Computation of the Dipole Moments of Proteins

Jan Antosiewicz
Department of Chemistry and Biochemistry, and Department of Pharmacology, University of California at San Diego, La Jolla, CA 92093-0365 USA

ABSTRACT A simple and computationally feasible procedure for the calculation of net charges and dipole moments of proteins at arbitrary pH and salt conditions is described. The method is intended to provide data that may be compared to the results of transient electric dichroism experiments on protein solutions. The procedure consists of three major steps: (i) calculation of self energies and interaction energies for ionizable groups in the protein by using the finite-difference Poisson-Boltzmann method, (ii) determination of the position of the center of diffusion (to which the calculated dipole moment refers) and the extinction coefficient tensor for the protein, and (iii) generation of the equilibrium distribution of protonation states of the protein by a Monte Carlo procedure, from which mean and root-mean-square dipole moments and optical anisotropies are calculated.

The procedure is applied to 12 proteins. It is shown that it gives hydrodynamic and electrical parameters for proteins in good agreement with experimental data.

INTRODUCTION

The investigation of dipole moments of proteins was initiated more than 50 years ago (see, e.g., Cohn and Edsall, 1943), and since then many experimental and theoretical studies have been reported (Kirkwood and Shumaker, 1952; Scheider, 1965; Orttung, 1968, 1969; Schlecht, 1969; Grant et al., 1978; Barlow and Thornton, 1986; Porschke, 1987; Antosiewicz and Porschke, 1989a, 1995; Takashima and Asami, 1992, 1993; Takashima, 1993).

Comparison of dipole moments calculated on the basis of assumed models of proteins with experimental results provides a good test of theoretical approaches and consequently has a significant influence on many areas of molecular biology. Dipole moments are also important for proper interpretation of the results of electrooptical relaxation experiments in terms of macromolecular structures (O'Konsky et al., 1959; Porschke, 1987; Porschke and Antosiewicz, 1990).

A protein molecule immersed in a solution may be considered as a set of electrical charges in a certain spatial arrangement, embedded in low dielectric material. Because of the exchange of free ionic species with the environment, the charge configuration and hence the dipole moment vector of the protein fluctuate. Both the spatial distribution of charges and its fluctuations contribute to the permanent dipole moment of the molecule (Scheider, 1965; Orttung, 1968; Antosiewicz and Porschke, 1993). The possible influence of an external electric field on the charge distribution within a protein is referred to as the polarizability of the molecule and contributes to the induced dipole moment. It seems that the polarizability of proteins is usually negligible (Porschke, 1987; Porschke et al., 1988; Antosiewicz and Porschke 1989a; Takashima and Asami, 1993).

Dipole moments of proteins can be determined experimentally by means of measurements of dielectric constants of their solutions as a function of the frequency of the applied electrical field (frequency domain method) (Oncley, 1943; Takashima and Asami, 1993) or from electrooptical transients of their solutions after application of a pulse of external electric field (time domain method) (Fredericq and Houssier, 1973; Porschke, 1987).

In frequency domain techniques, the dipole moment is derived from the dispersion of the dielectric constant of the solution when the frequency of applied electrical fields approaches the intrinsic frequency of orientation of polyelectrolyte particles. The difference between low- and high-frequency limiting dielectric increments, ΔD_0 and ΔD_∞, is connected to the dipole moment m of the particles through the equation (Oncley, 1943)

$$m^2 = \frac{9000\, kTM}{4\pi Nh}\left(\frac{\Delta D_0}{g} - \frac{\Delta D_\infty}{g}\right) \qquad (1)$$

where M is the molecular weight of the protein, g is the concentration of protein in grams per liter of solution, N is Avogadro's number, kT is the Boltzmann constant times the absolute temperature, and h is an empirical parameter that might be considered a correction factor for internal fields (Onsager, 1936). This parameter was calibrated to be 5.8 based on the measurements with simple amino acids whose dipole moments are well established (Oncley, 1943; Takashima and Asami, 1993). Because of the empirical nature of h, the dipole moment of the molecule, calculated according to Eq. 1, may be considered a good estimate of the dipole moment in solution. It is in general somewhat greater than the moment in the gas phase (Kirkwood, 1939).

Received for publication 15 May 1995 and in final form 22 June 1995.

Address reprint requests to Dr. Jan Antosiewicz, Department of Chemistry and Biochemistry, University of California, 9500 Gilman Dr., La Jolla, CA 92093-0365. Tel.: 619-534-3038; Fax: 619-534-7042; E-mail: jantosie@chemcca9.ucsd.edu.

Dr. Antosiewicz is on leave from the Department of Biophysics, University of Warsaw, Warsaw, Poland.

0006-3495/95/10/1344/11 $2.00

In time domain techniques, the transient electric dichroism or electric birefringence of solutions exposed to electric field pulses is measured. The first is easier to interpret in terms of molecular parameters, and our discussion is therefore limited to this method. Linear electric dichroism refers to the differential absorption of plane-polarized light by a sample containing optically anisotropic particles oriented by an external electric field (Norden, 1978). The results of extinction anisotropy measurements are usually represented in the form of the reduced linear dichroism:

$$\xi(t) = \frac{A_{\parallel}(t) - A_{\perp}(t)}{\bar{A}} \tag{2}$$

where $A_{\parallel}$ and $A_{\perp}$ are the absorbances of light polarized parallel and perpendicular to the field vector, respectively; $\bar{A}$ is the isotropic absorbance.

Linear dichroism of the solution at saturation ξ_{sat} depends on the electric and optical properties of the molecules under investigation (O'Konski, et al., 1959; Holcomb and Tinoco, Jr. 1963; Diekmann et al., 1982; Porschke, 1985). For pure permanent moment, the dichroism at saturation reads

$$\xi_{sat}(E) = \xi_{\infty} \cdot \left(1 - \frac{3(\coth\beta - 1/\beta)}{\beta}\right) \tag{3}$$

where ξ_{∞} is the reduced dichroism at infinitely large field, $\beta = mE/kT$, E is the external electric field strength, and m is the permanent dipole moment. In addition to electric field dependence of saturation dichroism, proper assessment of the dipole moment also requires the analysis of time constants of electrooptical rise curves (Porschke, 1987).

The dipole moment obtained from electro-optical experiments through the use of Eq. 3 also requires correction for internal field effects. This field is a sum of the cavity field and the reaction field (Onsager, 1936). The reaction field is due to polarization of the surrounding solvent by the molecule. This field polarizes the molecule, thus increasing its dipole moment. Because this field is directed along the dipole moment of the molecule, it does not participate in forming an orienting electric torque. The torque is entirely due to the cavity field. For spherical molecules, the cavity field E_i reads (Onsager, 1936; Fröhlich, 1958; Böttcher, 1973)

$$E_i = \frac{3D_1}{2D_1 + 1} E_e \tag{4}$$

where E_e is the applied external electric field strength and D_1 is the dielectric constant of the solvent. Thus for water solvent at 293 K and globular proteins the dipole moments resulting from electrooptical experiments should be divided by 1.5. If we consider a protein as a set of charges immersed in a dielectric with dielectric constant D_2, then the dipole moment m calculated solely on the basis of the charge distribution is related to the internal dipole moment m_i obtained from electrooptical measurements (with the directing field correction expressed by Eq. 4) by Eq. (Fröhlich, 1958) 5:

$$m_i = \frac{2D_1 + 1}{2D_1 + D_2} m \tag{5}$$

Thus we may see that the dipole moment corrected for the directing field effect is what it should be compared to computed dipole moments, as in aqueous solutions D_1 is of the order of 80, whereas D_2 is close to 4.

It was shown (Scheider, 1965; Antosiewicz and Porschke, 1993) that, depending on the rate of proton fluctuations, dielectric relaxation experiments and electrooptical experiments result in mean to root-mean-square dipole moments. Electro-optical experiments also give information about the orientation of the resulting dipole in the particle-fixed coordinate system.

Because dipole moments are directly obtainable from experiment, they provide good tests for theoretical methods. The advent of the Poisson-Boltzmann method for computing detailed electrostatic fields in and around macromolecules (Warwicker and Watson, 1982; Klapper et al., 1986; Gilson et al., 1988; Davis et al., 1991) and developments in molecular hydrodynamics of polymer solutions (Garcia de la Torre and Bloomfield, 1981) and in understanding of the interaction of macromolecules with electromagnetic radiation (Norden, 1978) make it possible to build a procedure for the calculation of dipole moments of proteins at arbitrary conditions, which may be directly compared with results of dielectric and electrooptical experiments.

In the present work we describe such a procedure and its application to calculating the dipole moments of 12 proteins, for most of which experimental data exist. The procedure is built from the following parts:

1. The finite-difference Poisson-Boltzmann (FDPB) method is used to calculate the self energies and interaction energies of the ionizable groups in the protein (Antosiewicz et al., 1994).
2. Translational and rotational diffusional tensors (Garcia de la Torre and Bloomfield, 1981), the center of diffusion (Harvey and Garcia de la Torre, 1980), and the extinction coefficients tensor (Norden, 1978; Antosiewicz and Porschke, 1989a) of the protein are calculated.
3. The energies are used in a Monte Carlo procedure (Metropolis et al., 1953; Antosiewicz and Porschke, 1989a) to generate an equilibrium distribution of protonation states from which electro-optical parameters are calculated.

MATERIALS AND METHODS

Calculation of dipole moment

The dipole moment, **m**, of a charge distribution is defined (see, e.g., Böttcher, 1973) as

$$\mathbf{m} = \sum_j q_j \mathbf{r}_j \tag{6}$$

where $\mathbf{r}_j$ is the position vector of charge q_j in a coordinate system fixed on the particle. Atomic coordinates for many proteins are available from Protein Data Bank (Bernstein et al., 1977). Calculation of a protein's dipole moment thus requires partial charges for all of its atoms. The dipole of a protein, in a certain protonation state, may be represented as a sum of two contributions: the first due to charges of all amino acids in neutral state, and the second due to formal charges of ionizable groups:

$$\mathbf{m} = \sum_{j=1}^{M} x_j \gamma_j \mathbf{r}_j + \mathbf{m}_{\text{neutral}} \tag{7}$$

where the summation over j counts only atoms chosen as ionizable sites (see below). The ionization variable x_j is equal to 0 for a neutral group and 1 for an ionized group. γ_j labels type of site, equal to -1 for anionic sites and $+1$ for cationic sites. $\mathbf{m}_{\text{neutral}}$ can be calculated either from partial charges for all amino acids or more approximately by summing the dipole moments of all of the protein's peptide bonds. Each peptide bond may be considered as a dipole of 3.5 Debye units directed from C atom to O atom of the peptide group (Hol, 1985). Usually $\mathbf{m}_{\text{neutral}}$ is small (Barlow and Thornton, 1986). However, α-helical fragments of proteins are known to have a substantial dipole moment. $\mathbf{m}_{\text{neutral}}$ is origin independent, but the dipole moment due to charges of ionizable groups is usually dependent on the choice of origin. We choose as the origin of the coordinate system the center of diffusion of the protein (Harvey and Garcia de la Torre, 1980).

Energy of a protein in a given protonation state

The first step in the procedure consists of evaluation of electrostatic energies of the protein in different protonation states. The energy of each state is composed of the sum of energy contributions characterizing the protonation of a single site with all the remaining sites neutral and interaction energies between the sites. We use a simple model of ionization that consists of adding a single positive or negative charge to a certain atom of the titratable amino acid, chosen as the titratable site. It is assumed that the difference in the titration behavior of an ionizable site in a protein and in a model amino acid compound can be accounted for by calculating the difference in the electrostatic work of altering the site charge from the unprotonated to the protonated state in the protein and the work of making the same alteration in the model compound.

The total electrostatic energy difference between the ionized and neutral states of an ionizable group in the model compound or the protein can be written in the form

$$\Delta G^{(\text{el})} = \frac{1}{2}\sum_i \sum_j q_i' \phi_{ji}' - \frac{1}{2}\sum_i \sum_j q_i \phi_{ji} \tag{8}$$

where q_i is the partial charge on the ith atom of the model compound or the protein, ϕ_{ji} is the potential at the location of charge i due to charge j, symbols without a prime refer to the neutral state, symbols with a prime refer to the ionized state, and the summation is taken over all atoms of the model compound or the whole protein. For convenience, the atom corresponding to a given ionization site is numbered 1 both in the model compound and in the protein. Notice that in this case, for $i, j > 1$ we have $q_i' = q_i$ and $\phi_{ji} = \phi_{ji}$. The electrostatic energy of ionization of the site in the protein relative to that in the model compound then reads

$$\Delta\Delta G_1 = \frac{1}{2}(q_1' - q_1)^2(\Psi_{11} - \Phi_{11}) + (q_1' - q_1) \cdot \sum_{j=1}^{N} q_i \Psi_{1j} - (q_1' - q_1)\sum_{j=1}^{n} q_i \Phi_{1j} \tag{9}$$

where Ψ_{1j} is the potential at charge j due to a unit positive charge at site 1 in the protein, Φ is the equivalent quantity in the model compound, N is the number of atoms in the protein, and n is the number of atoms in the model compound. Notice that $(q_1' - q_1)$ is -1 or $+1$, depending on the type of the site. Similar equations hold for any other atom chosen as the ionizable site. The electrostatic energy difference $\Delta\Delta G_i$ is assumed to alter the pK_a of the ionizable group in the protein with the remaining sites neutral, relative to the pK_a in the corresponding model compound. This assumption leads to the concept of intrinsic pK_a defined as (Tanford and Kirkwood, 1957)

$$pK_{a_{\text{intrinsic,i}}} = pK_{a_{\text{model,i}}} - \gamma_i \Delta\Delta G_i / 2.303\, RT \tag{10}$$

where $\gamma_i \equiv (q_i' - q_i)$ is -1 for an acidic site and $+1$ for a basic site. However, $\Delta\Delta G_i$ is not the only electrostatic factor influencing the pK_a of a titratable site in a protein. The second important factor is the electrostatic interaction between the titrable sites. For example, the electrostatic energy of the protein with two sites, i and j, ionized is

$$\Delta G_{i,j} = \Delta\Delta G_i + \Delta\Delta G_j + \gamma_i \gamma_j \Psi_{ij} \tag{11}$$

where Ψ_{ij} is the interaction energy between unit positive charges located at the positions of titratable sites i and j. This equation may be used to calculate the energy of the protein in an arbitrary protonation state. First notice that when we consider the protonation of a model amino acid compound with an experimentally known $pK_{a_{\text{model}}}$,

$$\Delta G = 2.303\, RT(\text{pH} - pK_{a_{\text{model}}}) \tag{12}$$

This is the free energy of protonation, relative to the deprotonated state (Poland, 1978). For the site in the protein environment with all other sites neutral the above equation changes to

$$\Delta G = 2.303\, RT(\text{pH} - pK_{a_{\text{intrinsic}}}) \tag{13}$$

If instead of the neutral protein, the deprotonated protein is chosen as the reference state, the interaction term in Eq. 11 must be modified

$$\Delta G_{i,j,\text{interaction}} = (q_i q_j - q_i^{\circ} q_j^{\circ})\Psi_{ij} \tag{14}$$

Now charges q_i and q_i° are idealized site charges in a given state and in the deprotonated state, respectively. They are idealized in the sense that for an acidic site they simply have $q_i^{\circ} = -1$ and in the protonated state $q_i = 0$, and for a basic site $q_i^{\circ} = 0$ and in the protonated state $q_i = +1$. The above results may be generalized to the case of a protein with many titratable sites and any protonation state:

$$\Delta G(\text{pH}, x_1, \ldots, x_M) = 2.303\, RT \sum_{i=1}^{M} x_i(\text{pH} - pK_{a_{\text{intrinsic,i}}}) + \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} (q_i q_j - q_i^{\circ} q_j^{\circ})\Psi_{ij} \tag{15}$$

where parameter x_i is 1 when the site is protonated; otherwise it is 0. This parameter is implicitly present also in the second term on the right side of Eq. 15 through charge q_i because for $x_i = 0$ $q_i = q_i^{\circ}$. Notice that now the indices i and j refer to titratable sites, and not to all of the partial charges of the protein.

For each ionizable group i, a unit charge is placed at the protonation site, charges of all other atoms are set to zero, and two sets of finite difference Poisson-Boltzmann calculations are performed: one with a single unit charge on the ionization site of the model compound in isolation, and a second with a single unit charge on the same site in the neutral protein. This is sufficient to get all data required for Eq. 15. The results of these calculations are organized in a file described under Monte Carlo Procedure,

below. Details of this methodology are presented elsewhere (Antosiewicz et al., 1994).

Calculation of the extinction coefficient tensor of a protein

Electromagnetic radiation is able to produce a transition between two molecular energy states provided that there exists a molecular electric or magnetic moment with which the radiation field can interact (Norden, 1978). In the case of a small molecule the absorption intensity is proportional to the square of the scalar product between the electric field vector of the light and the above-mentioned moment vector. The latter vector is considered to be attached to the molecular chromophore and is called the transition moment. Light polarized parallel to the transition moment has the maximum probability of absorption in the region of the spectral activity of the molecule, and conversely, if the polarization of the light is perpendicular to the transition moment no absorption can take place (Norden, 1978). These properties of a single chromophore may be mathematically described by using a diagonal extinction tensor of the form

$$\hat{\epsilon}_{\text{local}} = \begin{pmatrix} 0. & 0. & 0. \\ 0. & 0. & 0. \\ 0. & 0. & 3\bar{\epsilon} \end{pmatrix} \tag{16}$$

where $\bar{\epsilon}$ is the measured extinction coefficient of the molecule at a given wavelength. The form of the tensor corresponds to a local coordinate system with the z-axis along the transition moment vector. The direction of the latter is known in advance from, for example, experimental studies of the isolated chromophores oriented in films or crystals, or estimated with confidence using quantum mechanical calculations. Although the absorbance of chromophores depends on their environment (Rizzo and Schellman, 1984), as a first approximation it may be assumed that the extinction tensor of the whole protein can be determined as a tensorial sum of all of the isolated chromophore's contributions giving a molecular extinction tensor, say $\hat{\epsilon}$. Given $\hat{\epsilon}$ and the dipole moment vector $\mathbf{m}$, the reduced limiting dichroism at infinite electric field may be easily evaluated. If the incident light is directed along the x'-axis and the uniform electric field is directed along the z'-axis of the laboratory coordinate system, $A_\parallel$ and $A_\perp$ from Eq. 2 may be expressed by

$$A_\parallel = cl\langle\epsilon_{z'z'}\rangle \quad \text{and} \quad A_\perp = cl\left\langle\frac{\epsilon_{x'x'} + \epsilon_{y'y'}}{2}\right\rangle. \tag{17}$$

Thus finally, the reduced dichroism reads

$$\xi = \frac{3\langle\epsilon_{z'z'}\rangle - \text{tr}\,\hat{\epsilon}}{2\bar{\epsilon}} \tag{18}$$

where tr $\hat{\epsilon}$ is the sum of diagonal elements of the tensor (which is invariant under rotations), c is the molar concentration of particles, and l is the optical path length. In the above equations $\langle\ \rangle$ means average value over all particles present in the solution.

At infinite electric field strength, the dipole moment vector of the particle is aligned along the direction of the external electric field. Thus the reduced limiting dichroism is calculated from Eq. 18 with

$$\epsilon_{z'z',\infty} = \mathbf{e}^T\hat{\epsilon}\mathbf{e} \tag{19}$$

where $\mathbf{e} = \mathbf{m}/m$ is a unit vector in the direction of the dipole moment of the particle in the coordinate system fixed to the particle, and superscript T indicates a vector written as a row.

For absorption of UV radiation by proteins at around 280 nm, transitions in tryptophan and tyrosine are significant (Cantor and Timasheff, 1982). The directions of transition moments and extinction coefficients were taken from papers by Yamamoto and Tanaka (1972), Philips and Levy (1986), Maki et al. (1978), and Fasman (1976) and are summarized in Table 1.

TABLE 1 Atoms approximating orientation of transition moments and corresponding extinction coefficients $\bar{\epsilon}$, in units of M^{-1} cm^{-1} × 10^{-3}, in the direction of the transition moment for tryptophan and tyrosine at 280 nm

Amino acid	Direction	$\bar{\epsilon}$
Tryptophan	NE → CE3	4.8
Tryptophan	CG → CZ2	0.8
Tyrosine	CD1 → CD2	1.2

Names of atoms as in Protein Data Bank files (Bernstein et al., 1977).

Calculation of the center of diffusion

Each macromolecule immersed in a solution undergoes random translational and rotational motions because of the constant bombardment by solvent molecules. In very dilute solutions collisions between macromolecules are rare, and we neglect them in the following discussion. These random motions are characterized by the diffusion tensor of the molecule (Einstein, 1905; Brenner, 1965),

$$\hat{D} = \begin{pmatrix} \Xi_{t,O} & \Xi_{c,O}^T \\ \Xi_{c,O} & \Xi_r \end{pmatrix} \tag{20}$$

where $\Xi_{t,O}$ is the translational diffusion tensor, Ξ_r is the rotational diffusion tensor, and $\Xi_{c,O}$ is the translational-rotational coupling diffusion tensor. Subscript O indicates tensors depending on the location of the origin of the particle coordinate system, and superscript T refers to the transposed tensor. Translational and rotational tensors are always symmetric, but the coupling tensor is symmetric only when the origin of the coordinate system is located at the center of diffusion (CD) (Harvey and Garcia de la Torre, 1980). When the components of the diffusion tensors are known in an arbitrary coordinate system, the center of diffusion has coordinates (Harvey and Garcia de la Torre, 1980):

$$\begin{aligned} x_{CD} &= \frac{\Xi_{c,23} - \Xi_{c,32}}{\Xi_{r,22} + \Xi_{r,33}} \\ y_{CD} &= \frac{\Xi_{c,31} - \Xi_{c,13}}{\Xi_{r,11} + \Xi_{r,33}} \\ z_{CD} &= \frac{\Xi_{c,12} - \Xi_{c,21}}{\Xi_{r,11} + \Xi_{r,22}} \end{aligned} \tag{21}$$

The diffusional tensors of proteins can be calculated by means of a bead model approach, applicable to particles of arbitrary shape (Garcia de la Torre and Bloomfield, 1981; Garcia de la Torre and Rodes, 1983; Antosiewicz and Porschke, 1989a,b). Because the program used for these calculations was limited to a maximal number of 240 beads, the number of amino acids represented by a single bead depends on the total number of residues. The center of each bead was located at the center of coordinates of the nonhydrogen atoms of the set of amino acids. Many of these beads overlap. Since the hydrodynamic theory for overlapping beads has been developed only for beads of equal radii (Garcia de la Torre and Bloomfield, 1981), all beads were assigned the same radius. This radius was defined as the sum of the radius of a water molecule (1.4 Å) and the mean maximal distance of any atom from the center of its bead. This procedure has been shown to yield diffusion coefficients of biopolymers that are in reasonably good agreement with experimental values (Antosiewicz and Porschke, 1989b), and this is further confirmed by the present work.

Monte Carlo procedure and calculation of the electro-optical parameters

The Monte Carlo program used in the present study is a modification of a program described elsewhere (Antosiewicz and Porschke, 1989a). The old

version used self energies and interaction energies for ionizable sites determined by the Tanford-Kirkwood method (1957). The present version uses self energies and interaction energies for such sites calculated according to the treatment presented above (see Energy of a Protein in a Given Protonation State). The input data used by this program are the same as those used by the program Hybrid written by Gilson (1993) for very fast evaluation of mean charges, pK_a values, and electrostatic energies for proteins. The Monte Carlo procedure is much slower but provides information about mean dipole moments, root-mean-square dipole moments, mean charges, pK_a values of ionizable groups, limiting reduced dichroism for mean dipole moments, and mean limiting dichroism for all accepted protonation states, and lists a certain number of protonation states with the lowest energies found during the Monte Carlo search. The input data are organized as follows:

$$M$$

$$pK_{1,\text{model}}\ \gamma_1\ \Delta\Delta G_1\ 1$$

$$\Psi_{1,2}$$

$$\vdots$$

$$\Psi_{1,M}$$

$$pK_{2,\text{model}}\ \gamma_2\ \Delta\Delta G_2\ 2$$

$$\Psi_{2,3}$$

$$\vdots$$

$$\Psi_{2,M}$$

$$\vdots$$

$$pK_{M,\text{model}}\ \gamma_M\ \Delta\Delta G_M\ M$$

The first line contains the number of ionizable sites. The second line contains the model compound pK_a of site 1, γ_1, $\Delta\Delta G_1$ (see Eq. 9), and the site index. The next $M - 1$ lines contain the Ψ_{1j} interaction terms (see Eq. 11) for $j = 2, \ldots, M$. Blocks of analogous data for the subsequent sites 2, . . ., M follow. These quantities are used by our Monte Carlo program for generating the equilibrium distribution of protonation states of the protein for a given temperature, salt concentration, and pH.

The program uses the Metropolis algorithm for sampling protonation states (Metropolis et al., 1953). The equilibrium distribution is generated based on the free energy of states. For the generation of the initial distribution we use the following rule: if the intrinsic pK_a of an ionizable site is lower than the pH then the site is assumed to be deprotonated; otherwise it is protonated. Having defined the initial state of the protein, we calculate its Gibbs free energy ΔG_1 according to Eq. 15. Now to sample a second state we use a uniform {0–1} random number generator for deciding wheather an ionizable group is altered. If the kth random number is larger than a certain parameter s (taken to be between 0 and 1), then the kth group changes its state of protonation; otherwise it remains unchanged. With such a procedure an average $(1 - s)$ fraction of sites of the protein in each Monte Carlo step change their protonation state. Subsequently we calculate ΔG_2. If $\Delta G_2 \leq \Delta G_1$ then the new state 2 is accepted, but if $\Delta G_2 > \Delta G_1$ then state 2 is accepted with probability $\exp[(\Delta G_1 - \Delta G_2)/RT]$. Accepting state 2 means that this state will be taken as the reference state in the next Monte Carlo step. If state 2 is rejected it means that state 1 is kept as the reference state. In this way we subsequently sample and consider further states of the protein. Because the first sampled state may be far from equilibrium, collecting of states for the equilibrium distribution only starts after a large number of Monte Carlo steps have been done. The value of the parameter s is adjusted so that new states are accepted with approximately 50% probability.

Computational methods and parameters

Atomic radii were set to 0.5σ, where the σ values are those of the optimized potentials for liquid simulations (OPLS) nonbonded parameter set (Jorgensen and Tirado-Rives, 1988). Because the OPLS parameters do not include charges for the neutral forms of each ionizable residue, these were taken from CHARMm Version 22.0 (Brooks et al., 1982, Molecular Simulations, Inc.) (polar hydrogen-only parameter set). Charges for metal ions (present in some investigated proteins) were +2 for Fe, Zn, Ca, and Mn. These charges were included in the Poisson-Boltzmann calculations and contribute to the estimated isoelectric points and dipole moments. Polar hydrogen atoms were added to Protein Data Bank files by means of the CHARMm HBUILD command (Brunger and Karplus, 1988). Ionization was represented as the addition of a ± 1 proton charge to a single atom in each group: the C atom of the main chain C-terminus; the N atom of the main chain N-terminus; CG of Asp; CD of Glu; CZ of Arg; NZ of Lys; ND1 or NE2 of His; OH of Tyr; and SG of Cys. In all cases except ribonuclease A, the neutral form of histidine has the proton on ND1, and NE2 is the protonation site. For ribonuclease A the best model from a previous study (Antosiewicz et al., 1994) was used: in this case the neutral form has the proton on NE2 and ND1 is the protonation site.

The following initial pK_a values were used: C-terminus 3.8; N-terminus 7.5; Asp 4.0; Glu 4.4; Arg 12.0; Lys 10.4; His 6.3; Tyr 9.6; Cys 8.3 (Nozaki and Tanford, 1967; Stryer, 1981); and heme propionic acid 4.0 (Matthew et al., 1979).

All calculations used a temperature of 293 K, 10 or 150 mM ionic strength of monovalent electrolyte with a 2.0 ÅA Stern layer, and a solvent dielectric constant of 80. A protein dielectric constant of 15 was used. It was shown previously that values of the order of 15–20 give much better agreement between computed and experimental pK_a values of titratable groups than the more realistic value of 4 (Antosiewicz et al., 1994).

All finite-difference calculations were carried out using the program UHBD (Davis et al., 1991). The Richards probe-accessible surface definition (Richards, 1977) of the dielectric surface was used (Gilson et al., 1988). The probe sphere radius was 1.4 ÅA, and each atom-sphere was assigned a starting set of 300 surface dots. Dielectric boundary "smoothing" was used as described elsewhere (Gilson et al., 1993; Davis and McCammon, 1991).

Monte Carlo calculations were done using 100,000 sampling steps, with 10,000 sampling steps used for equilibration.

Sources of experimental data

Experimental values for the molar absorption coefficients of proteins investigated in the present work were taken from the *Handbook of Biochemistry and Molecular Biology* (Fasman, 1976). Data on translational diffusion coefficients were taken from the *Handbook of Biochemistry* (Sober, 1968), *Proteins* (Creighton, 1993), *Biophysical Chemistry* (Cantor and Schimmel, 1980), and Gaigalas et al. (1992). Experimental values of isoelectric points were taken from works by Malamud and Drysdale (1978), Hames and Rickwood (1981), Tanford and Roxby (1972), and Gorbunoff (1984). Experimental dipole moments were taken from the works of Takashima and Asami (1992, 1993), Schlecht et al. (1969), Grant et al. (1978), Keefe and Grant (1974), South and Grant (1972), and Antosiewicz and Porschke (1989a).

Experimentally determined values of diffusion coefficients, extinction coefficients, and isoelectric points are listed in Table 2. It should be noted that isoelectric points of proteins are relatively difficult to measure and reported values could be inaccurate. Moreover, the same protein from different species may have different isoelectric points; thus it is important that calculations and experiments are done for the same protein from the same species. These species are not specified in the work of Takashima and Asami (1993), our major source of the experimental dipole moments. However, discussion with Prof. Takashima established that only for the cases of phospholipase A2, carboxypeptidase A, and concanavalin A were isoelectric points and dipole moments obtained experimentally for species different from those used in the present calculations.

TABLE 2 Experimental and computed values of translational diffusion coefficients,* $D_{20,w}$ (in units of $cm^2s^{-1} \times 10^{-7}$), molar absorption coefficients at 280 nm,‡ $\bar{\epsilon}$ (in units of $M^{-1}\ cm^{-1} \times 10^{-4}$), and isoelectric points,§ pI (in units of pH)

	$D_{20,w}$		$\bar{\epsilon}$		pI	
Protein	Expl	Comp	Expl	Comp	Expl	Comp
Lysozyme	11.2	11.7	3.60–3.90	3.7	11.2	11.10
Lysozyme	11.2	11.7	3.60–3.90	3.7	11.2	11.05
Subtilisin	9.04	9.27	3.10–3.20	2.9		6.60
Phospholipase A2	12.6	11.5	2.06	1.4	7.4‖	8.20
Carboxypept. A	8.68	8.45	5.90	6.2		8.50
Concanavalin A	5.43–5.60	9.46	7.75–7.98	3.1	4.5–5.5	6.45
Ribonuclease A	10.2–13.1	11.6	0.88–1.14	0.7	9.3	9.65
Alcohol deh. mon.	5.96–6.50	7.79	3.44–3.83	1.6	8.7–9.3	9.80
Alcohol deh. dim.	5.96–6.50	5.82	3.44–3.83	3.2	8.7–9.3	9.80
Myoglobin oxy	11.3	10.8	1.54**	1.5	7.7–8.1	7.50
Myoglobin deoxy	11.3	10.9	1.54**	1.5	7.7–8.1	7.35
Chymotrypsin	10.20	9.71	4.46–5.00	5.0	8.8	9.10
Chymotrypsinogen	9.50–10.2	9.30	5.15	5.0	9.2–9.6	9.75

* Sober, 1968; Creighton, 1993.
‡ Fasman, 1976.
§ Malamud and Drysdale, 1978; Hames and Rickwood, 1981.
‖ The enzyme from rat spleen.
** The value for apo-myoglobin.

Protein structures

All calculations were based upon crystallographic coordinate sets. The Protein Data Bank (Bernstein et al., 1977) provided coordinates for triclinic hen egg white lysozyme (2LZT) (Ramanadham et al., 1981), tetragonal hen egg white lysozyme (1LYZ) (Diamond, 1974), *Bacilus amyloliquefaciens* subtilisin (2SBT) (Drenth et al., 1972), cow α-chymotrypsin (4CHA) (Tsukada and Blow, 1985), bovine pancreas chymotrypsinogen A (2CGA) (Wang et al., 1985), bovine pancreas phospholipase A2 (1BP2) (Dijkstra et al., 1981), bovine pancreas carboxypeptidase A (5CPA) (Rees et al., 1983), Jack bean concanavalin A (3CNA) (Hardman and Ainsworth, 1972), bovine pancreas ribonuclease A (3RN3) (Howlin et al., 1989), horse liver alcohol dehydrogenase (8ADH) (Colonna-Cesari et al., 1986), sperm whale oxymyoglobin (1MBO) (Phillips and Schoenborn, 1981), and sperm whale deoxymyoglobin (5MBN) (Takano, 1984).

Unobserved heavy atoms in the α-chymotrypsin structure were built graphically with the program INSIGHT (Biosym Technologies, 1992), as described elsewhere (Antosiewicz et al., 1994).

RESULTS AND DISCUSSION

The rotational diffusion of macromolecules occurs around their centers of diffusion, the degree of orientation under the influence of the external electric field pulse is determined by their dipole moments relative to this point, and the optical anisotropy is determined by the extinction coefficient tensor. Thus, any calculation of the dipole moment and optical anisotropy of the solution of investigated macromolecules should be preceded by calculation of their diffusional properties and extinction coefficient tensor. Moreover, as measurements frequently are done at isoelectric points, they should also be calculated. The following subsection briefly discusses these points. Following this, the results of the calculations of dipole moments are presented.

Diffusion coefficients, extinction coefficients, and isoelectric points

The computed values of translational diffusion coefficients, average extinction coefficients at 280 nm, and isoelectric points are listed in Table 2, together with the values obtained from experiment. Comparison of the experimental and calculated mean translational diffusion coefficients shows that the relatively simple computational method used in the present work is reliable. Deviations of the theoretical values from experimental values are below 10% for all but one case. The exception is concanavalin A. Although no indication of the formation of dimers or higher aggregates by this protein was indicated in the pdb structural file, it seems probable that the aggregation may be influencing the experimental results. For alcohol dehydrogenase we see that the calculated diffusion coefficient for the monomer of the protein is much larger than the experimental value. However, in the pdb file there is an indication that this protein exists in solution in the form of a dimer (Colonna-Cesari et al., 1986). The computed diffusion coefficient for the dimer of alcohol dehydrogenase is in excellent agreement with the experimentally determined value. It is worth mentioning that the accuracy of our simple approach is comparable to that of the more rigorous theoretical computation of diffusion coefficients presented by Brune and Kim (1993).

The good agreement between the calculated and theoretical diffusion coefficients suggests that the calculated positions of the centers of diffusion are reliable. The distances from the center of diffusion (CD) to the center of mass (CM) for each protein were also computed. It appears (data not shown) that for the proteins investigated in the present study, and probably for all globular proteins, deviations of the CD from the CM are

small. The largest noted value was 1.5 ÅA (for concanavalin A, ribonuclease A, and alcohol dehydrogenase monomer), but for the remaining proteins, these distances were below 1 ÅA. However, it may be worth recalling that a charge of 1 unit of elementary charge gives a contribution of 5 Debye units when the reference point is shifted by 1 ÅA.

The computed values of the molar extinction coefficients at 280 nm are also in good agreement with the experimental values. Again there is a discrepancy for concanavalin A, but this is consistent with the discrepancy observed for the values of the diffusion coefficient. For alcohol dehydrogenase the extinction coefficient calculated for the dimeric form of the enzyme is again in very good agreement with the experimental value. In the case of myoglobin, because the heme group is expected to influence the absorption at 280 nm, the computed value was compared to that of the apoprotein. In summary, we may conclude that our relatively simple model in which the effects of the environment on the chromophore were neglected is quite reliable. So we may expect that optical anisotropy calculations for the proteins will be of similar accuracy.

The experimental dipole moments reported by Takashima and Asami (1992, 1993) were obtained at the isoelectric points of the proteins. Comparison between the calculated and experimental isoelectric points shows reasonably good agreement (see Table 2). Again, a significant difference is obtained for concanavalin A. This is the only protein for which none of the computed results agree with the experimental data. Discrepancies between the experimental and calculated isoelectric points for carboxypeptidase A and phospholipase A2 result from the fact that they refer to the same protein but from different species.

Dipole moments due to partial charges and peptide bonds

The dipole moment of a protein due only to the contribution of peptide bonds may serve as an estimate of the dipole moment with all amino acids in their neutral state. For comparison the neutral form dipole moment may also be computed on the basis of the partial charges of the atoms of all amino acids in the neutral state. However, it seems useful to keep the first method of estimation of the neutral form dipole moment, because of its natural reference to the α-helical fragments of the protein, which are known to give the largest contribution to the neutral form dipole (Hol, 1985). Table 3 shows the dipole moments calculated by the above two methods, the angles between the computed dipole moments, the charges of the protein resulting from the null model at the isoelectric points estimated by Poisson-Boltzmann methods, and the dipole moments due to ionizable sites and metal ions if present. In the case of myoglobin the charge of the Fe was included as part of the heme group, thus contributing to the neutral form dipole.

The dipole moments of the neutral form of the proteins presented in Table 3 constitute between 10% and 80% of the value due to the ionizable sites. They are therefore not always negligible. Moreover, the difference between the two methods of calculating the neutral form dipole is sometimes quite significant. This is interesting, and it suggests that for more accurate analyses of the dipole moments of proteins, the partial charges of all atoms should be considered in calculating the dipole moments of the neutral forms. This, however, requires additional analysis of their dependence on different data sets available for these partial charges and so is not discussed further in the present work. Taking into account the fact that the partial charges of atoms are not known unambigously, we may consider these differences as giving some estimate of the uncertainties in the results of the calculations. It is also interesting that in some cases the angle between the computed dipole moments of the neutral form is large. This is important for the interpretation of electro-optical measurement, which provides an estimate of the direction of molecular dipole moment. Fortunately, for all cases shown in Table 3, a large neutral-form

TABLE 3 Calculated values of the dipole moments of the proteins due only to peptide bonds, m_{pb}, the dipole moments of the neutral forms of the proteins due to the partial charges, m_{pc}, the angle between these dipole moments, $\alpha_{pb,pc}$, the charge of the proteins according to the null model at the isoelectric points calculated according to the Poisson-Boltzmann model, $q_{nm,pI}$, and the mean dipole moments due to ionizable sites at pI plus the contribution of coordinated metal ions, if applicable, according to the Poisson-Boltzmann model, $m_{is,ac}$

Protein	m_{pb}	m_{pc}	$\alpha_{pb,pc}$	$q_{nm,pI}$	$m_{is,ac}$
Lysozyme triclinic	51	76	38.3	−2.2	97
Lysozyme tetragonal	50	65	2.8	−1.9	72
Subtilisin	217	263	6.7	−2.1	426
Phospholipase A2	31	56	22.2	−0.2	152
Carboxypeptidase A	193	217	5.3	−0.5	415
Concanavalin A	39	76	22.9	−5.0	262
Ribonuclease A	38	46	48.8	−1.7	371
Alcohol deh. mon.	189	202	7.2	−1.6	623
Alcohol deh. dim.	283	296	0	−3.1	936
Myoglobin oxy	79	88	7.6	−0.5	217
Myoglobin deoxy	79	80	15.4	−0.4	241
Chymotrypsin	70	102	23.1	−1.6	484
Chymotrypsinogen A	82	104	22.1	−1.8	452

The ionic strength is 10 mM.

dipole moment is characterized by a small angle between the moments calculated by both methods.

From Table 3 it can be also seen that protein charges estimated on the basis of a null model at the pH corresponding to the isoelectric points obtained from Poisson-Boltzmann calculations are in some cases quite substantial.

Dipole moments at isoelectric points

Table 4 presents dipole moments of the proteins together with the corresponding limiting reduced dichroisms. All are estimated at the isoelectric points of the proteins obtained from the computed titration curves. The dipole moments listed are the dipole moment according to the null model (i.e., all ionizable sites have their model compound pK_a values, and there is no interaction between ionizable sites in the protein), the mean dipole moment obtained by the Monte Carlo method, and the root-mean-square dipole moment, also resulting from the same Monte Carlo run. All are obtained as the sum of the dipole moment resulting from ionizable sites and the neutral form dipole moment, approximated by the sum of the peptide bond dipoles.

Because we do not know the rate at which proteins exchange protons with solvent, we should expect that the experimental dipole moment falls between the calculated mean and root-mean-square values. Depending on the magnitude of the dipole moments, a difference of 20 Debye units for smaller dipoles, and of 50 Debye units for larger dipoles, should not be considered as showing inconsistency between the experimental and computed values.

Comparing the calculated and experimental dipole moments, we should keep in mind that the results from dielectric relaxation measurements (e.g., those reported in the work by Takashima and Asami, 1993) are frequently obtained in the absence of a buffer; thus the conditions of the experiments are not well controlled. From the data in Table 4 we may see that satisfactory agreement between the experimental dipole moments and those computed with the methodology based on the Poisson-Boltzmann equation is observed for lysozyme, phospholipase A2, ribonuclease A, myoglobin, and chymotrypsin. Previously this methodology was shown to give good results for hemoglobins (Antosiewicz and Porschke, 1995). However, in some cases the null model seems to perform better than the more sophisticated method, but final conclusions require additional investigations because of the conditions mentioned above.

In principle, the limiting reduced dichroism would provide a further test of the computational methodology. In particular, it is interesting that quite different values of the limiting reduced dichroism can be obtained by the different models, despite the fact that they give similar dipole moments (see Table 4).

Dependence of electrical parameters on pH and ionic strength

Another interesting problem that may be studied is the dependence of the dipole and charge on pH and ionic strength. South and Grant (1972) reported results for horse and sperm whale myoglobin as a function of pH. The dipole moment of the sperm whale myoglobin was 150–160 Debye units for the pH range 6–8. For the horse myoglobin the change was more pronounced, from 150 D at pH 5.5 to almost 200 D at pH 7.5. Another study of pH dependence of the dipole moment of sperm whale myoglobin was presented by Schlecht (1969). The pH dependencies of the dipole moment from the above two papers and that calculated in the present work are shown in Fig. 1. It may be seen that there is a pronounced difference between the two experimental results below pH 7. The calculated data agree very well with those of South and Grant and with those of Schlecht above pH 7.

Another pH study was reported by Antosiewicz and Porschke (1989a) for chymotrypsin. The previous theoretical study with the Tanford-Kirkwood model appeared to fail for pH 4.2, and the null model seemed to be better. The

TABLE 4 Calculated values of the null model dipole moment, m_{nm}, the mean dipole moment, m_m, and the root-mean-square dipole moment, m_{rms}, and the corresponding limiting reduced dichroisms, ξ_{nm}, ξ_m, ξ_{rms}, respectively

Protein	m_{nm}	ξ_{nm}	m_m	ξ_m	m_{rms}	ξ_{rms}	m_{exp}
Lysozyme triclinic	185	−.017	115	0.134	156	0.042	122
Lysozyme tetragonal	141	0.207	85	−.082	141	−.059	122
Subtilisin	294	0.268	226	0.118	241	0.082	341
Phospholipase A2	177	0.385	146	0.344	154	0.339	141
Carboxypeptidase A	623	−.168	554	−.153	568	−.165	637
Concanavalin A	290	0.106	250	−.179	275	−.180	411
Ribonuclease A	434	0.297	355	0.364	370	0.340	350
Alcohol deh. mon.	815	0.229	554	−.077	594	−.079	
Alcohol deh. dim.	741	−.762	653	−.761	754	−.568	
Myoglobin oxy	158	−.702	153	−.615	172	−.473	167
Myoglobin deoxy	184	−.562	179	−.690	195	−.606	
Chymotrypsin	550	0.211	506	0.214	517	0.221	480
Chymotrypsinogen A	444	−.130	400	0.006	424	0.001	

For comparison, experimental values are also given, m_{exp}. All values of dipole moments are given in Debye units and refer to isoelectric points and are relative to the center of diffusion. When the protein binds an ion like Mg^{2+}, etc., this is included in the dipole of the protein.

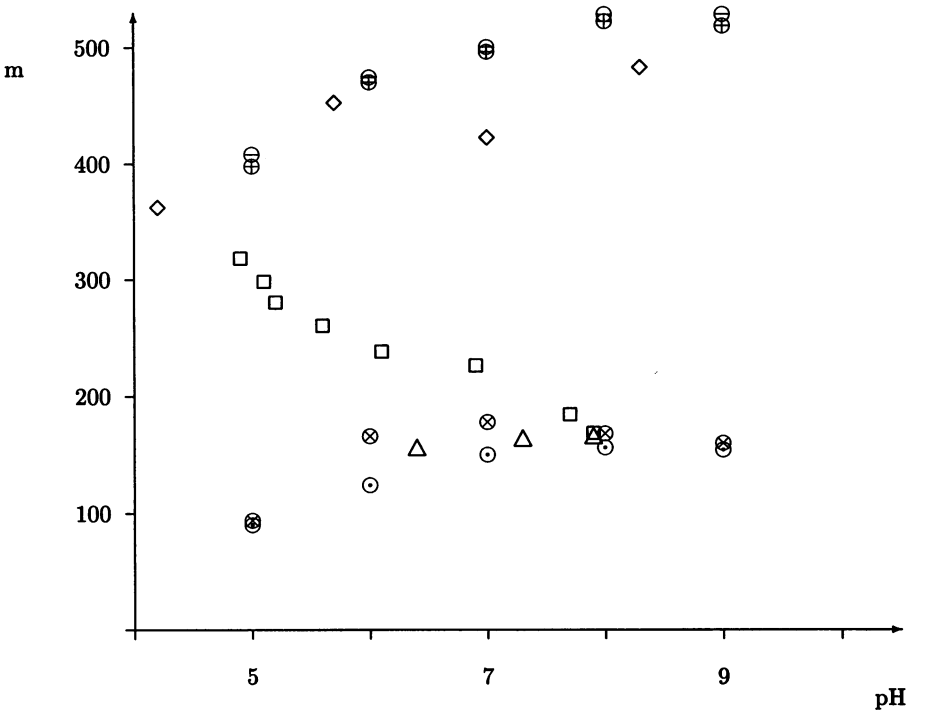

FIGURE 1 Comparison of calculated and experimental dipole moments of sperm whale oxymyoglobin (lower part) and bovine α-chymotrypsin (upper part), as a function of pH. Dipole moments are given in Debye units. Oxymyoglobin: ⊙, calculated mean dipole moment; ⊗, calculated root-mean-square dipole moment; △, experimental data from South and Grant (1972); □, experimental data from Schlecht (1969). α-chymotrypsin; ⊕, calculated mean dipole moment; ⊖, calculated root-mean-square dipole moment; ◇, experimental data from Antosiewicz and Porschke (1989a).

results reported here, however, are in very good agreement with the experimental data (Fig. 1).

For most of the proteins investigated in this work, experimental data for a wider range of pH values are not available. Our calculations indicate that dipole moments can vary substantially with pH, with the most common result being a maximum in the magnitude of the dipole moment around pH 7 (data not shown). We also performed calculations at different ionic strengths, with the most interesting result being that higher strength allows the protein to bear larger positive charge below the isoelectric point and larger negative charge above the isoelectric point. Isoelectric points are not affected by the ionic strength.

CONCLUSIONS

We believe that this work provides a reliable method for the computation of the dipole moments of proteins. This method includes recent developments in the finite-difference Poisson-Boltzmann methodology, which allows us to treat proteins of arbitrary shape and at arbitrary experimental conditions. This method is computationally fast: for a protein with a few hundred amino acids, computations on a workstation can be performed within a few hours.

This methodology, together with transient electric dichroism experiments, can form a powerful tool for studying the electrostatic properties of proteins. The possibility of monitoring dipole moments and their orientation within molecular axes, over a wide range of pH and ionic strengths, provides a strong test for the theoretical approach. Moreover, this methodology can be used for studies of the overall structure of large macromolecules in solution by comparison of experimental and calculated electric dichroism transients (see, e.g., Porschke et al., 1988).

I thank Professors J. A. McCammon, D. Porschke, and S. Takashima and Dr. A. Elcock for helpful discussions and for reading the manuscript.

This work was supported by a grant to Prof. J. A. McCammon from NIH and by State Committee for Scientific Research, Poland, within the grant KBN-4.0078.91.01. Computations of the diffusion tensors and the centers of diffusion for investigated proteins were performed at the facilities of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH, Göttingen, Germany.

REFERENCES

Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Theoretical prediction of pH-dependent properties of proteins. *J. Mol. Biol.* 238: 415–436.

Antosiewicz, J., and D. Porschke. 1989a. The nature of protein dipole moments: experimental and calculated permanent dipole of α-chymotrypsin. *Biochemistry.* 28:10077–10078.

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

Antosiewicz, J., and D. Porschke. 1993. Brownian dynamics simulation of electrooptical transients for complex macrodipoles. *J. Phys. Chem.* 97: 2767–2773.

Antosiewicz, J., and D. Porschke. 1995. Electrostatics of hemoglobins from measurements of the electric dichroism and computer simulations. *Biophys. J.* 68:655–664.

Barlow, D. J., and J. M. Thornton. 1986. The distribution of charged groups in proteins. *Biopolymers.* 25:1717–1733.

Bernstein, F. C., T. F. Koettzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. J. Tasumi. 1977. The protein data bank: a computer-based archival file for molecular structures. *J. Mol. Biol.* 123:557–594.

Biosym Technologies. 1992. INSIGHT II. Biosym Technologies Inc, San Diego, CA.

Böttcher, C. J. F. 1973. Theory of Electric Polarization, Vol. 1. Elsevier, Amsterdam.

Brenner, H. 1965. Coupling between the translational and rotational brownian motions of rigid particles of arbitrary shape. *J. Colloid. Sci.* 20:104–122.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1982. CHARMm: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

Brune, D., and S. Kim. 1993. Predicting protein diffusion coefficients. *Proc. Natl. Acad. Sci. USA.* 90:3835–3839.

Brunger, A. T., and M. Karplus. 1988. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins Struct. Funct. Genet.* 4:148–156.

Cantor, C. R., and P. R. Schimmel. 1980. Biophysical chemistry, Part II. W. H. Freeman and Company, San Francisco.

Cantor, C. R., and S. N. Timasheff. 1982. Optical spectroscopy of proteins. *In* The Proteins, Vol. 5. H. Neurath and R. L. Hill, editors. Academic Press, New York. 145–306.

Cohn, E. J., and J. T. Edsall. 1943. Proteins, Amino Acids and Peptides as Ions and Dipolar Ions. Reinhold Publishing, New York.

Colonna-Cesari, F., D. Perahia, M. Karplus, H. Eklund, C. I. Branden, and O. Tapia. 1986. Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode. *J. Biol. Chem.* 261:15273–15280.

Creighton, T. E. 1993. Proteins. Structures and Molecular Properties, 2nd ed. W. H. Freeman, New York. 264–270.

Davis, M. E., J. D. Madura, B. A. Luty, and J. A. McCammon. 1991. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comp. Phys. Commun.* 62:187–197.

Davis, M. E., and J. A. McCammon. 1991. Dielectric boundary smoothing in finite difference solutions of the Poisson equation: an approach to improve accuracy and convergence. *J. Comput. Chem.* 12:909–912.

Diamond, R. 1974. Real-space refinement of the structure of hen egg-white lysozyme. *J. Mol. Biol.* 82:371–391.

Diekmann, S., W. Hillen, M. Jung, R. D. Wells, and D. Porschke. 1982. Electric properties and structure of DNA restriction fragments from measurements of the electric dichroism. *Biophys. Chem.* 15:157–167.

Dijkstra, B. W., K. H. Kalk, W. G. J. Hol, and J. Drenth. 1981. Structure of bovine pancreatic phospholipase A2 at 1.7 angstroms resolution. *J. Mol. Biol.* 147:97–123.

Drenth, J., W. G. J. Hol, J. N. Jansonius, and R. Koekoek. 1972. A comparison of the three-dimensional structures of subtilisin/BPN$* and subtilisin novo. *Cold Spring Harb. Symp. Quant. Biol.* 36:107–116.

Einstein, A. 1905. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Physik.* 17:549–560.

Fasman, G. D., ed. 1976. CRC Handbook of Biochemistry and Molecular Biology, 3rd ed, Vol. 1, p. 187, Vol. 2, pp. 406–508. CRC Press, Cleveland, OH.

Fredericq, E., and C. Houssier. 1973. Electric Dichroism and Electric Birefringence. Clarendon Press, Oxford.

Fröhlich, H. 1958. Theory of Dielectrics. Oxford University Press, Oxford.

Gaigalas, A. K., J. B. Hubbard, M. McCurley, and S. Woo. 1992. Diffusion of bovine serum albumin in aqueous solutions. *J. Phys. Chem.* 96: 2355–2359.

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

Gilson, M. K. 1993. Multiple-site titration and molecular modeling: two rapid methods for computing energies and forces for ionizable groups in proteins. *Proteins Struct. Funct. Genet.* 15:266–282.

Gilson, M. K., M. E. Davis, B. A. Luty, and J. A. McCammon. 1993. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. *J. Phys. Chem.* 97:3591–3600.

Gilson, M. K., K. A. Sharp, and B. H. Honig. 1988. Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comput. Chem.* 9:327–335.

Gorbunoff, M. J. 1984. The interactions of proteins with hydroxyapatite. I. Role of protein charge and structure. *Anal. Biochem.* 136:425–432.

Grant, E. H., R. J. Sheppard, and G. P. South. 1978. Dielectric Behaviour of Biological Molecules in Solution. Clarendon Press, Oxford. 207–213.

Hames, B. D., and D. Rickwood. 1981. Gel electrophoresis of proteins: a practical approach. IRL Press Limited, Oxford.

Hardman, K. D., and C. F. Ainsworth. 1972. Structure of concanavalin A at 2.4 angstroms resolution. *Biochemistry.* 11:4910–4919.

Harvey, S. C., and J. Garcia de la Torre. 1980. Coordinate systems for modeling the hydrodynamic resistance and diffusion coefficients of irregularly shaped rigid macromolecules. *Macromolecules.* 13:960–964.

Hol, W. G. J. 1985. The role of the α-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.* 45:149–195.

Holcomb, D. N., and I. Tinoco. Jr. 1963. Electrical birefringence at high fields. *J. Chem. Phys.* 67:2691–2698.

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. Sect. A.* 45:851–861.

Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

Keefe, S. E., and E. H. Grant. 1974. Dipole moment and relaxation time of ribonuclease. *Phys. Med. Biol.* 19:701–707.

Kirkwood, J. G. 1939. The dielectric polarization of polar liquids. *J. Chem. Phys.* 7:911–919.

Kirkwood, J. G., and J. B. Shumaker. 1952. The influence of dipole moment fluctuations on the dielectric increment of proteins in solutions. *Proc. Natl. Acad. Sci. USA.* 38:855–862.

Klapper, I., R. Hagstrom, R. Fine, K. Sharp, and B. Honig. 1986. Focusing of electric fields in the active site of Cu,Zn superoxide dismutase. *Proteins Struct. Funct. Genet.* 1:47–79.

Maki, I., K. Kitaura, and K. Nishimoto. 1978. The effect of Σ-II interaction on the electronic structure of molecules in the excited states. II. Oscillator strength. *Bull. Chem. Soc. Jpn.* 51:401–406.

Malamud, D., and J. W. Drysdale, 1978. Isoelectric points of proteins: a table. *Anal. Biochem.* 86:620–647.

Matthew, J. B., G. I. H. Hanania, and F. R. N. Gurd. 1979. Electrostatic effects in hemoglobin. I. Hydrogen ion equilibria in human deoxy- and oxyhemoglobin A. *Biochemistry.* 18:1919–1928.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Molecular Simulations. 1992. Polar hydrogen parameter set for CHARMm Version 22. Molecular Simulations, Waltham, MA.

Norden, B. 1978. Applications of linear dichroism spectroscopy. *Appl. Spectrosc. Rev.* 14:157–248.

Nozaki, Y., and C. Tanford. 1967. Examination of titration behavior. *Methods Enzymol.* 11:715–734.

O'Konsky, C., K. Yoshioka, and W. H. Orttung. 1959. Electric properties of macromolecules. IV. Determination of electric and optical parameters from saturation of electric birefringence in solutions. *J. Phys. Chem.* 63:1558–1565.

Oncley, J. L. 1943. The electric moments and relaxation times of proteins as measured from their influence upon the dielectric constants of solutions. *In* Proteins, Amino Acids and Peptides. E. J. Cohn and J. T. Edsall, editors. Hafner Publishing, New York.

Onsager, L. 1936. Electric moments of molecules in liquids. *J. Am. Chem. Soc.* 58:1486–1493.

Orttung, W. H. 1968. Anisotropy of proton fluctuations and the Kerr effect of protein solutions. Theoretical considerations. *J. Phys. Chem.* 72: 4058–4066.

Orttung, W. H. 1969. Calculation of the mean-square dipole moment and proton fluctuation. *J. Phys. Chem.* 73:418–423.

Philips, L. A., and D. H. Levy. 1986. The rotationally resolved electronic spectrum of indole in the gas phase. *J. Chem. Phys.* 85:1327–1332.

Phillips, S. E. V., and B. P. Schoenborn. 1981. Neutron diffraction reveals oxygen-histidine hydrogen bond in oxymyoglobin. *Nature.* 292:81–82.

Poland, D. 1978. Cooperative Equilibria in Physical Biochemistry. Clarendon Press, Oxford.

Porschke, D. 1985. Effects of electric fields on biopolymers. *Annu. Rev. Phys. Chem.* 36:159–178.

Porschke, D. 1987. Electric, optical and hydrodynamic parameters of lac repressor from measurements of the electrical dichroism. High permanent dipole moment associated with the protein. *Biophys. Chem.* 28: 137–147.

Porschke, D., and J. Antosiewicz. 1990. Permanent dipole moment of tRNAs and variation of their structure in solution. *Biophys. J.* 58: 403–411.

Porschke, D., K. Tovar, and J. Antosiewicz. 1988. Structure of the Tet repressor and Tet repressor-operator complexes in solution from electrooptical measurements and hydrodynamic simulations. *Biochemistry.* 27:4674–4679.

Ramanadham, M., L. C. Sieker, and L. H. Jensen. 1981. Structure of triclinic lysozyme and its Cu(2+) complex at 2 Angstroms resolution. *Acta Crystallogr. A* 37(Suppl.):33.

Rees, D. C., M. Lewis, and W. N. Lipscomb. 1983. Refined crystal structure of carboxypeptidase A at 1.54 angstroms resolution. *J. Mol. Biol.* 168:367–387.

Richards, F. M. 1977. Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* 6:151–176.

Rizzo, V., and J. A. Schellman. 1984. Matrix-method calculation of linear and circular dichroism spectra of nucleic acids and polynucleotides. *Biopolymers.* 23:435–470.

Scheider, W. 1965. Dielectric relaxation of molecules with fluctuating dipole moment. *Biophys. J.* 5:617–628.

Schlecht, P. 1969. Dielectric properties of hemoglobin and myoglobin. II. Dipole moment of sperm whale myoglobin. *Biopolymers.* 8:757–765.

Sober, H. A. 1968. Handbook of Biochemistry. Selected Data for Molecular Biology. The Chemical and Rubber Co, Cleveland.

South, G. P., and E. H. Grant. 1972. Dielectric dispersion and dipole moment of myoglobin in water. *Proc. R. Soc. Lond. A.* 328:371–387.

Stryer, L. 1981. Biochemistry. 2nd ed. W. H. Freeman and Co, New York.

Takano, T. 1984. Refinement of myoglobin and cytochrome C. *In* Methods and Applications in Crystallographic Computing. S. R. Hall and T. Ashida, editors. Oxford University Press, Oxford. 262.

Takashima, S. 1993. Use of protein database for the computation of the dipole moments of normal and abnormal hemoglobins. *Biophys. J.* 64:1550–1558.

Takashima, S., and K. Asami. 1992. Measurement and calculation of the dipole moment of small proteins. *In* Colloid and Molecular Electrooptics 1991. B. R. Jennings and S. P. Stoylov, editors. Institute of Physics Publishing, Bristol.

Takashima, S., and K. Asami. 1993. Calculation and measurement of the dipole moment of small proteins: use of protein data base. *Biopolymers.* 33:59–68.

Tanford, C., and J. G. Kirkwood. 1957. Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* 79: 5333–5339.

Tanford, C., and R. Roxby. 1972. Interpretation of protein titration curves. Application to lysozyme. *Biochemistry.* 11:2192–2198.

Tsukada, H., and D. M. Blow. 1985. Structure of α-chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.* 184:703–711.

Wang, D., W. Bode, and R. Huber. 1985. Bovine chymotrypsinogen A. X-ray crystal structure analysis and refinement of a new crystal form at 1.8 angstroms resolution. *J. Mol. Biol.* 185:595–624.

Warwicker, J., and H. C. Watson. 1982. Calculation of the electric potential in the active site cleft due to α-helix dipoles. *J. Mol. Biol.* 157:671–679.

Yamamoto, Y., and J. Tanaka. 1972. Polarized absorption spectra of crystals of indole and its related compounds. *Bull. Chem. Soc. Jpn.* 45:1362–1366.