Solution Structure of Biopolymers: A New Method of Constructing a Bead Model

Ewa Banachowicz, Jacek Gapiński, and Adam Patkowski
Molecular Biophysics Laboratory, Institute of Physics, Adam Mickiewicz University, Umultowska 85, 61-614 Poznań, Poland

ABSTRACT We propose a new, automated method of converting crystallographic data into a bead model used for the calculations of hydrodynamic properties of rigid macromolecules. Two types of molecules are considered: nucleic acids and small proteins. A bead model of short DNA fragments has been constructed in which each nucleotide is represented by two identical, partially overlapping spheres: one for the base and one for the sugar and phosphate group. The optimum radius σ = 5.0 Å was chosen on the basis of a comparison of the calculated translational diffusion coefficients (D_T) and the rotational relaxation times (τ_R) with the corresponding experimental data for B-DNA fragments of 8, 12, and 20 basepairs. This value was assumed for the calculation D_T and τ_R of $tRNA^{Phe}$. Better agreement with the experimental data was achieved for slightly larger σ = 5.7 Å. A similar procedure was applied to small proteins. Bead models were constructed such that each amino acid was represented by a single sphere or a pair of identical, partially overlapping spheres, depending on the amino acid's size. Experimental data of D_T of small proteins were used to establish the optimum value of σ = 4.5 Å for amino acids. The lack of experimental data on τ_R for proteins restricted the tests to the translational diffusion properties.

INTRODUCTION

The structure of biopolymers has been a subject of investigation ever since they were discovered. So far, the most accurate methods for structure determination have been x-ray diffraction techniques. These techniques are, however, restricted to crystallizable molecules, and it has not been established if the molecular structure in a crystal is identical to that in solution. Many methods have been developed or adopted to probe biopolymer properties in solution. The most common are dynamic light scattering (Berne and Pecora, 1976), transient electric birefringence (Eden and Elias, 1983), and nuclear magnetic resonance and ultracentrifugation (Cantor and Schimmel, 1980). Hydrodynamic properties measured by these methods do not allow construction of molecular models with atomic resolution (except NMR for small proteins), but still, a lot of effort has been put into obtaining as much information as possible from the experimental results. Hydrodynamic properties, for example translational D_T and rotational D_R diffusion coefficients, are converted into size and shape parameters of model structures. The simplest model is a sphere with only one parameter: radius R. The famous Stokes-Einstein relation $D_T = kT/(6\pi\eta R)$ allows the calculation of a radius of the model sphere. However, until 1966 exact relations were restricted to spheres, ellipsoids of revolution (Perrin, 1936), and approximate for cylinders (Broersma, 1960a, b), which do not reflect the actual shapes of real macromolecules. Attempts to apply models of arbitrary shape are based on the so-called bead modeling technique, where the actual shape is approximated by an ensemble of spheres (beads) of given radii and positions. The choice of the bead size depends on the molecule size and the accuracy required. The basis for the method is the theory of irreversible transport properties of rigid macromolecules developed by Kirkwood (1954). Its present form is a result of extensive work of several groups of researchers (see Garcia de la Torre and Bloomfield, 1981 for details). The first applications to simple structures were aimed at testing the theory. In 1975 Teller and de Haen carried out calculations of diffusion coefficient for globular proteins taking into account only the surface atoms of the molecule (Teller and de Haen, 1975). In a subsequent paper (Teller et al., 1979) they report on the calculations in which the shape of the protein is modeled by a layer of beads coating the molecule (shell model). A similar approach was applied later by Pastor and co-workers (Pastor and Karplus, 1988; Venable and Pastor, 1988). The high numerical cost of this method strongly restricts its application to smaller biopolymers. Bloomfield and Garcia de la Torre modeled proteins in the form of large domains so that the total number of beads was limited to a few tens. From the point of view of accuracy, the atomic model seems to be a better choice, but the numerical cost restricts its applicability to molecules consisting of no more than a few thousand atoms. However, a model consisting of a few large domains is probably too crude to follow small differences in D_T and the construction of a bead model is rather arbitrary. Garcia de la Torre et al.(1994a) proposed a good solution for DNA fragments: each bead corresponds to a single nucleotide. The procedure was fully reproducible because the beads' centers were placed according to the helix equation.

We should also mention here works of Porschke (e.g., Antosiewicz and Porschke, 1989) devoted to improving different bead models of biological macromolecules, and the more general approach of Felderhof, e.g., analytical calculations of multisubunit shell frictional properties (in

Received for publication 9 April 1999 and in final form 20 October 1999.
Address reprint requests to Dr. Jacek Gapiński, A. Mickiewicz University, Institute of Physics, Umultowska 85, 61-614 Poznań, Poland. Tel.: 48-61-8273265; Fax: 48-61-8257758; E-mail: gapior@amu.edu.pl.

0006-3495/00/01/70/09 $2.00

McCammon et al., 1975), and dynamics of interacting particles (in Felderhof, 1996, and references therein).

Recently, a few reports appeared devoted to different methods of constructing bead models of biological macromolecules. Byron (1997) developed a method based on the local density of atoms. The space occupied by the molecule was divided into small cubes. In each of them there is a bead, the size and position of which is determined by the position of all the atoms in that cube. Such an approach allows an easy adjustment of the model resolution (cube size) to the molecule size. Spotorno et al. (1997) developed a whole set of computer programs (BEAMS) for generation, visualization, and computation of the hydrodynamic and conformational properties of bead models of proteins. The main aim of their method was to construct a bead model from low-resolution experimental or simulated data as a tool for testing hypothetical molecule conformations. Hellwig et al. (1997) constructed a bead model of a medium-size protein placing beads in the positions of C^{α} atoms. The calculated D_T values were then compared to experimental data. An alternative method for the calculation of hydrodynamic properties was proposed by Zhou (1995a,b). He estimated them through the relations between translational friction and capacitance and between intrinsic viscosity and polarizability. Those electrostatic properties were, in turn, calculated by means of the boundary-element technique.

Encouraged by those results, we tried to develop a fully automatic algorithm that would convert any crystallographic data into an appropriate bead model. Since we intended to apply it to macromolecules of arbitrary shape, the algorithm could not be based on any equation (as, for example, the helix for DNA), but only on the atomic coordinates. Nucleic acids seemed to be much simpler than proteins in terms of structure and homogeneity, so we decided to begin with short DNA fragments. Having experimental data of both translational and rotational diffusion coefficients of different DNA fragments, we tried to adjust our algorithm to provide satisfactory agreement between the measured and calculated values regardless of the DNA length. A similar procedure was performed for proteins, although due to much more complex structure and lack of extensive experimental data, our attempts are still subject to constant improvement. Also, the lack of τ_R experimental data in this case restricted the tests to the translational diffusion coefficients.

THEORY

Hydrodynamic interactions

The origin of the bead modeling theory was well described in Garcia de la Torre and Bloomfield, 1981. Here we give only a brief description of the method.

Let us consider a rigid assembly of N beads immersed in a continuous medium. The actual velocity v_i' of the ith bead is changed due to hydrodynamic interactions with other beads:

$$v_i' = v_i - \sum_{j=1,j\neq i}^{N} \mathbf{T}_{ij} F_j, \tag{1}$$

where v_i is the unperturbed velocity of the ith bead, F_i is the frictional force acting on the ith bead, and $\mathbf{T}_{ij}$ is the tensor describing hydrodynamic interactions between the ith and jth beads. A simple multiplication by the ith bead's frictional coefficient ζ_i gives a similar equation for the frictional force F_i:

$$F_i = \zeta_i v_i - \zeta_i \sum_{j=1,j\neq i}^{N} \mathbf{T}_{ij} F_j \tag{2}$$

where ζ_i is given by Stokes' law ($\zeta_i = 6\pi\eta_0\sigma_I$), η_0 being the viscosity of the solvent and σ_i the radius of the ith bead. For the sake of convenience, the so-called shielding tensors $\mathbf{G}_i$ are introduced

$$F_i = \zeta_i \mathbf{G}_i v_i \tag{3}$$

from which the translational friction tensor Ξ may be calculated:

$$\Xi = \sum_{i=1}^{N} \zeta_i \mathbf{G}_i \tag{4}$$

The translational diffusion coefficient D_T of the whole assembly (molecule) is given by

$$D_T = \frac{kT}{f_T} \tag{5}$$

where k is Boltzmann's constant, T is temperature, and f_T is the translational friction coefficient defined as

$$f_T = 3/Tr(\Xi^{-1}) \tag{6}$$

As the model is analytically insoluble for more then two beads, the main problem of the theory is to determine the most accurate form of the hydrodynamic interactions tensor $\mathbf{T}_{ij}$. For the case of separate beads we applied the form given by Garcia de la Torre and Bloomfield (1981):

$$\mathbf{T}_{ij} = (8\pi\eta_0 R_{ij})^{-1}\left[\mathbf{I} + \frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{R_{ij}^2} + \frac{(\sigma_i^2 + \sigma_j^2)}{R_{ij}^2}\left(\tfrac{1}{3}\mathbf{I} - \frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{R_{ij}^2}\right)\right], \tag{7}$$

while for overlapping elements (of the same radius) the formula developed by Rotne and Prager (1969) was applied:

$$\mathbf{T}_{ij} = (6\pi\eta_0\sigma)^{-1}\left[\left(1 - \frac{9}{32}\frac{R_{ij}}{\sigma}\right)\mathbf{I} + \frac{3}{32}\frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{R_{ij}\sigma}\right]. \tag{8}$$

MATERIALS AND METHODS

The program HI4, kindly provided by Venable and Pastor, was used for the calculations of the model's hydrodynamic properties. The source code was modified to include the Rotne-Prager interaction tensor for the case of overlapping spheres. All the calculations were performed on a regular PC using Microsoft Fortran Compiler for Windows. For some of the structures a large RAM was required (up to 256 MB). Values of D_T and τ_R (taken as $\tau_\perp$) were calculated for models with different bead sizes. Molecular structures of appropriate DNA fragments were generated using a computer program (HyperChem, Hypercube Inc., Windows version 4.0).

Construction of a bead model

In order to take full advantage of the atomic coordinates we decided to group atoms and place the beads at the geometrical centers of the groups

$$\mathbf{r}_j = \sum_{i=1}^{n_j} \mathbf{r}_j^i / n_j \qquad (9)$$

where the vector $\mathbf{r}_j$ points at the jth bead's center and n_j is the number of atoms in the jth group of atoms. We decided to include all the atoms (hydrogens) and use small groups (10–30 atoms) to maintain the actual shape of the molecule surface. Since our preliminary calculations for such models showed that beads have to overlap to give reasonable results, we restricted our approach to the case of identical beads.

All our computer programs for constructing the bead models were written in Turbo Pascal for DOS. The source code and executable versions will be available upon request from the corresponding author.

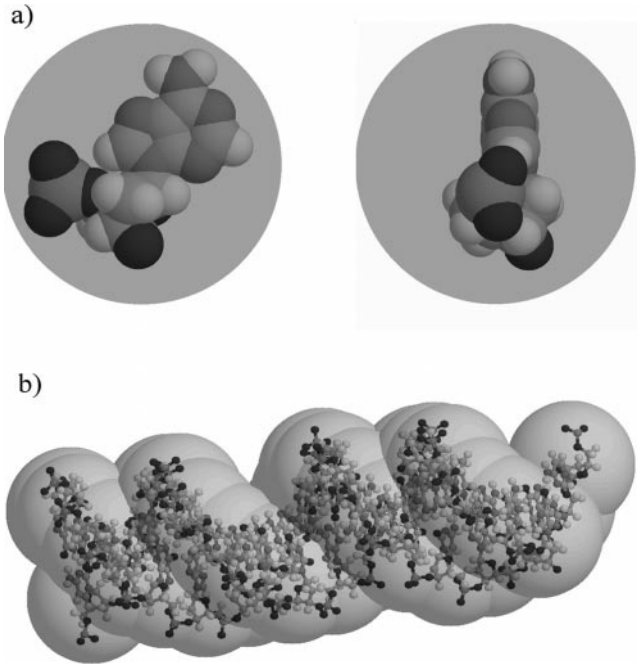

FIGURE 1 Graphical representation of the single-bead model of nucleic acids. (*a*) A schematic view of the arrangement of adenine atoms inside a model bead in two projections; (*b*) a schematic view of a B-DNA 20-mer in a ball-and-stick representation immersed in its bead model.

RESULTS

Nucleic acids

Fig. 1 *a* shows schematic views of a nucleotide immersed in its model bead. To test the procedure applicability, we compared the calculated values of D_T and τ_R with experimental dynamic light scattering data available for short DNA fragments (8, 12, and 20 basepairs, Eimer and Pecora, 1991). The choice of the experimental method used for comparison with the bead model calculations was a consequence of the results of Eimer et al. (1990). In their article they compared relaxation times obtained from NMR and depolarized dynamic light scattering measurements of short oligonucleotides to the reorientation times of model cylinders calculated using the formulas derived by Broersma and Garcia de la Torre. It was found that light scattering data describe the reorientation of the whole molecule around its shorter axis, while the NMR cross-relaxation times were superpositions of reorientational times around both axes and the internal local motions of the molecules. A schematic view of the DNA 20-mer surrounded by its bead model is shown in Fig. 1 *b*. We calculated D_T and τ_R for different values of bead radius σ in the range 6 Å $< \sigma <$ 8.5 Å. On the basis of the results presented in Fig. 2, a value of σ = 7.3 Å was chosen for further calculations.

The same procedure was applied to the $tRNA^{Phe}$ molecule (for experimental data see Patkowski et al., 1990a, and for the structure model see Sussman et al., 1978). The results of the calculations shown in Fig. 3 fit the experimental data for the same value of σ = 7.3 Å very well. At first this result seems to be quite reasonable and we regarded it as a positive verification of our approach. However, hydration of RNAs and DNAs can be different, depending on particular RNA local structure. Looking at Fig. 1 *a* one can see that there is a lot of empty space inside the bead, especially out of the base plane. It does not matter in the case of a double-stranded structure of DNA because this part is mostly buried inside the model, and due to the overlap of the beads cancels out in calculations. But in the case of RNA that empty space may become a source of error because of different packing. To minimize this effect and to be able to take into account different sizes of nucleotides containing purines and pyrimidines, we applied a "double bead" model both for DNAs and tRNA. In this model each nucleotide was divided into two groups of atoms: one containing the base and the other one containing the sugar and the phosphate group (Fig. 4). Due to the method's constraints, both beads had to have the same size. Again, calculations were performed for different σ values 3 Å $< \sigma <$ 8 Å. The results, presented in Fig. 5, show the same tendency as for the "single bead" model: shorter DNA fragments seem to be slower in reality than their models for the bead radius corresponding to best-fit value of longer ones. Although the error bars almost overlap, we think this effect could be relevant as a result of the "ends effect." All results have been summarized in Table 1.

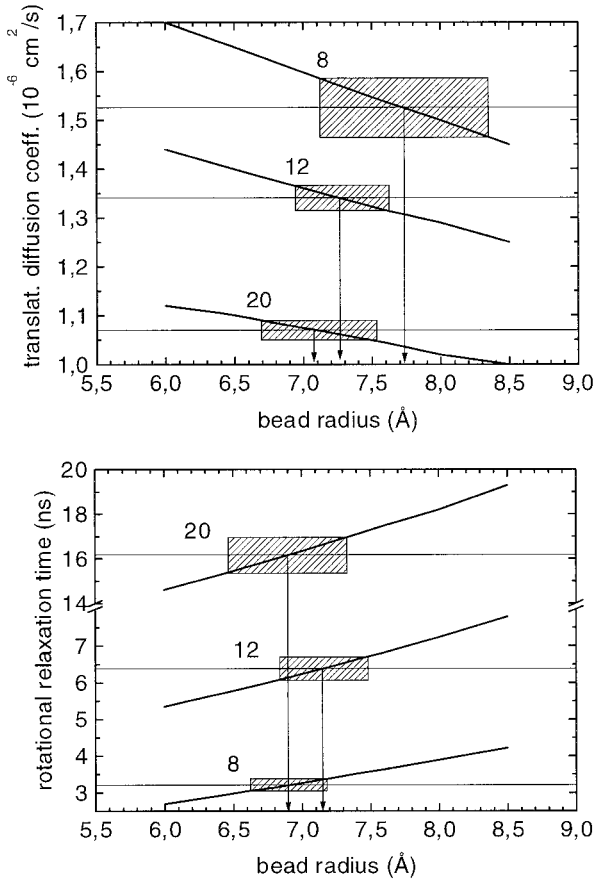

FIGURE 2 Results of single-bead model calculations of translational diffusion coefficients and rotational relaxation times for short DNA fragments; the numbers in the figure denote the number of basepairs; The rectangles represent the error bars defined by the experimental errors (height) and projected on the bead radius (width). Horizontal lines denote the experimental values.

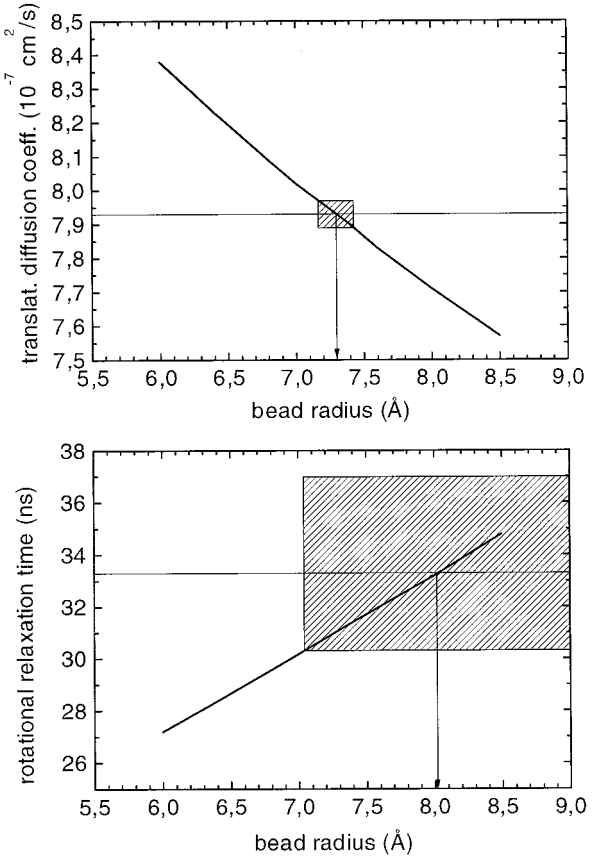

FIGURE 3 Results of single-bead model calculations of translational diffusion coefficients and rotational relaxation times for $tRNA^{Phe}$. Description of the details as in Fig. 2.

Small proteins

The procedure applied to model small proteins is similar to the one applied to nucleic acids. Also in this case we tried both single- and double-bead models. A single bead model was recently applied by Eimer's group (Hellwig et al. 1997) in studies of a large (220 kDa) MoFe protein from *Azotobacter vinelandii* (PDB code 2MIN, Peters et al., 1997). In their model, centers of beads were placed in the position of C^{α} atoms. We adopted this method for our calculations as the single bead model. However, we are convinced that the nonuniform size and shape of the amino acids (Fig. 6) requires a more detailed representation in the bead modeling procedure. We decided to develop a "mixed model" in which eight of the smaller amino acids (Gly, Ala, Val, Ser, Thr, Cys, Pro, Asp) are modeled as single spheres while all the others are modeled as two partially overlapping spheres. The criterion, although a little arbitrary, was the value of an average distance of all the amino acid atoms from their geometrical center (below 2 Å). Again, the bead center is positioned at the geometrical center of all the atoms in a group. In the case of eight small amino acids the group of atoms is equivalent to the residuum and the bead center is placed in the geometrical center of all its atoms. For larger amino acids one group is always composed of 10 atoms: the backbone atoms and the surrounding of the C^{β} atom, while the second one contains the rest of the atoms. In all cases the center of the first bead is very close to the C^{α} atom. The center of the other one is placed at the geometrical center of the remaining group of atoms. Fig. 6 shows the effects of application of such a model. In Fig. 6 *a* we present a small amino acid modeled with a single sphere and one of the largest amino acids modeled with two, almost separated spheres. A model of a small protein is shown in Fig. 6 *b*.

For tests we could use only those proteins for which both crystallographic structure and experimental values of diffusion coefficients were available. Six small proteins were chosen for the test calculations: bovine trypsin inhibitor (bovine pancrease), profilin (*Acanthamoeba castellanii*), insulin (hexamer, pig), ribonuclease A (bovine pancrease), lysozyme (turkey egg white), and cellulase (*humicola insolens*). The crystallographic data were taken from the Brookhaven Data Bank. Basic information about the pro-

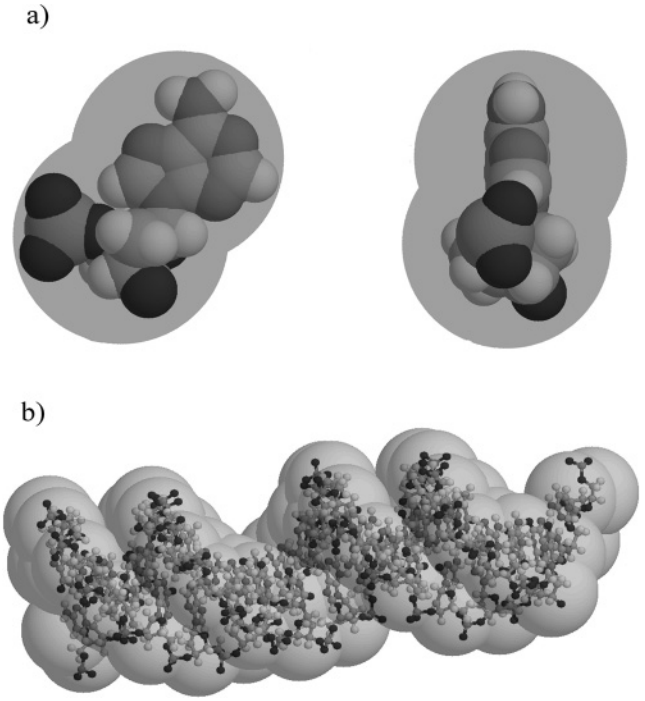

FIGURE 4 Graphical representation of the double-bead model of nucleic acids. (*a*) A schematic view of the arrangement of adenine atoms inside a model bead in two projections; (*b*) a schematic view of a B-DNA 20-mer in ball-and-stick representation immersed in its bead model.

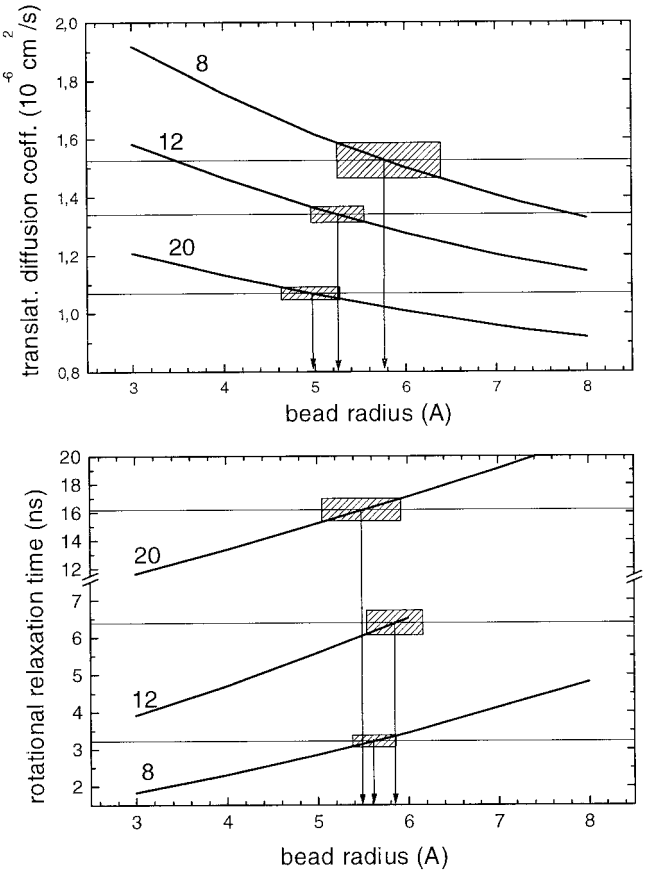

FIGURE 5 Results of double-bead model calculations for short DNA fragments. Details as in Fig. 2.

teins are summarized in Table 2. Due to the lack of experimental data on rotational relaxation, only the values of D_T were calculated for different values of σ and compared to the experimental data (Fig. 7). A substantial error of the experimental data is reflected in the uncertainty of the σ_{best} value: ± 0.3 up to ± 0.8 Å (7–18%). The average value of $\sigma_{av} = 4.5$ Å fits the experimental data for most of the proteins tested.

It seems that the introduction of the second sphere is important only for small proteins, for which a detailed structure of the surface may play a big role, while for the large ones one sphere per amino acid, centered, e.g., at the position of the C^{α} atom, is sufficient. We checked both methods on the series of six small proteins. The numbers presented in the first two columns of Table 3 support that observation. For smaller proteins (e.g., BPTI, profilin, lysozyme) the difference between these two methods is of an order of 5%, while for insulin hexamer (306 amino acids) the difference decreases to much below 1%. It is clear that for larger structures it is mainly the overall shape that influences the friction coefficient, while for smaller proteins every detail of the surface may have a substantial influence on its hydrodynamic properties.

The choice of the test proteins was strongly affected by three factors: availability of crystallographic and experimental data and the size of the molecule (hardware limitations). The latter problem can be partially solved by applying algorithms calculating the solvent availability at all the points of the surface. We took advantage of the ASC (Analytical Surface Calculation) computer program (version 2.12) written by Eisenhaber (Eisenhaber and Argos, 1993; Eisenhaber et al., 1995). Crystallographic data were converted into our bead model and the resultant files in the xyzr format were analyzed by the ASC program. Beads unavailable to the solvent were rejected from the structure and hydrodynamic calculations were performed on the modified files. The results, presented in the "ASC dbl." column of Table 3, show that the diffusion coefficient of the modified structure is either exactly equal (insulin) or higher by a fraction of a percent (smaller proteins) than D_T of the original model. The fact that such a difference exists, although only beads with zero surface exposed to the solvent were rejected, probably comes from the effect of the finite solvent molecule size used in the ASC calculations. The algorithms calculating hydrodynamic properties of bead models do not take into account the atomic structure of water. Even the smallest holes and pores are penetrated by a continuous solvent.

TABLE 1 Translational diffusion coefficients and rotational relaxation times of short B-DNA fragments and tRNA^{Phe}

		8-mer	12-mer	20-mer	tRNA^{Phe}
Experimental data	D^T_{exp} (10^{-6} cm^2/s)	1.53 ± 0.05*	1.34 ± 0.03*	1.07 ± 0.02†	0.793 ± 0.004†
	τ^R_{exp} (ns)	3.22 ± 0.16*	6.39 ± 0.32*	16.2 ± 0.8*	33 ± 3‡
Single bead model	$D^T_{7.3\ Å}$ (10^{-6} cm^2/s)	1.569	1.337	1.057	0.793
	σ^T_{best} (Å)	7.7 ± 0.6	7.3 ± 0.4	7.1 ± 0.6	7.3 ± 0.1
	$\tau^R_{7.3\ Å}$ (ns)	3.46	6.53	16.9	31.1
	σ^R_{best} (Å)	6.9 ± 0.3	7.13 ± 0.35	6.9 ± 0.4	8.0 ± 1
Double bead model	$D^T_{5.0\ Å}$ (10^{-6} cm^2/s)	1.614	1.362	1.066	
	σ^T_{best} (Å)	5.8 ± 0.6	5.3 ± 0.3	5.0 ± 0.3	5.7 ± 0.1
	$\tau^R_{5.0\ Å}$ (ns)	2.84	5.58	15.2	
	σ^R_{best} (Å)	5.6 ± 0.3	5.85 ± 0.35	5.5 ± 0.4	6.5 ± 1

D^T_{exp} (τ^R_{exp}), experimental value of translational diffusion coefficient D_T (rotational relaxation time τ_R); $D^T_{7.3\ Å}$ ($\tau^R_{7.3\ Å}$), calculated value of D_T (τ_R) for single bead model with σ = 7.3 Å; σ^T_{best} (σ^R_{best}), radii of the beads of the best-fit model calculated from translational (rotational) diffusion data. The respective symbols with subscript "5.0 Å" correspond to the double bead model.

*Eimer and Pecora, 1991.

†Liu et al., 1998.

‡Patkowski et al., 1990a.

Our test calculations have been also extended to the case of large proteins. Two large structures were chosen: MoFe protein investigated recently by Hellwig et al. (1997) with the molecular weight of 220 kd, and ferritin in the *apo* form (no iron) with the molecular weight of 474 kd. Ferritin is a multi-domain protein and the crystallographic data are available only for a single domain (PDB code 1IER, Granier et al., 1997). We took advantage of a data file generated at Washington University, in which symmetry relations were used to generate a whole molecule model from coordinates of atoms of a single domain. Our hardware did not allow for calculations on the original double bead models, and ASC algorithms had to be used for both molecules. In the case of the MoFe protein it was possible to reject only beads with zero available surface and perform D_T calculations (1300 beads). In the case of ferritin over 80% of beads had to be rejected, so we couldn't avoid rejecting beads available to the solvent. In order to estimate the error of such a procedure we performed a series of calculations for the MoFe protein rejecting beads with different areas of surface accessible to the solvent. For the value used in the case of ferritin (33 Å^2) the difference in calculated D_T amounted to 0.54% of that obtained when only beads with no accessible surface were rejected. The results, presented in the last two rows of Table 3, show a tendency of the model calculations to overestimate the mobility of large molecules. In the case of ferritin we were able to perform a comparison with simple geometrical considerations. It is generally accepted that ferritin has the form of a spherical protein shell covering the mineral, iron-rich core. The radius of the cavity is estimated to be 40 Å (Stryer, 1995). Taking a molecular weight of 474 kd and the typical partial specific volume of proteins v = 0.73 cm^3/g (Cantor and Schimmel), we get a value of 58.5 Å for the outer radius, which is >10 Å smaller then the hydrodynamic radius R_h = 69 Å corresponding to the measured diffusion coefficient. The very good quality of the sample (second cumulant below 0.05) and the use of a multi-tau correlator (ALV5000) practically exclude any

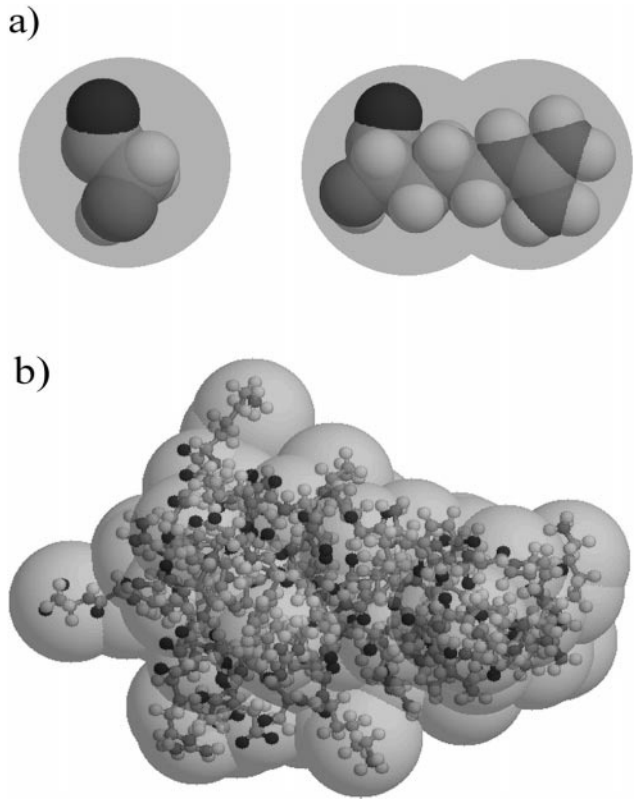

FIGURE 6 Graphical representation of the bead model of proteins. (*a*) A schematic view of glycine and arginine atom arrangement inside the model beads; (*b*) a schematic view of a small protein (BPTI) immersed in its bead model.

TABLE 2 Basic data of the small test proteins

	BPTI	Profilin	Insulin	Ribonuclease A	Lysozyme	Cellulase
PDB code	4PTI*	1PNE[†]	1AI0[‡]	3RN3[§]	1LZ3[¶]	2ENG[‖]
D_{exp} (10^{-6} cm^2/s)	1.44 ±0.03**	1.06 ± 0.02[††]	0.79 ± 0.02[‡‡]	1.17 ± 0.05[¶¶]	1.11 ± 0.05[¶¶]	0.98 ± 0.02[‖‖]
M_w (kDa)	6.4	14.8	34.4	13.6	14	22.0
N_{res}	58	139	306	124	129	210
N_{beads}	86	202	479	181	190	293

D_{exp}, experimental value of translational diffusion coefficient; M_w, molecular weight; N_{res}, number of residues (amino acids); N_{beads}, number of beads in bead model.
*Marquart et al., 1983.
[†]Cedergren-Zeppezauer et al., 1994.
[‡]Takahara et al., 1995.
[§]Howlin et al., 1989.
[¶]Harata, 1993.
[‖]Davies et al., 1993.
**Gallagher and Woodward, 1989.
[††]Patkowski et al., 1990b.
[‡‡]Hvidt, 1991.
[¶¶]Kuntz and Kauzmann, 1974.
[‖‖]E. Banachowicz, C. Boisset, M. Schülein, R. Borsali, B. Henrissat, and A. Patkowski, manuscript in preparation.

possibility of experimental artifacts, such as the presence of ferritin dimers or larger aggregates, which might increase the best-fit hydrodynamic radius. To verify this discrepancy we also measured the distance between the two most distant molecules in the atomic model of the ferritin molecule. The result of 135 Å corresponds more to the measured value than to the calculated one. We think that the source of discrepancy lies in two facts: 1) the domains have the form of long barrels that cannot form a smooth spherical surface, which results in the presence of many edges and holes in the protein shell of ferritin; and 2) the same applies to the inner cavity for which the diameter of 80 Å is the smallest one and so the calculations above have wrong assumptions. A uniformly distributed water shell on the surface, which is incorporated in our model, seems not to work properly in the conditions where large portions of water form kinds of puddles that fill all the grooves and holes in the surface of large multi-domain molecules. The same argumentation applies to the MoFe protein, which consists of four subunits bound in such a way that a lot of water can be immobilized in the space between them. Hydration of smaller molecules probably cannot be as efficient simply because there is not enough space to form large grooves or holes in the structure of the molecule.

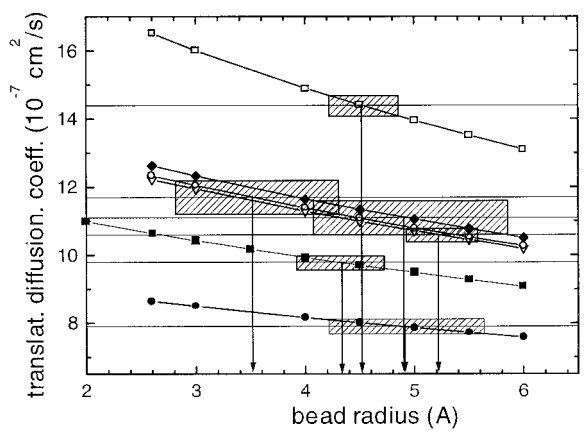

FIGURE 7 Results of bead model calculations for small proteins. The horizontal lines denote experimental values. *Symbols:* □, BPTI; ●, insulin hexamer; ■, cellulase; ♦, lysozyme; ▽, profilin; ○, ribonuclease A. The rectangles represent experimental error bars.

CONCLUSIONS

In this paper we have presented a systematic automated procedure for constructing bead models of biomacromolecules from the atomic coordinates. Despite the model's simplicity, the results presented here are quite promising. Comparing the hydrodynamic properties calculated from these models with experimental data, one can verify or optimize the ternary structure of biopolymers in solution. We are aware of the basic problem in our approach; that is, the nonuniform distribution of water on the surface of macromolecules, which is not implemented in the model. Improved versions of the method including calculations of hydration properties of macromolecules are already considered in our group. More model proteins should be considered in the testing procedure, preferably with experimental data from one source.

Some numerical problems arise with the growing size of the molecules. The amount of computer memory required to calculate D_T is proportional to $(3N)^2$, N being the number of beads. Hence, for large N, the solution of the problem becomes more expensive and time-consuming. Nevertheless, with the increasing capabilities of computers today, this restriction will probably disappear soon. It is possible and quite reasonable in terms of run time to analyze molecules modeled using 1000 beads even on medium-size

TABLE 3 Comparison of calculation results (translational diffusion coefficient D_T [10^{-6} cm^2/s]) for different bead models

Protein	Single bead	Double bead	ASC* dbl.	Experimental value	σ_{best} (Å)
BPTI	1.541	1.441	1.446	1.44 ± 0.03	4.5 ± 0.3
Profilin	1.136	1.099	1.104	1.06 ± 0.02	5.2 ± 0.4
Insulin	0.798	0.801	0.800	0.79 ± 0.02	4.9 ± 0.7
Ribonuclease A	1.156	1.108	1.111	1.17 ± 0.05	3.5 ± 0.8
Lysozyme	1.181	1.134	1.139	1.11 ± 0.05	4.9 ± 0.8
Cellulase	1.018	0.970	0.985	0.98 ± 0.02	4.4 ± 0.4
MoFe	0.462†	—	0.441	0.40 ± 0.02†	—
Ferritin	—	—	0.330	0.31 ± 0.01‡	—

In the "single-bead model" spheres (only one for each amino acid) are positioned on the C^α atoms. Total bead radius is 6.6 Å (as in Hellwig et al., 1997). "Double-bead model" corresponds to the model described in this paper with the total bead radius of 4.5 Å. The "ASC double" denotes results obtained after removing all solvent-inaccessible beads from the double bead model. Symbol σ_{best} denotes the value of σ corresponding to the experimental value of D_T.
*Analytical surface calculation.
†Hillweg et al., 1997.
‡Data from our lab, to be published.

workstations. The size of the largest variable is $3N \times 3N \times 16$ (double precision 16 bytes per number). For $N = 1000$ the variable size amounts to 140 MB. In principle, at the time of writing this paper, a fast PC equipped with sufficient RAM amount (256 MB) can also handle such a task.

Partial financial support of Volkswagen Stiftung, Federal Republic of Germany, and the US-Polish Maria Sklodowska-Curie Joint Fund II Grant MEN/NSF-96-254 is gratefully acknowledged.

REFERENCES

Antosiewicz, J., and D. Porschke. 1989. Volume correction for bead model simulations of rotational friction coefficients of macromolecules. *J. Phys. Chem.* 93:5301–5305.

Berne, B., and R. Pecora. 1976. Dynamic Light Scattering. John Wiley and Sons, New York.

Broersma, S. 1960a. Rotational diffusion constant of a cylindrical particle. *J. Chem. Phys.* 32:1626–1631.

Broersma, S. 1960b. Viscous force constant for a closed cylinder. *J. Chem. Phys.* 32:1632–1635.

Byron, O. 1997. Construction of hydrodynamic bead models from high-resolution x-ray crystallographic and nuclear magnetic resonance data. *Biophys. J.* 72:408–415.

Cantor, P., and R. Schimmel. 1980. Biophysical Chemistry, vol. II. Techniques for the Study of Biological Structure and Function. W. H. Freeman and Co., New York.

Cedergren-Zeppezauer, E. S., N. C. W. Goonesekere, M. D. Rozycki, J. C. Myslik, Z. Dauter, U. Lindberg, and C. E. Schutt. 1994. Crystallization and structure determination of bovine profilin at 2.0 Angstroms resolution. *J. Mol. Biol.* 240:459–475.

Davies, G. J., G. G. Dodson, R. E. Hubbard, S. P. Tolley, Z. Dauter, K. S. Wilson, C. Hjort, J. M. Mikkelsen, G. Rasmussen, and M. Schulein. 1993. Structure and function of endoglucanase V. *Nature.* 365:362–364.

Eden, D., and J. G. Elias. 1983. *In* Measurement of Suspended Particles by Quasi-Elastic Light Scattering. B. Dahneke, editor. Wiley-Interscience, New York.

Eimer, W., and R. Pecora. 1991. Rotational and translational diffusion of short rodlike molecules in solution: Oligonucleotides. *J. Chem. Phys.* 94:2324–2329.

Eimer, W., J. R. Williamson, S. G. Boxer, and R. Pecora. 1990. Characterization of the overall and internal dynamics of short oligonucleotides by depolarized dynamic light scattering and NMR relaxation measurements. *Biochemistry.* 29:799–811.

Eisenhaber, F., and P. Argos. 1993. Improved strategy in analytic surface calculation for molecular systems: handling of singularities and computational efficiency. *J. Comp. Chem.* 14:1272–1280.

Eisenhaber, F., P. Lijnzaad, P. Argos, C. Sander, and M. Scharf. 1995. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comp. Chem.* 16:273–284.

Felderhof, B. U. 1996. Dynamics of colloidal suspensions. *Physica B.* 228:27–32.

Gallagher, W. H., and C. K. Woodward. 1989. The concentration dependence of the diffusion coefficient for bovine pancreatic trypsin inhibitor: a dynamic light scattering study of a small protein. *Biopolymers.* 28: 2001–2023.

Garcia de la Torre, J., and V. A. Bloomfield. 1981. Hydrodynamic properties of complex, rigid, biological macromolecules. Theory and applications. *Q. Rev. Biophys.* 14:81–139.

Garcia de la Torre, J., S. Navarro, and M. C. Lopez Martinez. 1994a. Hydrodynamic properties of a double-helical model for DNA. *Biophys. J.* 66:1573–1579.

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

Granier, T., B. Gallois, A. Dautant, B. Langlois D'Estaintot, and G. Precigoux. 1996. Preliminary results on x-ray diffraction study of the tetragonal form of native horse spleen ferritin. *Acta Crystallogr. D.* 52:594–596.

Harata, K. 1993. X-ray structure of monoclinic turkey egg lysozyme at 1.3 Å resolution. *Acta Crystallogr. D.* 49:497–504.

Hellwig, T., W. Eimer, E. Krahn, K. Schneider, and A. Müller. 1997. Hydrodynamic properties of nitrogenase: the MoFe protein from *Azotobacter vinelandii* studied by dynamic light scattering and hydrodynamic modeling. *BBA.* 1337:311–318.

Howlin, B., D. S. Moss, and G. W. Harris. 1989. Segmented anisotropic refinement of bovine ribonuclease A by the application of the rigid-body TLS Model. *Acta Crystallogr. A.* 45:851–860.

Hvidt, S. 1991. Insulin association in neutral solutions studied by light scattering. *Biophys. Chem.* 39:205–213.

Kirkwood, J. G. 1954. The general theory of irreversible processes in solution of macromolecules. *J. Polym. Sci.* 12:1–12.

Kuntz, I. D., and W. Kauzmann. 1974. Hydration of protein and polypeptides. *Adv. Protein Chem.* 28:236–345.

Liu, H., L. Skibinska, J. Gapinski, A. Patkowski, E. W. Fischer, and R. Pecora. 1998. Effect of electrostatic interactions on the structure and dynamics of a model polyelectrolyte. I. Diffusion. *J. Chem. Phys.* 109:7556–7566.

Marquart, M., J. Walter, J. Deisenhofer, W. Bode, and R. Huber. 1983. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallogr. B*.39: 480.

McCammon, J. A., J. M. Deutch, and B. U. Felderhof. 1975. Frictional properties of multisubunit structures. *Biopolymers.* 14:2613–2623.

Pastor, R. W., and M. Karplus. 1988. Parametrization of the friction constant for stochastic simulations of polymers. *J. Phys. Chem.* 92: 2636–2641.

Patkowski, A., W. Eimer, and Th. Dorfmüller. 1990a. A polarized and depolarized dynamic light scattering study of the $tRNA^{Phe}$ conformation in solution. *Biopolymers.* 30:93–105.

Patkowski, A., J. Seils, F. Biss, B. Jockusch, and Th. Dorfmüller. 1990b. Size and shape parameters of the actin-binding protein profilin in solution. A depolarized and polarized dynamic light scattering study. *Biopolymers.* 30:219–222.

Perrin, F. 1936. Mouvement brownien d'un ellipsoide. II. Rotation libre et depolarisation des fluorescences. Translation et diffusion de molecules ellipsoidales. *J. Phys. Radium.* 7:1–11.

Peters, J. W., M. H. Stowell, S. M. Soltis, M. G. Finnegan, M. K. Johnson, and D. C. Rees. 1997. Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry.* 36:1181–1187.

Rotne, J., and S. Prager. 1969. Variational treatment of hydrodynamic interaction in polymers. *J. Chem. Phys.* 50:4831–4837.

Spotorno, B., L. Piccinini, G. Tassara, C. Ruggiero, M. Nardini, F. Molina, and M. Rocco. 1997. BEAMS (BEAds Modelling System): a set of computer programs for the generation, the visualization and the computation of the hydrodynamic and conformational properties of bead models of proteins. *Eur. Biophys J.* 25:373–384.

Stryer, L. 1995. Biochemistry. W. H. Freeman and Co., New York.

Sussman, J. L., S. R. Holbrook, R. W. Warrant, G. M. Church, and S. H. Kim. 1978. Crystal structure of yeast phenylalanine tRNA. I. Crystallographic refinement. *J. Mol. Biol.* 123:607.

Takahara, P. M., A. C. Rosenzweig, C. A. Frederick, and S. J. Llipard. 1995. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature.* 377:649.

Teller, D. C., and C. de Haen. 1975. The sedimentation behavior of oligomeric proteins. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* 34:598.

Teller, D. C., E. Swanson, and C. de Haen. 1979. The translational friction coefficient of proteins. *Adv. Enzymol.* 61:103–124.

Venable, R. M., and R. W. Pastor. 1988. Frictional models for stochastic simulations of proteins. *Biopolymers.* 27:1001–1014.

Zhou, H. X. 1995a. Calculation of translational friction and intrinsic viscosity. 1. General formulation for arbitrarily shaped particles. *Biophys J.* 69:2286–2297.

Zhou, H. X. 1995b. Calculation of translational friction and intrinsic viscosity. 2. Application to globular proteins. *Biophys J.* 69:2298–2303.