this value from $R_{T,calc}$ using the Svedberg equation, Eq. 6:

$$s = s_{calc} = 10^8 \left(\frac{M - M\bar{v}\rho_{20,w}}{N_A 6\pi\eta_0 R_{T,calc}} \right), \tag{6}$$

where $\rho_{20,w}$ is the water density at 20°C (g mL^{-1}), N_A is Avogadro's number, η_0 is water viscosity at 20°C (poise), the 10^8 term converts $R_{T,calc}$ from Å to cm to convert the answer to the correct order of magnitude for the sedimentation coefficient, and other terms are as previously defined.

D_T (cm^2 s^{-1}) and D_R (s^{-1}) are the translational and rotational diffusion coefficients calculated from their respective hydrodynamic radii as follows:

$$D_T = D_{T,calc} = \frac{k_B T}{6\pi\eta_0 R_{T,calc}}, \text{ and} \tag{7}$$

$$D_R = D_{R,calc} = \frac{k_B T}{8\pi\eta_0 R^3_{R,calc}}, \tag{8}$$

where k_B is Boltzmann's constant and other terms are as previously described. The temperature (T) employed in these calculations is 20°C because this is the standard used for hydrodynamic modeling.

The intrinsic viscosity, $[\eta]$, is defined by Eq. 9.

$$[\eta] = \frac{10^{-23} N_A \pi R^3_{T,calc}}{3M}. \tag{9}$$

By mathematical definition, the vertices of the convex hull are placed at the outermost surface atom centers. This has several implications for subsequent hydrodynamic calculations. First, an individual structure file contains a static, snapshot representation of the atomic coordinates. For proteins, the specific side chain conformations encoded by the PDB file will influence the corresponding convex hull for a particular protein. In contrast, side chain motion in solution would lead to an ensemble of side chain conformations and corresponding variations in the positions of the convex hull vertices. These alternate positions in turn lead to variations in the hull volumes. To average surface-exposed side chain conformations represented in a particular protein coordinate file, each side chain is represented by a single, unified pseudoatom located along the Cα – Cβ vector. This pseudoatom is displaced outward by a distance equal to the radius of a sphere corresponding to the respective side chain volume. To maintain the fold represented in the original PDB file, all

backbone atoms, including those that are surface-exposed, are unchanged. It is this latter pseudoatom model from which the initial convex hull is created. Images of initial convex hulls shown in the figures were created by using PyMOL (26).

To improve computational speed, a reduced atom model of nucleic acids is made by using only the nitrogen and oxygen atoms of the molecule. The difference in the volumes of convex hulls made with the reduced atom nucleic acid models compared to models that also contained carbon and phosphorus atoms is small. This is because almost all vertices of the convex hull are determined by either phosphate oxygen atoms or base nitrogen atoms in nucleic acid structures. The volume difference for DNA duplexes was 0.5%, for transfer RNA it was 0.8%, and for 50 s ribosomal subunit it was 0.01%. Interestingly, the sign of the difference was dependent on the structure, and a variation of this magnitude is consistent with the averaging effect of pseudoatom side chains in proteins.

A second issue for hydrodynamic modeling is that proteins and nucleic acids are hydrated in solution. This aspect is not taken into account when calculating the initial convex hull above because the input structural model is anhydrous. To account for hydration, the initial convex hull volume is increased by expanding the planes of the hull outward. As with other commonly used hydrodynamic modeling algorithms, the extent of expansion was empirically determined by optimizing agreement of the calculated parameters with experimental hydrodynamic data. The resulting total expanded hull volume is assumed to be the hydrodynamic volume of the molecule. The hydrodynamic radii are calculated from a sphere of volume equal to that of the expanded convex hull corrected by a shape factor as described below. Because hydration is manifested differently in translational and rotational transport, separate hull expansions were empirically determined and optimized for R_T and R_R as described in Results and Discussion.

To take full advantage of experimental hydrodynamic data, we used jackknife resampling (27) of the translational protein data set to test for bias caused by specific proteins during optimization of the hull expansion for hydration. Following this procedure, each protein in the data set in turn was omitted during determination of the optimal hull expansion. The R_T of the omitted protein was calculated by using the optimal hull expansion value determined by using all other proteins in the set. This calculated $R_{T,calc}$ was compared to experimental $R_{T,exp}$ to estimate the independent error.

Code execution

The HullRad method is available as a web server (http://www.hullrad.wordpress.com), and the Python script will be made freely available on-line upon publication. Qhull is freely available at http://www.qhull.org. The HullRad code is designed to be user-friendly from the command line and is suitable for batch calculation of multiple structures. Neither preparation of a configuration file nor manipulation of the molecular structure file is necessary. The same fitted parameters are used for all molecular sizes. A simple command with a PDB file name as input results in output of the format shown in Fig. 1 (PDB code: 5PTI shown for example).

An option for code execution includes output of the reduced atom structure used in construction of the convex hull as a PDB file. The README file in the code distribution contains instructions for installation, execution, and convex hull image generation. HullRad has been tested on Windows, Mac, and UNIX platforms. Prior installation of Python 2.7 is required.

Transformation of experimental coefficients

For ease of comparison between experimental values obtained by using different methods, all experimental hydrodynamic data were converted to equivalent hydrodynamic radii. Sedimentation coefficients were converted to equivalent $R_{T,exp}$ (in Å) according to the Svedberg relationship in Eq. 10. These values were provided in the literature at the standard of 20°C and pure water.

```
pdb5pti.ent
PROTEINASE INHIBITOR (TRYPSIN)  05-OCT-84   5PTI
 M                :         6518      g/mol
 v_bar            :        0.719      mL/g
 R(Anhydrous)     :        12.30      Angstroms
 Axial Ratio      :         1.66
 f/fo             :         1.16
 Dt               :      1.50e-06     cm^2/s
 R(Translation)   :        14.25      Angstroms
 s                :      1.13e-13     sec
 [eta]            :         2.80      cm^3/g
 Dr               :      4.05e+07     s^-1
 R(Rotation)      :        15.83      Angstroms
```

FIGURE 1 Sample output for execution of HullRad using the PDB file for bovine pancreatic trypsin inhibitor.

$$R_{T,exp} = 10^8\left(\frac{M - M\bar{\nu}\rho_{20,w}}{N_A 6\pi\eta_0 s_{20,w}}\right), \tag{10}$$

where the 10^8 term converts the value from cm to Å and other terms are as previously defined.

Experimental translational diffusion coefficients, D_T, were converted to equivalent $R_{T,exp}$ (in Å) according to Eq. 11.

$$R_{T,exp} = 10^8\left(\frac{k_B T}{6\pi\eta_0 D_T}\right). \tag{11}$$

Similarly, protein rotational correlation times were converted to equivalent $R_{R,exp}$ in Å by using the following relationship,

$$R_{R,exp} = 10^8\left(\sqrt[3]{\frac{3k_B\tau_c}{4\pi\eta_0}}\right), \tag{12}$$

where τ_c is the overall rotational correlation time (seconds) assuming isotropic rotation, and the other terms are as previously defined. Correlation times obtained by NMR are typically reported for the conditions of data collection and were converted to 20°C by using the known dependence of water viscosity on temperature by multiplying the experimentally observed τ by the factor $\eta_0/\eta_{T_{exp}}$ (28,29).

As a measure of protein backbone expansion of intrinsically disordered proteins (IDPs) (see Fig. 7), we used the radius of gyration calculated by Eq. 13.

$$R_{G,calc} = \sqrt{\frac{\sum_1^N (m_i r_i - COM)^2}{\sum_1^N m_i}}, \tag{13}$$

where N is the number of atoms, m_i represents the mass of each atom, r_i represents the Cartesian position of each atom and COM is the center of mass of the protein. This is an anhydrous, geometric R_G of the protein atoms alone and is not the same as the radius of gyration determined by scattering experiments. The radius of gyration determined by x-ray scattering includes effects from waters of hydration (30) and requires a consideration of explicit solvent positions for accurate calculation (31). The convex hull method is an empirical method that does not consider explicit waters of hydration and therefore does not predict experimental R_G.

Bead modeling calculations

Hydrodynamic parameters were independently calculated by using the program HYDROPRO Version 10 (14). For most structures, the atomic-level primary model was used with default values of 2.9 Å for the radius of the elements in the primary model and $\sigma = 6$. For the large complex of GroEL (M = 828,989) the atomic level calculation was unsuccessful, and we therefore used the residue-level primary model with a default radius of 4.8 Å. We used the default option to estimate σ by the program. Residue level calculations were also carried out for the benchmarking shown in Fig. 4. We used the default temperature and viscosity values of 20°C and 0.01 poise, respectively.

Protein data sets

Folded protein structures from previous hydrodynamic property predictions (12,32) were used as a starting point to assemble the sets of proteins employed. Additionally, larger proteins as well as proteins for which new experimental data were available were added to the set. When possible, updated experimental results were used. Some proteins were discarded if the confidence in experimental values was not high. Some structures used for translational property calculations had missing residues built in (Table 1), and these structures were generously provided by Olwyn Byron and Mattia Rocco. In the case of rotational diffusion, only proteins for which the rotational diffusion was determined by using NMR methods were considered. For DNA structures, four different lengths of duplexes (8–24 basepair), for which hydrodynamic properties have been published (33–36) were used for code validation. RNA molecules in isolation are flexible (37,38), so we used two different protein-RNA complexes—histidinyl-tRNA synthetase (39,40) and bacterial 50s ribosomal subunit—to test the method on known structures containing RNA that are expected to be less flexible.

These procedures resulted in three final data sets of folded structures to validate the convex hull method: one data set for which translational hydrodynamic properties of proteins have been experimentally determined (Table 1), one data set for which rotational properties of proteins have been experimentally determined (Table 2), and one data set of DNA duplexes and RNA-protein complexes for which several types of experimental data are available (Table 3). In addition, the largest current protein model in the PDB (HIV-1 capsid, PDB code 3J3Q) was used to test for any size limitations.

For analysis of conformationally heterogeneous proteins, a disordered state ensemble of α-synuclein (accession number PED9AAC) was

TABLE 1 Translational Hydrodynamic Radii for Folded Proteins

Protein (PDB ID)	M (Da)	Chains	$R_{T,\mathrm{exp}}$[a] (Å)	$R_{T,\mathrm{calc}}$ (Å)	% Signed Error	% Error (Jackknifed)	a/b[b]	Reference(s)
BPTI (5PTI)	6518	1	14.5	14.3	−1.61	−3.41	1.66	(65)
Cytochrome *c* (1HRC)	11,703	1	17.3[c]	17.2	−0.39	−0.56	1.16	(66–68)
RNase A (8RAT)	13,692	1	18.9	19.0	+0.37	+0.27	1.31	(65)
Lysozyme (1AKI)	14,315	1	18.6	18.5	−0.54	−0.75	1.51	(65)
LegHb (1LH1)	16,657	1	21.4	20.2	−5.72	−5.86	1.43	(69)
Apo-Mb (1MGN)	17,359	1	20.8	19.9	−4.19	−4.52	1.35	(70)
Soy TI (1AVU)[d]	19,984	1	22.6	21.9	−3.21	−3.43	1.10	(71)
Trypsinogen (1TGN)	22,620	1	22.1	22.1	−0.11	−0.25	1.15	(72)
β-trypsin (1TPO)	23,124	1	22.8	22.2	−2.74	−2.96	1.13	(73)
α-chymotrypsin (4CHA)[d]	25,041	1	22.7	22.9	+0.80	+0.72	1.15	(71)
Chymotrypsinogen (2CGA)[d]	25,670	1	22.6	23.3	+2.91	+2.91	1.08	(74)
Carbonic anhydrase (2CAB)[d]	28,757	1	24.1	23.9	−0.71	−0.84	1.36	(75)
Superoxide dismutase (2SOD)[d]	31,089	2	25.9	25.8	−0.58	−0.73	1.59	(76)
Pepsin (4PEP)[d]	34,516	1	24.6	26.2	+6.26	+6.38	1.37	(77)
β-lactoglobulin (1BEB)[d]	36,606	2	27.4	27.3	−0.52	−0.66	1.83	(78)
TPI (8TIM)	52,996	2	29.7	30.5	+2.42	+2.39	1.45	(66)
Hb (CO) (1HCO)	61,942	4	31.5	31.6	+0.24	+0.15	1.11	(79)
HSA (1AO6)	66,482	1	34.0	35.1	+3.21	+3.21	1.49	(80)
Alkaline phosphatase (1ALK)	94,082	2	37.6	36.5	−2.94	−3.10	1.67	(81)
Citrate synthase (1CTS)[d]	97,835	2	37.0	37.7	+2.06	+2.03	1.53	(82)
Inorganic PPase (2AU9)	1,17,361	6	37.6	39.4	+4.72	+4.77	1.23	(81)
Trp. synth. (1KFK)	1,38,595	4	44.4	45.8	+3.11	+3.11	2.20	(83)
Human IgG (1HZH)	1,433,37	4	55.1	54.9	−0.37	−0.44	2.07	(84)
Apo-G3PD (2GD1)[d]	1,43,743	4	42.9	43.3	+1.00	+0.95	1.16	(85)
Apo-LDH (5LDH)[d]	1,45,749	4	42.5	42.5	−0.07	−0.14	1.20	(86)
Aldolase (1ADO)[d]	1,56,776	4	47.6[c]	47.7	+0.12	+0.03	1.30	(87–89)
Holo calalase (4BLC)[d]	2,30,321	4	52.3	52.4	+0.17	+0.11	1.47	(90)
Xanthine oxidase (1FIQ)	2,67,770	6	54.5	54.5	−0.14	−0.21	1.78	(91)
β-galactosidase (4V40)	4,64,490	4	68.5	67.1	−2.02	−2.10	1.63	(92)
Apoferritin (3AJO)	5,120,84	24	67.4	64.7	−4.00	−4.11	1.00	(93)
Urease (3LA4)	5,397,00	6	65.8[c]	66.6	+1.26	+1.23	1.13	(94,95)
GroEL (2CGT)	8,28,989	21	82.8	86.3	+4.26	+4.27	1.40	(96)
YMV capsid (1AUY)	3,456,960	60	151.2	152.4	+0.28	+0.28	1.06	(97)
			mean % signed error		+0.10	−0.04		

Hb, hemoglobin; HSA, human serum albumin; IgG, immunoglobulin G; TI, trypsin inhibitor; TPI, triose phosphate isomerase; Trp. Synth., tryptophan synthase; YMV, Yellow Mosaic Virus.

[a]All experimental hydrodynamic coefficients as reported in the references were converted to R_T for comparison.

[b]Axial ratio of a prolate ellipsoid with volume equal to the initial convex hull volume.

[c]Average value of those reported in cited publications.

[d]PDB files for these structures provided by Mattia Rocco and Olwyn Byron. Missing residues were modeled as described in (32).

TABLE 2 Rotational Hydrodynamic Radii for Folded Proteins

Protein (PDB ID)	M (Da)	$R_{R,exp}$[a] (Å)	$R_{R,calc}$ (Å)	% Signed Error	a/b[b]	Reference
BPTI (5PTI)	6518	15.3	15.9	+3.73	1.66	(65)
Calbindin (1IG5)	8502	17.0	16.6	−2.29	1.10	(98)
Ubiquitin (1UBQ)	8566	16.8	17.3	+2.68	1.49	(7)
Cytochrome *c* (351C)	8696	17.1	16.8	−2.22	1.10	(99)
Plastocyanin (1PCS)	10,322	18.0	17.6	−2.06	1.19	(100)
Binase (1GOU)	12,214	19.7	19.3	−2.03	1.27	(101)
Ribonuclease A (1AQP)	13,692	20.0	20.4	+1.65	1.34	(65)
Azurin (1E5Y)	13,948	19.7	19.7	+0.25	1.31	(102)
Lysozyme (1AKI)	14,315	19.2	20.1	+5.06	1.51	(65)
Spo0F (2FSP)	14,230	20.1	20.2	+0.25	1.24	(103)
DHFR (1DYI)	18,003	22.6	22.0	−2.87	1.20	(104)
β-lactoglobulin (2AKQ)	17,104	22.1	21.3	−3.48	1.21	(105)
Apo-AK (4AKE)	23,589	26.1	26.2	+0.31	1.43	(106)
Holo-AK (1AKE)	24,616	23.7	23.9	+0.93	1.14	(106)
		mean % signed error		−0.08		

DHFR, dihydrofolate reductase.
[a]Experimental hydrodynamic coefficients (τ_c, D_R) as reported in the references were converted to R_R for comparison.
[b]Axial ratio of a prolate ellipsoid with volume equal to the initial convex hull volume.

downloaded from the Protein Ensemble Database (41). This ensemble contains 575 structures and was generated as described in (42).

Benchmarking

Benchmarking was carried out to measure and compare the execution speeds for HullRad, HYDROPRO, and SOMO. HYDROPRO speed was measured on an Intel ×64 machine running Ubuntu Linux; SOMO speed in Ultrascan (15) version 3.3 Rev 2202 was measured on a MacIntosh 2.2 GHz Intel core i7 laptop running OS X 10.11.6; HullRad speed was measured on both platforms. The following PDB files were used for this comparison purpose: 5PTI, 1HRC, 1AKI, 1AVU, 4CHA, 1BEB, 1AO6, 1CTS, 1ADO, 1FIQ, and 4V40.

RESULTS AND DISCUSSION

Convex hull method predicts translational hydrodynamic radius of folded proteins

A convex hull model of molecular structures was implemented in a program entitled HullRad. HullRad predicts the translational hydrodynamic radius, R_T, of folded proteins over a large range of protein sizes (6518–3,456,960 Da) and shapes. Fig. 2 shows representative images of initial convex hulls overlaid upon ribbon representations for select proteins spanning a range of molecular mass and shape. Fig. S2 shows a depiction of the unified side chain model together with the convex hull construct for the same proteins.

Table 1 lists the experimental set used to validate folded protein translational hydrodynamic radii calculations. The set consists of data derived from 33 nonhomologous proteins and represents independent measurements of sedimentation or diffusion coefficients derived from three different experimental methods. As described in the Methods, the joint comparison of these data is easily conducted by using the equivalent $R_{T,exp}$ value for each observable.

Fig. S3 *A* shows the optimized hydration shell thickness for translational diffusion as 2.8 Å. Incorporation of this expansion on the initial hull results in the final volumes used in the hydrodynamic calculations. Fig. 3 *A* shows

TABLE 3 Hydrodynamic Radii for Nucleic Acids and Complexes

Molecule (PDB ID)	M (Da)	$R_{T,exp}$[a] (Å)	$R_{T,calc}$ (Å)	% Signed Error	$R_{R,exp}$[a] (Å)	$R_{R,calc}$ (Å)	a/b[b]	Reference(s)
DNA duplex 8mer[c]	4850	14.1	13.8	−1.9	14.5	14.8	1.13	(33,65)
DNA duplex 12mer (5NT5)	7331	16.6[d]	16.4	−1.2	18.3	18.0	1.61	(33,35,36)
DNA duplex 20mer (5F9I)	12,288	19.7	20.5	+4.1	25.0	24.7	2.62	(33)
DNA duplex 24mer[e]	14,758	22.5	22.6	+0.4	ND	ND	2.92	(34)
tRNA-HisRS (4RDX)	70,463	37.4	37.8	+1.1	ND	ND	1.60	(39)
Bacteria ribosome 50S (5HKV)	1,263,178[f]	92.9	95.32	+2.6	ND	ND	1.26	(40)
		mean % signed error		+0.85				

ND, not determined.
[a]Experimental hydrodynamic coefficients as reported in the references were converted to R_T or R_R for comparison.
[b]Axial ratio of prolate ellipsoid with maximum dimension and volume equal to the respective convex hull.
[c]The structure in PDB file 5NT5 was edited to be equivalent to an 8mer.
[d]Average of values reported in cited references.
[e]The structure in PDB file 5F9I was used to model a 24mer with the same pitch.
[f]Molar mass of the structure in 5HKV. A reported actual molar mass of 1,364,954 g mol^{-1} (107) was used with a sedimentation coefficient of 50.0 and partial specific volume of 0.614 to calculate $R_{T,exp}$.

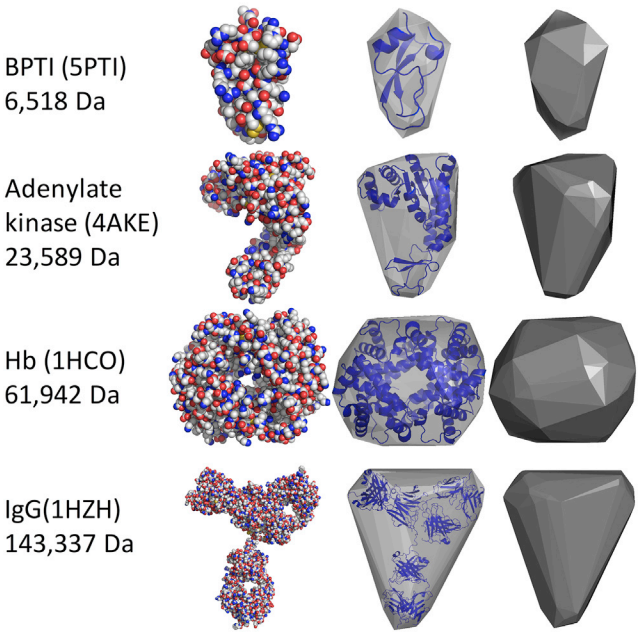


FIGURE 2 Convex hull representations of protein volumes. Four proteins of different sizes and shapes are depicted showing the relationship of a convex hull model to atomic space filling and backbone ribbon models. The left vertical column gives the PDB ID and molecular mass for each protein. The second vertical column shows an atomic sphere representation; the third column shows backbone cartoon diagram representations for each protein (*blue* online) enclosed by its semitransparent initial convex hull (*gray* online); the right vertical panel shows the convex hull model as a solid. This figure was created in PyMOL. The transparent and solid convex hull planes are CGOs created within PyMOL by using the convex hull vertices. A PyMOL script to generate these CGOs is included with the HullRad code. CGO, compiled graphic object. To see this figure in color, go online.

the excellent agreement between experimental and calculated values for this set. Fig. 3 *B* and Table 1 show the deviations of the calculated values from the experimental numbers. The percent signed error ($100\times(R_{T,calc}-R_{T,exp})/R_{T,exp}$) of HullRad as compared to experimental values ranges from −5.7 to +6.3 with a mean value of 0.1. The mean percent signed error for these independent jackknifed predictions is −0.04, with a range of −5.9 to +6.4. These minimum and maximum percent signed errors increased only slightly when the jackknife procedure was conducted, indicating that no one protein overly influences the fit. The original grouped errors and jackknifed independent errors are found to not be significantly different when using a *t*-test with a *t*-value of 0.19 and *p*-value of 0.43. Notably, all calculated values are well within the accuracy of experimentally determined hydrodynamic properties by analytical ultracentrifugation (43), one of the most accurate experimental methods available for determining hydrodynamic properties.

Fig. 3 *B* shows for comparison the percent signed deviations of calculated R_T from HYDROPRO. This program, too, has excellent accuracy except for the smallest proteins in the set where the HullRad method is slightly better when default values are used in the HYDROPRO program. The source of the discrepancy for small proteins can be explained by the default value of 2.9 Å for the primary element radius in the HYDROPRO program, which is optimized for a set of proteins with larger sizes than included here; more accurate calculated values could be obtained with a protein-size-optimized primary element radius. Rocco and Byron have previously shown that the SOMO implementation in Ultrascan as well as the standalone program ZENO reproduce hydrodynamic parameters within similar (∼5%) deviation from the experiment (32).

HullRad execution is fast and easy

Given comparable accuracy to other hydrodynamic prediction programs, two main attractions of HullRad are its ease of use and its speed of execution. Calculations involve running a simple Python script with a PDB file as input as described in the Methods. A custom user script can easily be written to run HullRad on a list of PDB files, which means that HullRad has the capability to analyze ensembles of structures generated, for example, in a molecular dynamics trajectory or a Monte Carlo simulation that seeks to explore conformational properties of members of an ensemble. Thousands of structures could easily be parsed in a short time as described below in the section on IDPs. In addition, HullRad handles large files. Although there is no experimental data for comparison, we validated the ability of HullRad to calculate hydrodynamic parameters of very large macromolecular assemblies using the HIV-1 capsid model, which is currently the largest atomic structure available in the Research Collaboratory for Structural Bioinformatics with 2,440,801 atoms (M = 34,795,461; PDB identifier [ID] = 3J3Q). The calculated R_T is 416 Å.

We further demonstrate the execution speed advantage of the HullRad algorithm by benchmarking its performance against both HYDROPRO and SOMO. Fig. 4 shows that HullRad output was generated on the millisecond time scale irrespective of protein size over the range tested (6518 to 464,490 Da). HullRad is orders of magnitude faster than either HYDROPRO or SOMO for all proteins, except HullRad is only twice as fast as HYDROPRO for the smallest test example (bovine pancreatic trypsin inhibitor [BPTI], M = 6518 Da). HullRad is also much faster than ZENO, for which calculation times have been reported to be on the order of minutes for a 100-residue protein (13).

Convex hull method predicts rotational hydrodynamic radii of folded proteins

Fig. 5 *A* shows that the convex hull method also accurately predicts isotropic rotational hydrodynamic radii

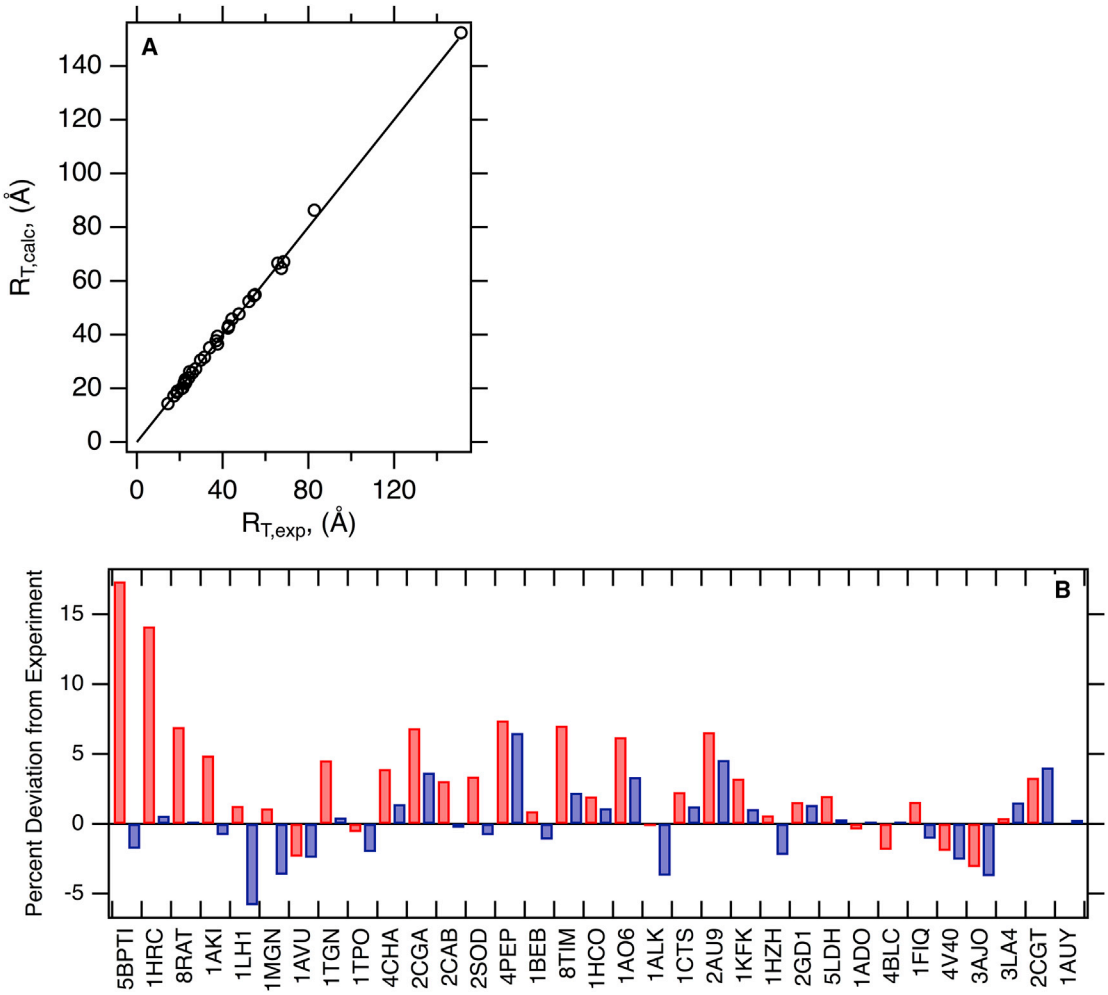


FIGURE 3 The convex hull method accurately predicts experimental protein translational hydrodynamic radii. (*A*) HullRad predicted $R_{T,calc}$ for folded proteins are plotted versus experimental $R_{T,exp}$ values. The black line represents a slope of one and intercept of zero; the correlation coefficient for a linear regression of the data (data not shown) is 0.998. (*B*) Percent signed errors of calculated to experimental R_T for folded proteins are plotted as bars. Dark gray (*blue* online) was calculated with HullRad; light gray (*red* online) was calculated with HYDROPRO. In this comparison, all HYDROPRO calculations used atomic-level beads except the structure 2CGT, for which residue level beads were used. Only HullRad was used to calculate hydrodynamic properties for the largest example, 12AUY. The individual proteins are labeled according to respective PDB IDs; common names, molecular masses, number of chains, and respective hydrodynamic radii are listed in Table 1. This figure was created in Igor. To see this figure in color, go online.

for a data set of folded protein structures. The experimental and calculated values are listed in Table 2, in which the mean percent signed error of calculated to experimental values is −0.08 with a range of −3.5 to +5.1. The individual protein prediction errors of calculated rotational radii to experimental values from the convex hull and bead modeling algorithms are graphically compared in Fig. 5 *B*. Both programs are accurate within ∼5% error.

Fig. S3 *B* shows the optimized hull expansion for calculation of rotational diffusion properties using a shape factor correction, $F_R = F_T^4$. The optimal hydration thickness shown in Fig. S3 *B* (4.3 Å) is larger than that required to accurately predict translational diffusion (2.8 Å) in agreement with the theoretical difference between translational and rotational hydration effect described in the theory section of Supporting Materials and Methods.

Convex hull method predicts hydrodynamic radii of nucleic acids and protein-nucleic acid complexes

Fig. 6 *A* shows four DNA duplex structures with their respective initial convex hull representations. HullRad predicts both the R_T and R_R of these DNA molecules with excellent accuracy as shown in Fig. 6 *B* and Table 3. Note that accurate prediction of the translational hydrodynamic radii for DNA is achieved with the same convex hull hydration expansion used for proteins as expected (44). Furthermore, HullRad predicts the translational hydrodynamic radii of the protein-RNA complexes, histidinyl-tRNA synthetase, and bacterial 50s

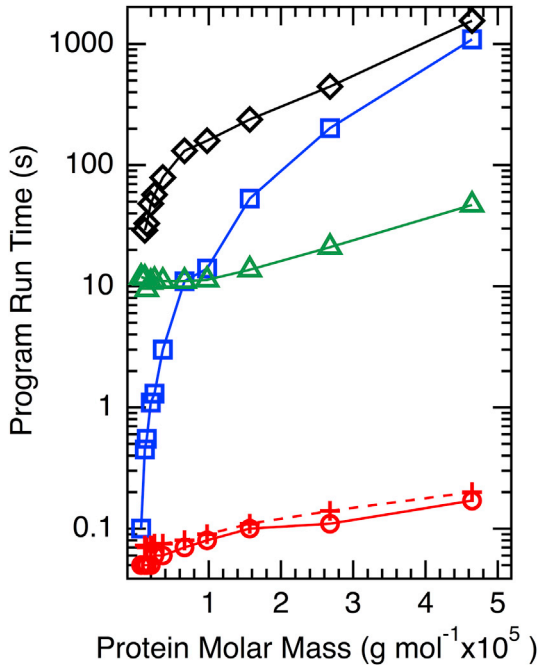


FIGURE 4 HullRad execution is orders of magnitude faster than other common hydrodynamic prediction algorithms for most proteins. The red circles and crosses are data for HullRad measured on an Intel ×64 machine running Ubuntu Linux and a MacOS laptop computer, respectively; the blue squares and green triangles are HYDROPRO runtimes on Linux using the default atom and residue options, respectively; black diamonds are SOMO runtimes on a Macintosh laptop. The points for each series are connected by lines for clarity. To see this figure in color, go online.

ribosomal subunit with excellent accuracy as shown in Table 3. HullRad calculates the hydrodynamic properties of the 50s ribosomal subunit structure with a molecular weight of 1.26 million in 0.57 s on a laptop computer running MacOS.

Convex hull method predicts translational hydrodynamic radii of IDPs

In addition to folded proteins with a defined three-dimensional structure, much of the proteome consists of polypeptides that do not form unique native structures (45). These proteins are referred to as IDPs. Proteins can have both stable, folded regions as well as intrinsically disordered regions. These intrinsically disordered regions are now recognized to be important for a number of functions: motif display sites, chaperones, scavengers, complex assembly, effector ligands, and entropic chains (45).

The sizes and shapes of IDPs are topics of intense interest (46,47). These proteins exist as conformationally heterogeneous ensembles in solution, and there are many efforts to measure both the average degree of structure and their dynamics. Approaches include experimentally measuring specific residue contacts, residue dipolar couplings, and/or hydrodynamic properties. Small angle x-ray scattering, analytical ultracentrifugation, and NMR are techniques frequently employed for this purpose. However, these experiments report ensemble averaged observables and rely on models for interpretation (46). The structural disorder inherent in IDPs creates a challenge to the generation of such structural models. In particular, large numbers of individual structural models may need to be created to accurately construct a distribution of conformations. Even though molecular simulations are able to generate large numbers of atomic coordinate models, a rapid method to calculate the desired physical parameters from this ensemble of structures is required (47).

The HullRad algorithm has the speed, accuracy, and ease of use to be well suited for this task. We demonstrate this using a conformationally heterogeneous ensemble of α-synuclein. This protein is a 140-residue (MW 14,462), natively disordered protein and is present in the protofibrils found in Parkinson's disease (48). The experimentally determined R_T for α-synuclein in 100 mM NaCl is 32.0 Å (49), indicating a slightly more compact ensemble than a protein of the same size denatured in urea (33.0 Å), but more extended than a premolten globule state (26.4 Å) (50). Allison et al. (42) used paramagnetic relaxation enhancement NMR spectroscopy to obtain 595 intermolecular distances for the natively unfolded α-synuclein. These inter-residue distances were used as ensemble-averaged restraints in multiple

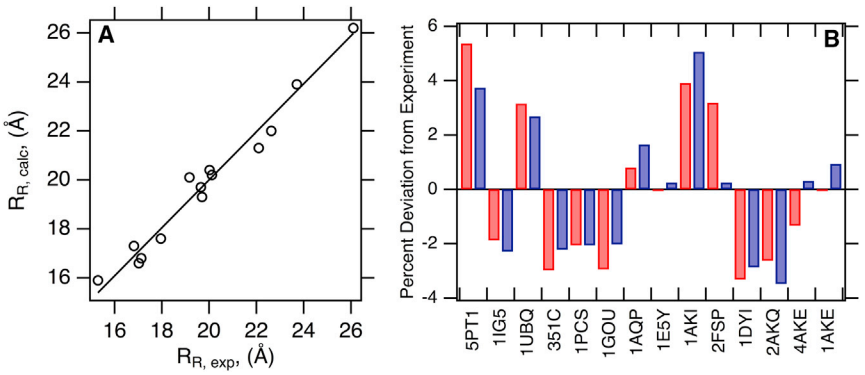


FIGURE 5 The convex hull method accurately predicts experimental rotational hydrodynamic radii. (*A*) HullRad-predicted R_R for the proteins listed in Table 2 are plotted versus experimental R_R values (*circles*). The line represents a slope of one and intercept of zero; the correlation coefficient for a linear regression of the data (data not shown) is 0.993. (*B*) Percent signed errors of calculated to experimental R_R for the list of proteins in Table 2 are plotted as bars. Dark gray (*blue* online) was calculated with the program HullRad; light gray (*red* online) was calculated with HYDROPRO. The individual proteins are labeled according to respective PDB IDs; common names, molecular masses, and respective hydrodynamic radii are listed in Table 2. To see this figure in color, go online.

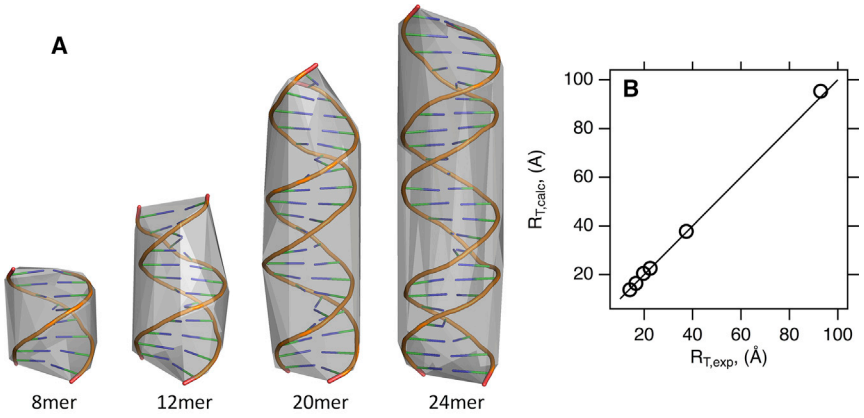


FIGURE 6 The convex hull method accurately predicts experimental hydrodynamic radii for nucleic acids and protein-nucleic acid complexes. (*A*) Shown are convex hull representations of DNA duplexes used for validation of HullRad. Each image shows the PyMOL backbone cartoon representation enclosed by its semitransparent initial convex hull. The PDB IDs, respective molecular masses, hydrodynamic radii, and ellipsoidal axial ratios are listed in Table 3. (*B*) A comparison of HullRad calculated and experimental hydrodynamic radii for nucleic acids and protein-nucleic acid complexes listed in Table 3. The line represents a slope of one and intercept of zero; the correlation coefficient for a linear regression of the data (data not shown) equals 0.999. To see this figure in color, go online.

replicate molecular dynamics simulations to generate a disordered state ensemble of 575 structures.

Fig. 7 shows a comparison of calculated effective R_T values for this ensemble using both HullRad and HYDROPRO. The results were sorted by compactness by using anhydrous R_G calculated from the protein atomic coordinates only. The calculated ensemble mean HYDROPRO hydrodynamic radius is ~0.7% larger than the HullRad hydrodynamic radius (mean $R_{T,calc}$ = 31.13 vs. 30.90 Å, respectively). A larger individual difference is apparent for the more compact structures. This is expected from the differences between these programs for small proteins shown in Fig. 3 *B*. The individual $R_{T,calc}$ values from both methods are highly correlated (see Fig. S4). HullRad processes these files much more quickly: computer runtimes on a typical ×86 chip for analysis of the 575 structures of α-synuclein are 24 s for HullRad and 27 min for HYDROPRO.

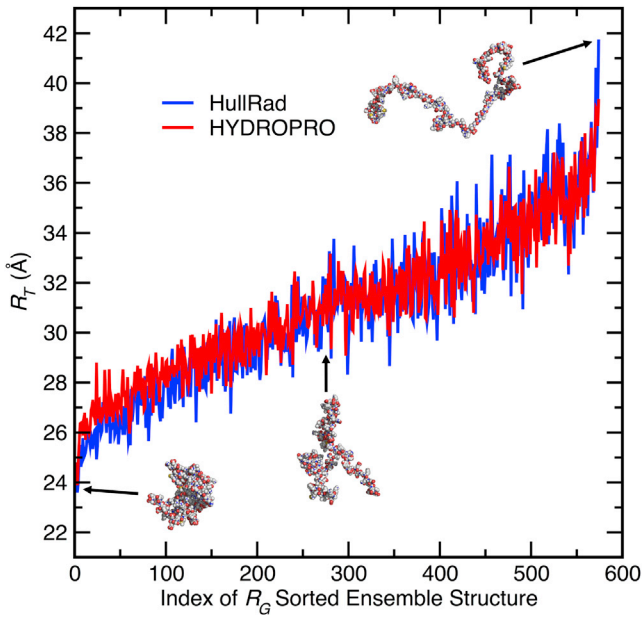


FIGURE 7 Prediction of hydrodynamic properties for a disordered protein ensemble. Shown are the effective translational hydrodynamic radii of disordered protein structures calculated using HullRad (*dark gray*, *blue* online) or HYDROPRO (*light gray*, *red* online). The structures are sorted by compactness using the anhydrous R_G calculated using Eq. 13. Three example structures representing different degrees of compactness are shown as atomic sphere models. The ensemble contains 575 generated structures of α-synuclein from the Protein Ensemble Database (accession number PED9AAC). To see this figure in color, go online.

The agreement shown in Fig. 7 and Fig. S4 indicate that the convex hull method works well even for the unusual, extended conformations found in such a disordered state ensemble. HullRad and HYDROPRO are similarly accurate for disordered structures.

The mean HullRad calculated $R_{T,calc}$ of 30.9 Å for the α-synuclein ensemble is smaller than the experimentally determined $R_{T,exp}$ of 32.0 Å. This difference suggests that modifying the simulation-generated α-synuclein ensemble to contain a higher proportion of extended structures would more closely mimic the experimental ensemble. Further exploration and culling of this ensemble will be easy to accomplish given the ease of use and execution speed inherent in HullRad.

Shape factors and ellipsoidal representations

Fig. 8 shows both the initial convex hull and ellipsoid of revolution models for the most elongated protein in our data sets (tryptophan synthetase) and the longest DNA duplex tested (24 basepair). The respective ellipsoids of revolution were constructed with the same maximum length and volume of the initial convex hull for calculation of a shape factor.

The translational diffusion shape factor used by HullRad is based on the Perrin equation for a prolate ellipsoid (24). The Perrin-derived shape factors for prolate and oblate ellipsoids are different by only 5.5% up to axial ratios of 10:1 (1). This axial ratio represents far more nonsphericity than encountered for most biomolecules studied by diffusion measurements. The maximum axial ratio observed in the above ensemble of intrinsically disordered α-synuclein is 4.62. This fact, together with the significant increase in

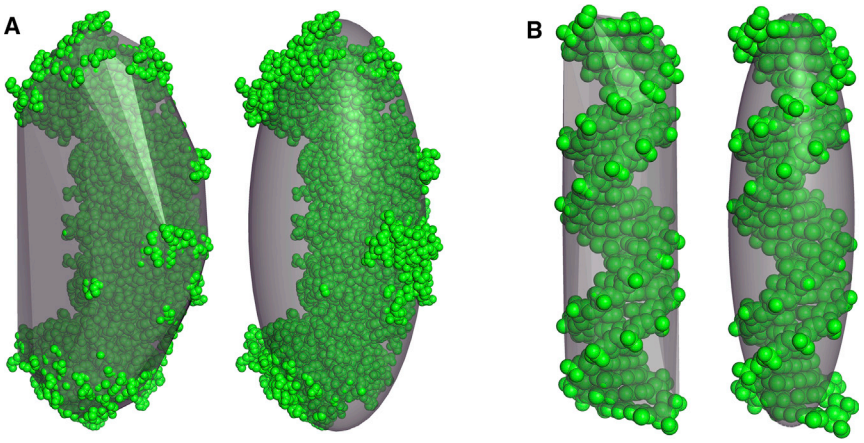


FIGURE 8 Comparison of convex hull and ellipsoid of revolution models. (*A*) Tryptophan synthetase as atomic spheres (*green* online). (*B*) Shown is a DNA duplex with 24 base pairs as atomic spheres. Each molecular structure is overlaid with a semitransparent (*gray* online) initial convex hull (*left*) or prolate ellipsoid of revolution (*right*) of volumes and maximum lengths equal to that of the respective convex hull. To see this figure in color, go online.

computing time necessary to calculate the best-fit general tri-axial ellipsoid of a molecule, led us to use the simple construction of a prolate ellipsoid described above to estimate the effect of elongation on translational diffusion. As shown in Tables 1 and 3, this approximation allowed accurate calculation of translational diffusion properties for all the molecules tested.

Calculation of the rotational shape factors for elongated molecules is more complicated than the corresponding translational factors because 1) rotations around different axes of the molecule can proceed at different rates, and rotation is best described by a rotational tensor (25); and 2) the rotational tensors are significantly different for prolate and oblate ellipsoidal shapes (51). HullRad is limited to the calculation only of isotropic tumbling. We found by empirical comparisons that a rotational shape factor correction, $F_R = F_T^4$, allowed accurate prediction of isotropic rotational diffusion radii for the most elongated, rigid structures for which experimental data is available (Fig. S5 and Table S1). Although an anisotropic diffusion model is required for accurate analysis of NMR rotational relaxation of highly nonspherical molecules, an isotropic model is sufficient for many biological molecules given the precision of relaxation data (12). The HullRad code prints a warning to users when a molecule has an axial ratio larger than 2.62, which is the largest in the experimental data available for comparison in this study.

Internal cavity and external crevice water are hydrodynamically equivalent

The convex hull transport model also illustrates the hydrodynamic equivalence of internally buried cavity water and surface crevice water as shown in Fig. 9. Fig. 9 *A* (*top row*) shows apoferritin, which is a hollow protein sphere containing a large water-filled cavity as illustrated by the cross-sectional view in the right panel. The hydrodynamic $R_{T,exp}$ of apoferritin is 64.71 Å, and the anhydrous protein volume R_0 is 51.83 Å. Because this molecule is nearly spherical, the difference in these two parameters is due to a combination of hydration and hydrodynamic water. Hydration water is inherently part of the diffusing molecule and is frequently modeled by a molecular surface expansion of 1.1–2.8 Å (8,12). Here, hydrodynamic water is that water not considered true hydration water but rather water that is entrained by the indentations in the molecular surface or trapped in internal cavities. Clearly, the water trapped in the cavity of apoferritin is hydrodynamic water in the sense that it diffuses with the protein even if, dynamically, it is able to exchange with bulk water.

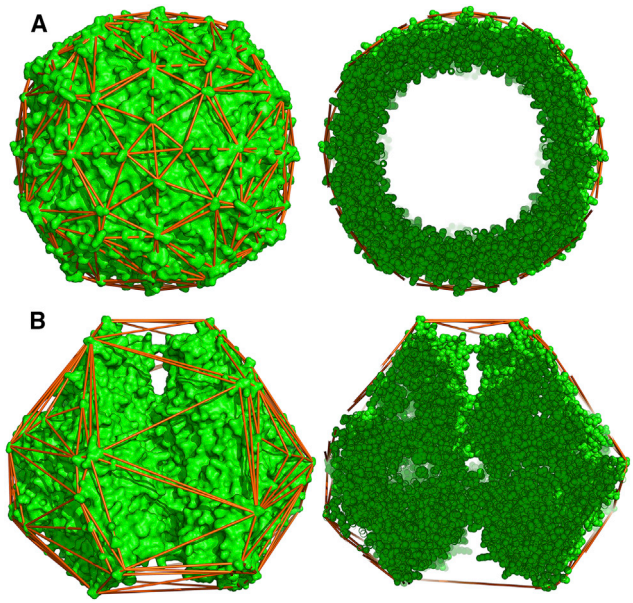


FIGURE 9 The hydrodynamic hydration water in internal cavities and surface crevices is equivalent. (*A*) shows apoferritin (PDB ID: 3AJO); (*B*) shows urease (PDB ID: 3LA4). In the left vertical panel, molecular surface renderings (*green* online) with convex hull edges as sticks (*orange* online) are shown. In the right vertical panel, 30 Å slabs of the same structures are shown as atomic spheres instead of molecular surfaces. This figure was created in PyMOL. To see this figure in color, go online.

Fig. 9 *B* (*bottom row*) shows that the structural arrangements and hydration patterns of urease are distinct from apoferritin, yet the implication for hydrodynamic volume is the same. Urease is a comparably sized protein with a similar hydrodynamic radius ($R_{T,exp}$ = 66.58 Å) and anhydrous protein volume radius (R_0 = 54.09 Å), yet it has no large internal cavity. In contrast, most of the urease hydrodynamic water is in medium-sized surface crevices encapsulated by the convex hull as shown in the right panel of Fig. 9 *B*. The observation that the convex hull model accurately predicts the translational diffusion of urease is evidence that the water distributed in surface crevices on this structure also affects the diffusional volume of the protein in a manner similar to the internal cavity water of apoferritin. We postulate that this protein-specific inclusion of different amounts of water in the convex hull transport model may contribute to its accuracy. A further discussion of hydration and hydrodynamic water is included in the Supporting Materials and Methods.

We note that the convex hull transport model recapitulates hydrodynamic properties based only on geometry of the protein. It does not take into account any of the chemical details of the surface atoms that are also known to determine the dynamics of hydration shell water (52). As with most other predictive methods, the convex hull method smooths and averages the effects of hydration volume. We recognize that methods to predict location-specific waters of hydration do exist (53–57), but these are too time consuming for the purpose of the rapid prediction discussed here.

CONCLUSIONS

The convex hull volume of a macromolecule accurately predicts the hydrodynamic volume of that molecule. A Python script called HullRad is provided, which calculates sedimentation coefficients, translational and rotational diffusion coefficients, and hydrodynamic radii for protein and nucleic acid molecules using a convex hull algorithm. The method is orders of magnitude faster than common methods and works for molecules with significant shape asymmetry and surface rugosity, including the extended structures of a disordered state ensemble. The same fitted parameters are used for molecules that extend over two orders of magnitude in size. Application of this method to the modification of computer-generated disordered state ensembles should provide a better agreement of the calculated and experimental ensembles of IDPs.

SUPPORTING MATERIAL

Supporting Materials and Methods, five figures, and one table are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)30065-1.

AUTHOR CONTRIBUTIONS

P.J.F. designed and performed the research. P.J.F. and K.G.F. wrote the paper.

ACKNOWLEDGMENTS

The authors thank Margaret Johnson, Ananya Majumdar, Walter F. Stafford III, James Wrabl, and the members of the K.G.F. lab for helpful discussions and Olwyn Byron and Mattia Rocco for providing their data set of protein structures.

Funded by the National Institutes of Health (grant No. R01GM079440) and National Science Foundation (grant No. MCB1412108 to K.G.F.).

SUPPORTING CITATIONS

References (58–64) appear in the Supporting Material.

REFERENCES

1. Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry: Part II: Techniques for the Study of Biological Structure and Function.. W. H. Freeman, New York, pp. 561–565.
2. Cole, J. L., J. W. Lary, …, T. M. Laue. 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol.* 84:143–179.
3. Pilz, I. 1982. Proteins. *In* Small Angle X-ray Scattering. O. Glatter and O. Kratky, eds. Academic Press, New York, pp. 239–293.
4. Serdyuk, I. N., N. R. Zaccai, and J. Zaccai. 2007. Methods in Molecular Biophysics.. Cambridge University Press, Cambridge, UK, pp. 481–504.
5. Badea, M. G., and L. Brand. 1979. Time-resolved fluorescence measurements. *Methods Enzymol.* 61:378–425.
6. Himmel, M. E., and P. G. Squire. 1988. Size exclusion parameters. *In* Aqueous Size Exclusion Chromatography. P. L. Dubin, ed. Elsevier, New York, pp. 3–22.
7. Tjandra, N., H. Kuboniwa, …, A. Bax. 1995. Rotational dynamics of calcium-free calmodulin studied by 15N-NMR relaxation measurements. *Eur. J. Biochem.* 230:1014–1024.
8. Aragon, S. 2004. A precise boundary element method for macromolecular transport properties. *J. Comput. Chem.* 25:1191–1205.
9. García De La Torre, J., M. L. Huertas, and B. Carrasco. 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78:719–730.
10. Rocco, M., and O. Byron. 2015. Hydrodynamic modeling and its application in AUC. *Methods Enzymol.* 562:81–108.
11. Harding, S. E., and A. J. Rowe. 1983. Modeling biological macromolecules in solution. II. The general tri-axial ellipsoid. *Biopolymers.* 22:1813–1829.
12. Ryabov, Y. E., C. Geraghty, …, D. Fushman. 2006. An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. *J. Am. Chem. Soc.* 128:15432–15444.
13. Kang, E. H., M. L. Mansfield, and J. F. Douglas. 2004. Numerical path integration technique for the calculation of transport properties of proteins. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:031918.
14. Ortega, A., D. Amorós, and J. García de la Torre. 2011. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101:892–898.
15. Demeler, B. 2005. Hydrodynamic methods. *In* Bioinformatics Basics: Applications in Biological Science and Medicine. H. Rashidi and L. Buehler, eds. CRC Press, Boca Raton, Florida, pp. 226–255.

16. Aurenhammer, F. 1991. Voronoi diagrams - a survey of a fundamental geometric data structure. *ACM Comput. Surv.* 23:345–405.
17. Babick, F., J. Mielke, …, V. D. Hodoroaba. 2016. How reliably can a material be classified as a nanomaterial? Available particle-sizing techniques at work. *J. Nanopart. Res.* 18:158.
18. Babick, F. 2016. Suspensions of Colloidal Particles and Aggregates.. Springer International Publishing, Switzerland, p. 171.
19. Barber, C. B., D. P. Dobkin, and H. Hannu. 1996. The Quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 22:469–483.
20. Cohn, E. J., and J. T. Edsall. 1943. Proteins, Amino Acids and Peptides as Ions and Dipolar Ions.. Reinhold Publ. Corp., New York, pp. 155–176.
21. Perkins, S. J. 1986. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* 157:169–180.
22. Nadassy, K., I. Tomás-Oliveira, …, S. J. Wodak. 2001. Standard atomic volumes in double-stranded DNA and packing in protein–DNA interfaces. *Nucleic Acids Res.* 29:3362–3376.
23. Voss, N. R., and M. Gerstein. 2005. Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly. *J. Mol. Biol.* 346:477–492.
24. Perrin, F. 1936. Mouvement Brownian d'un ellipsoide II. Rotation libre et dépolarisation des fluorescences. Translation et diffusion de molécules ellipsoïdales. *J. Phys. Radium.* 7:1–11.
25. Favro, L. D. 1960. Theory of the rotational brownian motion of a free rigid body. *Phys. Rev.* 119:53–62.
26. Delano, W. L. 2015. The PyMOL Molecular Graphics System, version 1.8. Schrödinger, LLC.
27. Efron, B., and C. Stein. 1981. The jackknife estimate of variance. *Ann. Stat.* 9:586–596.
28. Lavalette, D., C. Tétreau, …, Y. Blouquit. 1999. Microscopic viscosity and rotational diffusion of proteins in a macromolecular environment. *Biophys. J.* 76:2744–2751.
29. The International Association for the Properties of Water and Steam. 2008. Release on the IAPWS Formulation 2008 for the Viscosity of Ordinary Water Substance.
30. Perkins, S. J. 2001. X-ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits. *Biophys. Chem.* 93:129–139.
31. Knight, C. J., and J. S. Hub. 2015. WAXSiS: a web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics. *Nucleic Acids Res.* 43 (W1):W225–W230.
32. Rocco, M., and O. Byron. 2015. Computing translational diffusion and sedimentation coefficients: an evaluation of experimental data and programs. *Eur. Biophys. J.* 44:417–431.
33. Eimer, W., and R. Pecora. 1991. Rotational and translational diffusion of short rodlike molecules in solution: Oligonucleotides. *J. Chem. Phys.* 94:2324–2329.
34. Lapham, J., J. P. Rife, …, D. M. Crothers. 1997. Measurement of diffusion constants for nucleic acids by NMR. *J. Biomol. NMR.* 10:255–262.
35. Birchall, A. J., and A. N. Lane. 1990. Anisotropic rotation in nucleic acid fragments: significance for determination of structures from NMR data. *Eur. Biophys. J.* 19:73–78.
36. Bonifacio, G. F., T. Brown, …, A. N. Lane. 1997. Comparison of the electrophoretic and hydrodynamic properties of DNA and RNA oligonucleotide duplexes. *Biophys. J.* 73:1532–1538.
37. Porschke, D., and J. Antosiewicz. 1990. Permanent dipole moment of tRNA's and variation of their structure in solution. *Biophys. J.* 58:403–411.
38. Benítez, A. A., J. G. Hernández Cifre, …, J. G. de la Torre. 2015. Prediction of solution properties and dynamics of RNAs by means of Brownian dynamics simulation of coarse-grained models: ribosomal 5S RNA and phenylalanine transfer RNA. *BMC Biophys.* 8:11.
39. Bovee, M. L., W. Yan, …, C. S. Francklyn. 1999. tRNA discrimination at the binding step by a class II aminoacyl-tRNA synthetase. *Biochemistry.* 38:13725–13735.
40. Hill, W. E., J. W. Anderegg, and K. E. Van Holde. 1970. Effects of solvent environment and mode of preparation on the physical properties of ribosomes fron Escherichia coli. *J. Mol. Biol.* 53:107–121.
41. Varadi, M., S. Kosol, …, P. Tompa. 2014. pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. *Nucleic Acids Res.* 42:D326–335.
42. Allison, J. R., P. Varnai, …, M. Vendruscolo. 2009. Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc.* 131:18314–18326.
43. Zhao, H., R. Ghirlando, …, P. Schuck. 2015. A multilaboratory comparison of calibration accuracy and the performance of external references in analytical ultracentrifugation. *PLoS One.* 10:e0126420.
44. Fernandes, M. X., A. Ortega, …, J. García de la Torre. 2002. Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. *Nucleic Acids Res.* 30:1782–1788.
45. van der Lee, R., M. Buljan, …, M. M. Babu. 2014. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* 114:6589–6631.
46. Best, R. B. 2017. Computational and theoretical advances in studies of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* 42:147–154.
47. Nygaard, M., B. B. Kragelund, …, K. Lindorff-Larsen. 2017. An efficient method for estimating the hydrodynamic radius of disordered protein conformations. *Biophys. J.* 113:550–557.
48. Stefanis, L. 2012. α-Synuclein in Parkinson's disease. *Cold Spring Harb. Perspect. Med.* 2:a009399.
49. Binolfi, A., R. M. Rasia, …, C. O. Fernández. 2006. Interaction of alpha-synuclein with divalent metal ions reveals key differences: a link between structure, binding specificity and fibrillation enhancement. *J. Am. Chem. Soc.* 128:9893–9901.
50. Uversky, V. N. 2002. What does it mean to be natively unfolded? *Eur. J. Biochem.* 269:2–12.
51. Koenig, S. H. 1975. Brownian-motion of an ellipsoid - A correction to perrins results. *Biopolymers.* 14:2421–2423.
52. Qin, Y., L. Wang, and D. Zhong. 2016. Dynamics and mechanism of ultrafast water-protein interactions. *Proc. Natl. Acad. Sci. USA.* 113:8424–8429.
53. Venable, R. M., E. Hatcher, …, R. W. Pastor. 2010. Comparing simulated and experimental translation and rotation constants: range of validity for viscosity scaling. *J. Phys. Chem. B.* 114:12501–12507.
54. Makarov, V. A., B. K. Andrews, …, B. M. Pettitt. 2000. Residence times of water molecules in the hydration sites of myoglobin. *Biophys. J.* 79:2966–2974.
55. Virtanen, J. J., L. Makowski, …, K. F. Freed. 2010. Modeling the hydration layer around proteins: HyPred. *Biophys. J.* 99:1611–1619.
56. Neumayr, G., T. Rudas, and O. Steinhauser. 2010. Global and local Voronoi analysis of solvation shells of proteins. *J. Chem. Phys.* 133:084108.
57. Fogarty, A. C., and D. Laage. 2014. Water dynamics in protein hydration shells: the molecular origins of the dynamical perturbation. *J. Phys. Chem. B.* 118:7715–7729.
58. Kuntz, I. D., Jr., and W. Kauzmann. 1974. Hydration of proteins and polypeptides. *Adv. Protein Chem.* 28:239–345.
59. Aragon, S., and D. K. Hahn. 2006. Precise boundary element computation of protein transport properties: diffusion tensors, specific volume, and hydration. *Biophys. J.* 91:1591–1603.
60. Denisov, V. P., and B. Halle. 1996. Protein hydration dynamics in aqueous solution. *Faraday Discuss.* 103:227–244.

61. Halle, B., and M. Davidovic. 2003. Biomolecular hydration: from water dynamics to hydrodynamics. *Proc. Natl. Acad. Sci. USA.* 100:12135–12140.

62. Tanford, C. 1961. Physical Chemistry of Macromolecules.. Wiley, New York, p. 337.

63. Bird, R. B., W. E. Stewart, and E. N. Lightfoot. 2002. Transport Phenomena.. J. Wiley, New York.

64. Edwardes, D. 1893. Steady motion of a viscous liquid in which an ellipsoid is constrained to rotate about a principal axis. *Quarterly Journal of Pure and Applied Mathematics.* 26:70–78.

65. Rai, N., M. Nöllmann, …, M. Rocco. 2005. SOMO (SOlution MOdeler) differences between X-Ray- and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. *Structure.* 13:723–734.

66. Wilkins, D. K., S. B. Grimshaw, …, L. J. Smith. 1999. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry.* 38:16424–16431.

67. Walters, R. R., J. F. Graham, …, D. J. Anderson. 1984. Protein diffusion coefficient measurements by laminar flow analysis: method and applications. *Anal. Biochem.* 140:190–195.

68. Clark, S. M., D. G. Leaist, and L. Konermann. 2002. Taylor dispersion monitored by electrospray mass spectrometry: a novel approach for studying diffusion in solution. *Rapid Commun. Mass Spectrom.* 16:1454–1462.

69. Broughton, W. J., M. J. Dilworth, and C. A. Godfrey. 1972. Molecular properties of lupin and serradella leghaemoglobins. *Biochem. J.* 127:309–314.

70. Gast, K., H. Damaschun, …, G. Damaschun. 1994. Compactness of protein molten globules: temperature-induced structural changes of the apomyoglobin folding intermediate. *Eur. Biophys. J.* 23:297–305.

71. Hanlon, A. D., M. I. Larkin, and R. M. Reddick. 2010. Free-solution, label-free protein-protein interactions characterized by dynamic light scattering. *Biophys. J.* 98:297–304.

72. Tietze, F. 1953. Molecular-kinetic properties of crystalline trypsinogen. *J. Biol. Chem.* 204:1–11.

73. Cunningham, L. W., Jr. 1954. Molecular-kinetic properties of crystalline diisopropyl phosphoryl trypsin. *J. Biol. Chem.* 211:13–19.

74. Wilcox, P. E., J. Kraut, …, H. Neurath. 1957. The molecular weight of alpha-chymotrypsinogen. *Biochim. Biophys. Acta.* 24:72–78.

75. Armstrong, J. M., D. V. Myers, …, J. T. Edsall. 1966. Purification and properties of human erythrocyte carbonic anhydrases. *J. Biol. Chem.* 241:5137–5149.

76. Wood, E., D. Dalgleish, and W. Bannister. 1971. Bovine erythrocyte cupro-zinc protein. 2. Physicochemical properties and circular dichroism. *Eur. J. Biochem.* 18:187–193.

77. Edelhoch, H. 1957. The denaturation of pepsin. I. macromolecular changes. *J. Am. Chem. Soc.* 79:6100–6109.

78. Cecil, R., and A. G. Ogston. 1949. The sedimentation constant, diffusion constant and molecular weight of lactoglobulin. *Biochem. J.* 44:33–35.

79. Kawahara, K., A. G. Kirshner, and C. Tanford. 1965. DISSOCIATION OF HUMAN CO-hemoglobin by urea, guanidine hydrochloride, and other reagents. *Biochemistry.* 4:1203–1213.

80. Charlwood, P. A. 1952. Sedimentation and diffusion of human albumins. I. Normal human albumins at a low concentration. *Biochem. J.* 51:113–118.

81. Wong, S. C., D. C. Hall, and J. Josse. 1970. Constitutive inorganic pyrophosphatase of Escherichia coli. 3. Molecular weight and physical properties of the enzyme and its subunits. *J. Biol. Chem.* 245:4335–4345.

82. Wu, J. Y., and J. T. Yang. 1970. Physicochemical characterization of citrate synthase and its subunits. *J. Biol. Chem.* 245:212–218.

83. Lane, A. N., and K. Kirschner. 1983. The quaternary structure of tryptophan synthase from Escherichia coli. Fluorescence and hydrodynamic studies. *Eur. J. Biochem.* 129:675–684.

84. Jøssang, T., J. Feder, and E. Rosenqvist. 1988. Photon correlation spectroscopy of human IgG. *J. Protein Chem.* 7:165–171.

85. Jaenicke, R., and W. B. Gratzer. 1969. Cooperative and non-cooperative conformational effects of the coenzyme on yeast glyceraldehyde-3-phosphate dehydrogenase. *Eur. J. Biochem.* 10:158–164.

86. Jaenicke, R., E. Gregori, and M. Laepple. 1979. Conformational effects of coenzyme binding to porcine lactic dehydrogenase. *Biophys. Struct. Mech.* 6:57–65.

87. Christen, P., F. Goschke, …, A. Schmid. 1965. Uber die aldolase der Kaninchenleber molekulargewicht, dissoziation in untereinheiten. *Helv. Chim. Acta.* 48:1050–1056.

88. Kawahara, K. 1969. Evaluation of diffusion coefficients of proteins from sedimentation boundary curves. *Biochemistry.* 8:2551–2557.

89. Taylor, J. F., A. A. Green, and G. T. Cori. 1948. Crystalline aldolase. *J. Biol. Chem.* 173:591–604.

90. Samejima, T. 1959. Splitting of catalase molecule by alkali treatment. *J. Biochem.* 46:155–159.

91. Andrews, P., R. C. Bray, …, K. V. Shooter. 1964. The chemistry of xanthine oxidase. 11. Ultracentrifuge and gel-filtration studies on the milk enzyme. *Biochem. J.* 93:627–632.

92. Sund, H., and K. Weber. 1963. Studies on the lactose-splitting enzyme. XIII. Quantity and configuration of beta-galactosidase from E. Coli. *Biochem. Z.* 337:24–34.

93. May, C. A., J. K. Grady, …, N. D. Chasteen. 2010. The sedimentation properties of ferritins. New insights and analysis of methods of nanoparticle preparation. *Biochim. Biophys. Acta.* 1800:858–870.

94. Sumner, J. B., N. Gralen, and I. B. Eriksson-Quensel. 1938. The molecular weight of urease. *J. Biol. Chem.* 125:37–44.

95. Follmer, C., F. V. Pereira, …, C. R. Carlini. 2004. Jack bean urease (EC 3.5.1.5) aggregation monitored by dynamic and static light scattering. *Biophys. Chem.* 111:79–87.

96. Behlke, J., O. Ristau, and H. J. Schönfeld. 1997. Nucleotide-dependent complex formation between the Escherichia coli chaperonins GroEL and GroES studied under equilibrium conditions. *Biochemistry.* 36:5149–5156.

97. Harding, S. E., and P. Johnson. 1985. Physicochemical studies on turnip-yellow-mosaic virus. Homogeneity, relative molecular masses, hydrodynamic radii and concentration-dependence of parameters in non-dissociating solvents. *Biochem. J.* 231:549–555.

98. Kördel, J., N. J. Skelton, …, W. J. Chazin. 1992. Backbone dynamics of calcium-loaded calbindin D9k studied by two-dimensional proton-detected 15N NMR spectroscopy. *Biochemistry.* 31:4856–4866.

99. Russell, B. S., L. Zhong, …, K. L. Bren. 2003. Backbone dynamics and hydrogen exchange of Pseudomonas aeruginosa ferricytochrome c(551). *J. Biol. Inorg. Chem.* 8:156–166.

100. Bertini, I., D. A. Bryant, …, J. Zhao. 2001. Backbone dynamics of plastocyanin in both oxidation states. Solution structure of the reduced form and comparison with the oxidized state. *J. Biol. Chem.* 276:47217–47226.

101. Pang, Y., M. Buck, and E. R. Zuiderweg. 2002. Backbone dynamics of the ribonuclease binase active site area using multinuclear ((15)N and (13)CO) NMR relaxation and computational molecular dynamics. *Biochemistry.* 41:2655–2666.

102. Kalverda, A. P., M. Ubbink, …, G. W. Canters. 1999. Backbone dynamics of azurin in solution: slow conformational change associated with deprotonation of histidine 35. *Biochemistry.* 38:12690–12697.

103. Feher, V. A., J. W. Zapf, …, J. Cavanagh. 1997. High-resolution NMR structure and backbone dynamics of the Bacillus subtilis response regulator, Spo0F: implications for phosphorylation and molecular recognition. *Biochemistry.* 36:10015–10025.

104. Mauldin, R. V., P. J. Sapienza, …, A. L. Lee. 2012. Structure and dynamics of the G121V dihydrofolate reductase mutant: lessons from a transition-state inhibitor complex. *PLoS One.* 7:e33252.

105. Uhrínová, S., M. H. Smith, …, P. N. Barlow. 2000. Structural changes accompanying pH-induced dissociation of the beta-lactoglobulin dimer. *Biochemistry.* 39:3565–3574.

106. Shapiro, Y. E., and E. Meirovitch. 2009. Evidence for domain motion in proteins affecting global diffusion properties: a nuclear magnetic resonance study. *J. Phys. Chem. B.* 113:7003–7011.

107. Matzov, D., Z. Eyal, …, A. Yonath. 2017. Structural insights of lincosamides targeting the ribosome of Staphylococcus aureus. *Nucleic Acids Res.* 45:10284–10292.

Biophysical Journal, Volume *114*

Supplemental Information

HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties

Patrick J. Fleming and Karen G. Fleming

Supporting Information

HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties

Patrick J. Fleming and Karen G. Fleming

(2018) *Biophysical Journal*
https://doi.org/10.1016/j.bpj.2018.01.002

T. C. Jenkins Department of Biophysics, John Hopkins University, Baltimore, MD 21218, USA

1 Theory

1.1 Stokes-Einstein equations

The anhydrous diffusion coefficients of a molecule in solution are related to the equivalent size, anhydrous sphere by the Stokes-Einstein equations,

$$D_T = \frac{k_B T}{6\pi\eta_0 R_0} \tag{S1a}$$

$$D_R = \frac{k_B T}{8\pi\eta_0 R_0^3} \tag{S1b}$$

where D_T is the translational diffusion coefficient, D_R is the rotational diffusion coefficient, k_B is the Boltzmann constant, T is temperature, η_0 is the solvent viscosity, and R_0 is the radius of the equivalent size, anhydrous sphere.

For a protein or nucleic acid in solution the operative radius is not the radius, R_0, of a sphere equivalent to the anhydrous volume of the molecule, but rather the equivalent radius, R_H, of the hydrated molecule that includes any waters of hydration that transiently interact with the molecule and affect its diffusion, times a shape factor, F_S, to account for non-sphericity. As discussed below, both the hydrated radii and the shape factors are different for translational and rotational diffusion.

1.2 Hydration water considerations

The amount of hydration water, and therefore the expansion of R_H over R_0, has been a controversial topic. Kuntz and Kauzmann originally reviewed this topic and concluded that proteins may have between 0.3 and 0.6 grams of hydration water per gram of protein (1). Importantly, different proteins appear to have different apparent fractions of hydration water. This fact means that assumption of a uniform hydration fraction in the calculation of hydrodynamic properties would lead to variable errors depending on the specific protein. To account for hydration when calculating hydrodynamic coefficients from structure, various approaches have been employed, and most use an empirically derived best-fit uniform value for the hydration. In the boundary element method, the triangulated protein surface is expanded by some thickness optimized for all proteins in a data set (2). In the bead-modeling method the size of the beads is expanded to some radius also optimized for all proteins in the data set (3, 4) or molecular dynamics simulations may be used to identify specific hydration waters that are then used in the calculation (5). In an ellipsoid model a hydration layer equal to the diameter of a water molecule (2.8 Å) is added to the surface of the ellipsoid (6). And for the numerical path integration method individual residues are modeled as expanded spheres to account for hydration (7).

Such adjustments for hydration water imply that this water is rigidly bound to the surface of the protein and increases its apparent size. However, the diffusion rate of surface water has been shown by both computational (8) and experimental methods (9) to be orders of magnitude greater than the protein. Therefore, significant hydration water is

not rigidly bound to the surface of the protein. This apparent paradox may be reconciled by assuming that the hydration waters transiently interacting with the protein surface experience a modest viscosity enhancement (10).

In addition to the *hydration* waters described above, which may be considered an inherent part of the macromolecule, some water is also likely to be *hydrodynamically* part of the diffusing particle due to the shape irregularities of the macromolecule. There is evidence that water flow in surface crevices of irregularly shaped particle aggregates (analogous to the surface crevices on proteins or the major groove around DNA duplexes) is retarded (11). This latter situation would make the fraction of apparent hydration partially dependent on the shape and size of surface crevices. As stated by Tanford, "…sharp indentations on the surface, will naturally contain solvent, and…this "trapped" solvent will travel with the same velocity as the adjoining macromolecular substance…" (12) (cf. Figures 2, 8 and 9 in the main text).

1.3 Effective hydration layers are different for translational and rotational diffusion

A frequently unappreciated aspect of hydration water is that the hydration layer has different effects on translational and rotational diffusion (1). In order to correctly model the hydration layer in calculations of diffusion coefficients this difference must be taken into account. Consideration of the varied solvent velocities around a translating and rotating sphere explains the origin of this difference (13). The velocity components of a solvent molecule for the case of translational diffusion may be visualized as in Supplemental Figure 1 (adapted from (13)).

The apparent velocity of solvent at any point r around a sphere of radius R during translational diffusion at low Reynolds number conditions can be described by two vectors, v_r and v_θ,

$$v_r = v_\infty \left[1 - \frac{3}{2}\left(\frac{R}{r}\right) + \frac{1}{2}\left(\frac{R}{r}\right)^3\right] cos\theta \tag{S2a}$$

$$v_\theta = v_\infty \left[-1 + \frac{3}{4}\left(\frac{R}{r}\right) + \frac{1}{4}\left(\frac{R}{r}\right)^3\right] sin\theta \tag{S2b}$$

where v_∞ is either the velocity of the sphere in stationary bulk solvent or the velocity of the bulk fluid around a stationary sphere. The relative velocity of fluid at $r = R$ is zero (stick boundary conditions), and approaches v_∞ as $1/r$ moving away from the surface.

For rotational diffusion only, the solvent angular velocity at point r, relative to the sphere surface, is equal to,

$$v_\emptyset = v_\Omega R \frac{R^2}{r^2} sin\theta \tag{S3}$$

where v_Ω is the angular velocity of the sphere and the relative velocity approaches bulk values as $1/r^2$. Because the solvent velocity field around a rotating sphere decays as $1/r^2$,

but decays as 1/*r* around a translating sphere, the viscous energy dissipation must occur in a thinner shell around a rotating sphere. This results in an apparent *increased* effect of the solvation water on rotational diffusion and must be compensated appropriately when calculating hydrated radii for rotational diffusion.

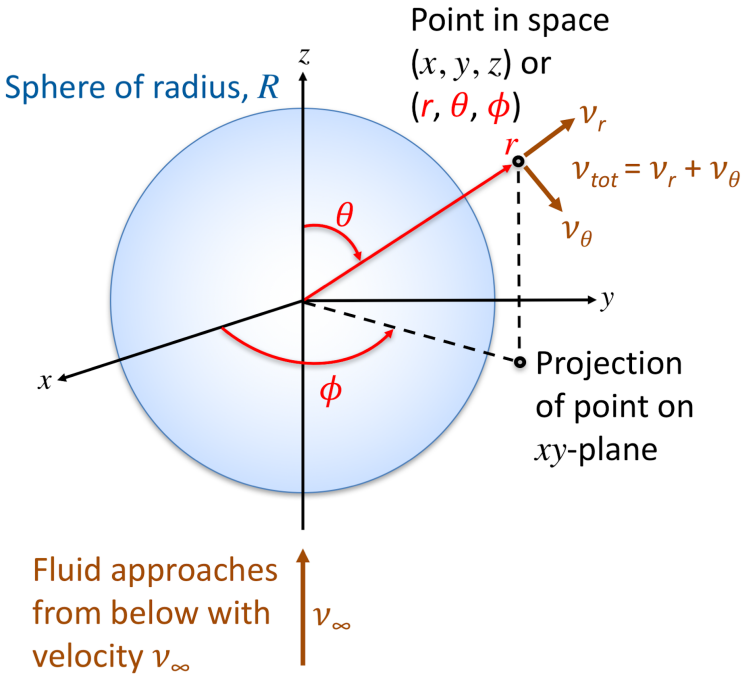
1.4 The effects of shape are different for translational and rotational diffusion

Most proteins are not perfectly spherical in shape and it is common to model this non-sphericity with ellipsoids, either ellipsoids of revolution where two axes are identical or general tri-axial ellipsoids where all three axes may be of different lengths. For identical volumes, an ellipsoid will have more surface area compared to a sphere and will have greater friction during diffusion. Analytical expressions for the dependence of the diffusional frictional coefficient on the axial ratio of ellipsoids of revolution have been worked out for both translational diffusion (14) and rotational diffusion (15, 16). For translational diffusion, the friction is averaged over random orientations of the ellipsoid and, for example, an axial ratio of 3 results in a ~10% increase in the frictional coefficient; similar effects are observed for both prolate and oblate ellipsoids with axial ratios <10 where the difference between prolate and oblate ellipsoids is 5.5%.

For rotational diffusion, the situation is more complex. There are two frictional coefficients, one for rotation about the major axis and one for rotation about the minor axis. The dependence of these coefficients on axial ratios is significantly different for prolate and oblate ellipsoids and the rotational frictional coefficients are generally greater than the translational coefficients. In the case of rotational diffusion, axial ratios of only 1.5 result in significant increases in friction (17). For elongated molecules, the rotational diffusion is better described by anisotropic tumbling rather than axially symmetric tumbling. Although it is not possible to analytically define a single rotational frictional coefficient, the practical outcome is that for NMR data analysis of small monomeric proteins an axially symmetric model is sufficient to describe the measured tumbling (6). Such a simplification may not be accurate for multi-domain proteins or very elongated structures. Therefore, as stated in the main text, the HullRad program prints a warning to the user concerning rotational properties if the axial ratio of a particular molecule is greater than that tested here by comparison to experimental data (a/b=2.62).

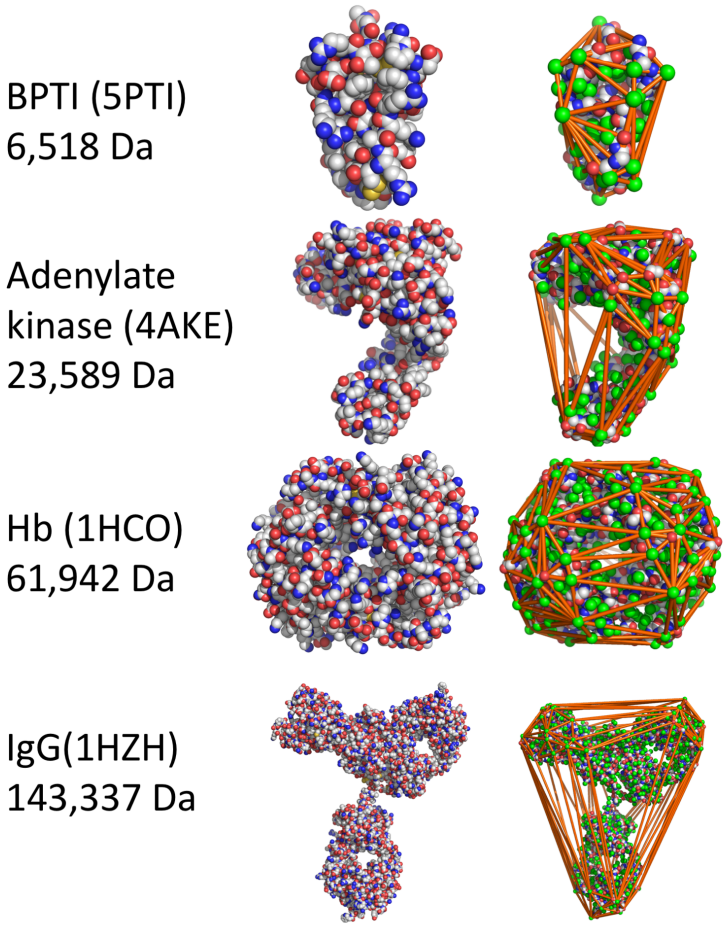
2 Supplemental Figures

2.1 Supplemental Figure 1



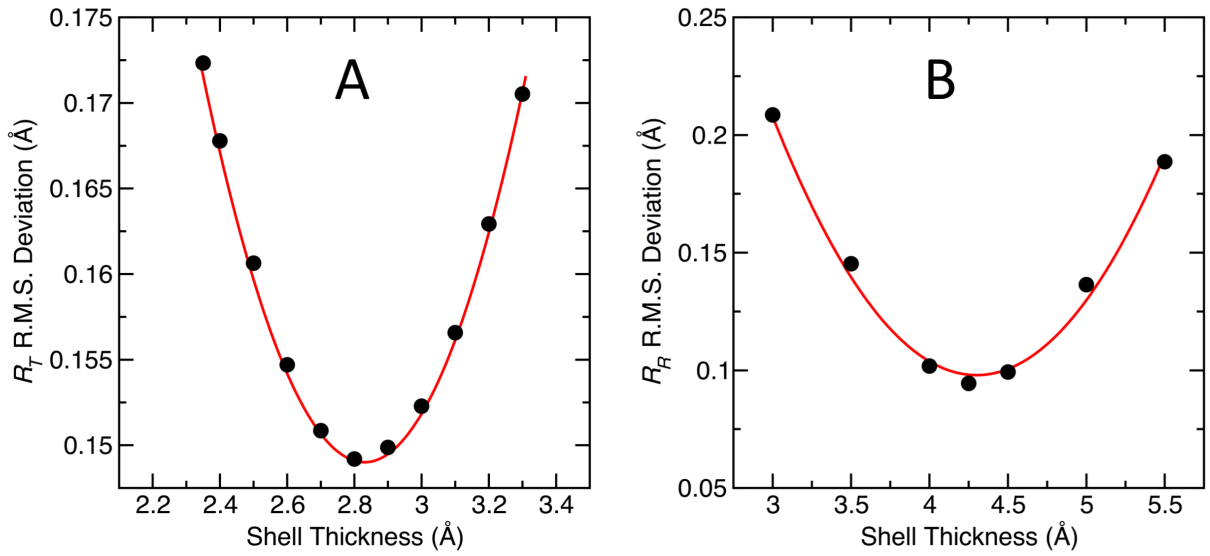
Supplemental Figure 1. Solvent velocity components around a sphere for flow at low Reynolds number. Here the fluid is flowing around a stationary sphere but the relative relationships are general and applicable to the case of a sphere diffusing in a stationary fluid. Adapted from (13). This figure was created in Microsoft PowerPoint.

2.2 Supplemental Figure 2



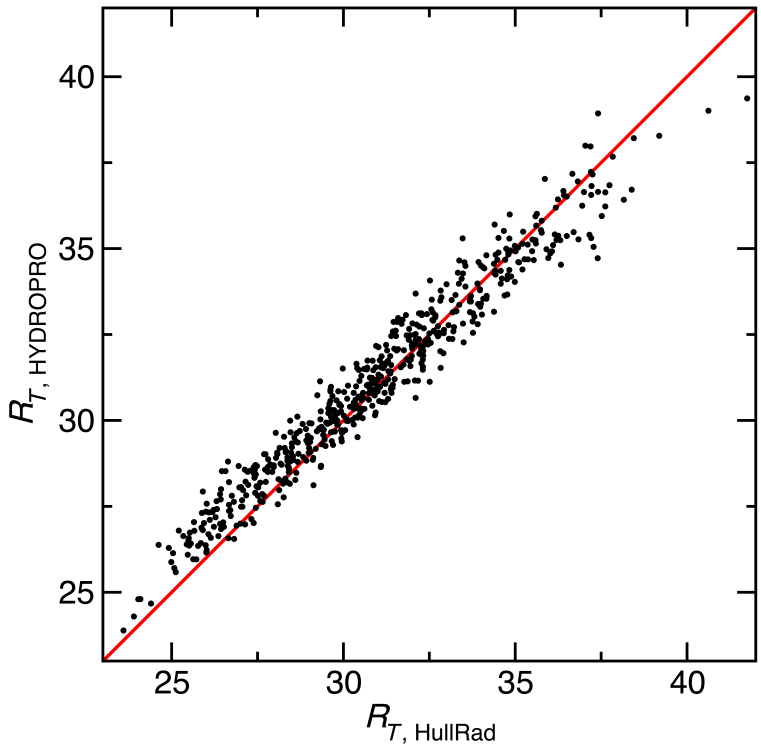
Supplemental Figure 2. Representation of the pseudo-atom side chain protein model. Left column: common name, PDB file and molecular mass; Middle column: atomic sphere representations; Right column: Unified atom side chain model (green side chain pseudo-atoms) with convex hull edges (orange sticks). Vertices of the convex hull are at the centers of the outer-most backbone atoms and side chain pseudo-atoms of the model.

2.3 Supplemental Figure 3



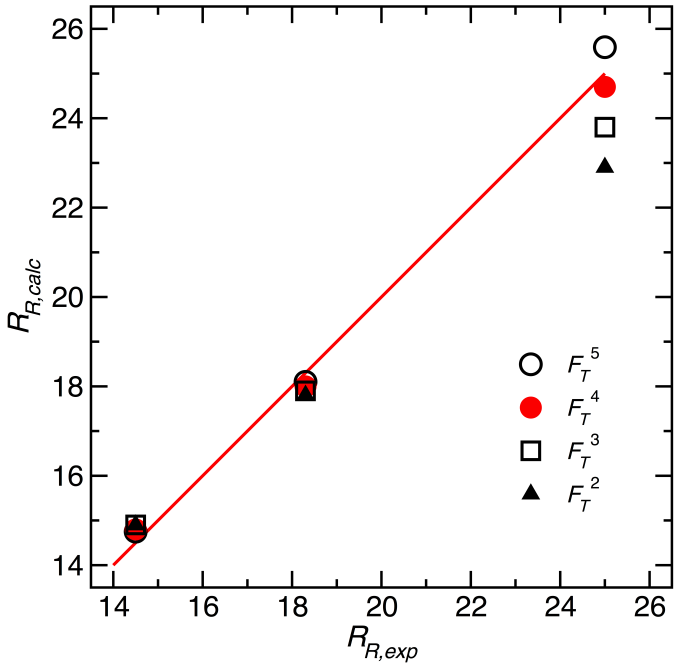
Supplemental Figure 3. Optimization of the hull expansion value to account for hydration. A. Translational hydrodynamic radii (R_T) corrected with a prolate ellipsoidal shape factor (F_T) were calculated for the proteins listed in Table 1 and the root mean squared deviations of convex hull to experimental values for different hydration shell thicknesses are shown as black circles. The data were fit to a quadratic expression (red line) and the minimum deviation is obtained at 2.83 Å shell thickness. B. Rotational hydrodynamic radii (R_R) corrected with a prolate ellipsoidal shape factor ($F_R = F_T^4$) were calculated for the proteins listed in Table 2 and the root mean squared deviations of convex hull to experimental values for different hydration shell thicknesses are shown as circles. The data were fit to a quadratic expression (red line) and the minimum deviation is obtained at 4.30 Å shell thickness.

2.4 Supplemental Figure 4



Supplemental Figure 4. Comparison of the effective translational hydrodynamic radii of disordered state ensemble protein structures calculated by HullRad and HYDROPRO. The red line represents a slope of one and intercept of zero; the correlation coefficient, *R*, for a linear regression of the data (not shown) is 0.95. The HYDROPRO hydrodynamic radii are, on average, ~0.7% larger than the HullRad hydrodynamic radii, but this is largely due to larger predicted values for compact structures. The ensemble contains 575 generated structures of α-synuclein from the Protein Ensemble Database (Accession number PED9AAC).

2.5 Supplemental Figure 5



Supplemental Figure 5. Comparison of calculated to experimental R_R using different rotational shape factors. The rotational hydrodynamic radius (R_R) was calculated using HullRad with different rotational shape factors as follows: Open circles, $F_R = F_T^5$; red solid circles, $F_R = F_T^4$; open square, $F_R = F_T^3$; solid triangle, $F_R = F_T^2$. The red line represents a slope of one and intercept of zero. The plotted data are from Supplemental Table 1.

3 Supplemental Tables

3.1 Supplemental Table 1

Shape Factor Correction for Rotational Diffusion[a]

DNA Duplex	$R_{R,\text{exp}}$	$R_{R,\text{calc}}{}^{b}$ $F_T{}^{5}$	$R_{R,\text{calc}}{}^{c}$ $F_T{}^{4}$	$R_{R,\text{calc}}{}^{d}$ $F_T{}^{3}$	$R_{R,\text{calc}}{}^{e}$ $F_T{}^{2}$	Axial Ratio[f]
8mer	14.5	14.8	14.8	14.9	15.0	1.13
12mer	18.3	18.1	18.0	17.9	17.8	1.61
20mer	25.0	25.6	24.7	23.8	22.9	2.62

[a]The rotational hydrodynamic radii for three DNA duplexes were calculated with different rotational shape factors. In each case the optimal hydration shell expansion was determined for the specific shape factor using the protein data set listed in Table 2 in the main text in a manner similar to that shown in Supplemental Figure 3B.

[b]A shell expansion of 4.2 Å and shape factor equivalent to $F_T{}^{5}$ as the F_R was used.

[c]A shell expansion of 4.3 Å and shape factor equivalent to $F_T{}^{4}$ as the F_R was used.

[d]A shell expansion of 4.4 Å and shape factor equivalent to $F_T{}^{3}$ as the F_R was used.

[e]A shell expansion of 4.5 Å and shape factor equivalent to $F_T{}^{2}$ as the F_R was used.

[f]Axial ratio of a prolate ellipsoid of revolution with volume equal to convex hull volume of the molecule.

4 References

1. Kuntz, I. D., Jr., and W. Kauzmann. 1974. Hydration of proteins and polypeptides. Adv Protein Chem 28:239-345.
2. Aragon, S., and D. K. Hahn. 2006. Precise boundary element computation of protein transport properties: Diffusion tensors, specific volume, and hydration. Biophys J 91:1591-1603.
3. Garcia De La Torre, J., M. L. Huertas, and B. Carrasco. 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J 78:719-730.
4. Rocco, M., and O. Byron. 2015. Hydrodynamic Modeling and Its Application in AUC. Methods Enzymol 562:81-108.
5. Venable, R. M., E. Hatcher, O. Guvench, A. D. Mackerell, Jr., and R. W. Pastor. 2010. Comparing simulated and experimental translation and rotation constants: range of validity for viscosity scaling. J Phys Chem B 114:12501-12507.
6. Ryabov, Y. E., C. Geraghty, A. Varshney, and D. Fushman. 2006. An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. J Am Chem Soc 128:15432-15444.
7. Kang, E. H., M. L. Mansfield, and J. F. Douglas. 2004. Numerical path integration technique for the calculation of transport properties of proteins. Phys Rev E Stat Nonlin Soft Matter Phys 69:031918.
8. Makarov, V. A., B. K. Andrews, P. E. Smith, and B. M. Pettitt. 2000. Residence times of water molecules in the hydration sites of myoglobin. Biophys J 79:2966-2974.
9. Denisov, V. P., and B. Halle. 1996. Protein hydration dynamics in aqueous solution. Faraday Discuss 103:227-244.
10. Halle, B., and M. Davidovic. 2003. Biomolecular hydration: from water dynamics to hydrodynamics. Proc Natl Acad Sci U S A 100:12135-12140.
11. Babick, F. 2016. Suspensions of Colloidal Particles and Aggregates. Springer International Publishing, Switzerland, 171.
12. Tanford, C. 1961. Physical Chemistry of Macromolecules. Wiley, New York, 337.
13. Bird, R. B., W. E. Stewart, and E. N. Lightfoot. 2002. Transport Phenomena. J. Wiley, New York,
14. Perrin, F. 1936. Mouvement Brownian d'un ellipsoide II. Rotation libre et dépolarisation des fluorescences. Translation et diffusion de molécules ellipsoïdales. Journal de Physique et Le Radium 7:1-11.
15. Edwardes, D. 1893. Steady motion of a viscous liquid in which an ellipsoid is constrained to rotate about a principal axis. The Quarerly Journal of Pure and Applied Mathematics 26:70-78.
16. Favro, L. D. 1960. Theory of the Rotational Brownian Motion of a Free Rigid Body. Phys Rev 119:53-62.
17. Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry. Part II. Techniques for the study of biological structure and function. W. H. Freeman, New York, 561-565.