Use of protein database for the computation of the dipole moments of normal and abnormal hemoglobins

Shiro Takashima
Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6392 USA

ABSTRACT Previously, we discussed the calculation of the dipole moments of small proteins using the three-dimensional protein database. Our results demonstrate that the calculated dipole moments are in acceptable agreement with measured values. We, however, noted the difficulty of the calculation with larger proteins, in particular those consisting of several subunits. Hemoglobin (Hb) is a protein having a molecular weight of 64,000 that consists of four subunits, a typical case where the computation was found to be difficult. To circumvent the difficulties, we calculated the dipole moment of each subunit separately. The dipole moment of the whole protein was calculated by the vectorial summation of subunit moments. With this method, the calculated net dipole moment is in good agreement with the experimental value. Our calculation shows that the dipole moment vectors of subunits are, by and large, antiparallel in tetramers causing partial cancellation of the net dipole moment. In addition to normal HbA, the dipole moment of abnormal HbS was calculated using an approximate computational technique. Because of the loss of two negative charges as a result of the replacement of glutamic acid with valine in β-chains, the dipole moment of HbS was found, experimentally and theoretically, to be significantly smaller than that of HbA.

INTRODUCTION

The calculation of the dipole moment of macromolecules, particularly globular proteins, had been considered extremely cumbersome. Only for some synthetic polyamino acids in a helical form where the internal regularity is very high had the computation of dipole moment had reasonable success (1). Globular proteins are, in general, the admixture of helical and random coil configurations and their internal structure is much more complex than simple helical polyamino acids. Because of the predominance of nonhelical configurations and the lack of internal regularity, the calculation of the dipole moments of globular proteins is more difficult than simple uncharged helical polymers. Because of the rapid progress in x-ray techniques, however, the structural intricacy that prevented the computation began to unravel for small protein molecules.

The dipole moment of protein molecules consists of three components: the first due to fixed charges on the surface, the second due to polar groups in main and side chains, and the last due to the mobile ion fluctuations. The computation of these components requires the knowledge of the coordinates of charge and dipolar groups in the molecule. x-Ray crystallographic analyses began with myoglobin and hemoglobin (Hb) by Kendrew et al. (2) and Perutz et al. (3) and, not surprisingly, the early computations of the dipole moments of globular proteins began with these proteins by several investigators (4–6). By and large, the calculated dipole moments of these proteins are in reasonable agreement with experimental results. The computation of dipole moment was extended, in recent years, to other proteins by Barlow and Thornton (7), Antosiewicz and Porschke (8), and Takashima and Asami (9, 10). These studies demonstrated that the methods of computation used by early workers can be used, with some modifications and/or refinements, for other proteins as long as their molecular weights are not excessively large. The protein database that provides the necessary three-dimensional coordinates of charged sites has been expanded rapidly in recent years to include more than a few hundred proteins (see Bernstein [11]). Combining the expanded database with markedly improved computational capabilities, the calculation of dipole moment of medium-sized proteins is no longer prohibitively difficult.

In addition to the calculation of the dipole moment of small proteins, Takashima et al. (unpublished results) attempted to extend the same technique to proteins with moderate molecular weights, such as alcohol dehydrogenase and human Hb. They found that the agreement becomes progressively worse as the molecular weight increases, in particular for those proteins having several subunits. Typically, Hb, which consists of four subunits, is one of the proteins that were not amenable to the method used by Takashima et al. in the previous work.

This paper will discuss the method and results of the computation of the dipole moments of normal and abnormal human Hb's, i.e., HbA and HbS, using, basically, a computational technique similar to the one discussed in the previous paper. As before, the calculated dipole moments of these proteins are compared with experimental results in order to validate the calculations. The previous data obtained by Oncley (12) for horse oxy Hb are not directly applicable to the present calculation. Likewise, no previous data are available for HbS in the concentration range of the current interest. Thus this paper consists of experiments and numerical calculations.

Address correspondence to Dr. Shiro Takashima, Department of Bioengineering, University of Pennsylvania, Hayden Hall, Suite 120, 240 South 33rd Street, Philadelphia, PA 19104-6392.

0006-3495/93/05/1550/09 $2.00

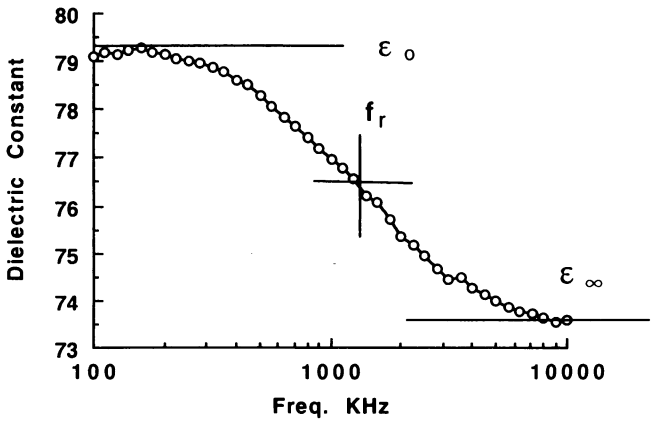

FIGURE 1 The frequency dependence of the dielectric constant of HbS solution. The abscissa is frequency in kHz and the ordinate is dielectric constant. Dielectric increment is defined as $\Delta\epsilon = \epsilon_0 - \epsilon_\infty$. f_r is relaxation frequency.

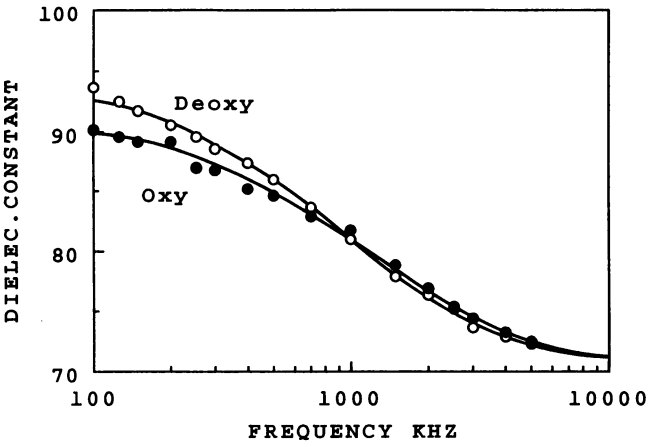

FIGURE 2 The frequency dependence of the dielectric constant of deoxy HbA (*open circles*) and oxy HbA (*filled circles*).

EXPERIMENTAL TECHNIQUES AND RESULTS

HbA was purified using a DEAE column at the Children's Hospital of Philadelphia and is chromatographically proven to be of a single component. HbS was purified by Dr. H. Mizukami at Wayne State University (Detroit, MI) and is also chromatographically pure.

In frequency domain techniques, dipole moment is calculated from measured dielectric constant (using an impedance analyzer model 4191A; Hewlett-Packard Co., Palo Alto, CA). The frequency range of measurements was between 10 kHz and 10 MHz. The dielectric measurement with a sample solution was followed by measurement with a matched reference solution with its conductance adjusted to that of sample solution. The purpose of this procedure is to facilitate the correction of measured sample capacitance for electrode polarization and other artifacts such as false frequency dependence. The dielectric constant was calculated using Eq. 1 (13):

$$\epsilon = A(\epsilon_w - 1) \quad \text{where} \quad A = (C_s - C_0)/(C_w - C_0), \quad (1)$$

where C_s, C_0, and C_w are the capacitances with samples, air, and matched salt solutions. ϵ_w is the dielectric constant of water, i.e., 78.5 at 25°C. The conductances of protein solutions are ~0.5–1 ms/cm, which is equivalent to 0.01 M KCl.

Typical results of the dielectric constant measurement with deoxy HbA and -S are shown in Figs. 1 and 2. The dielectric constant of deoxy Hb was measured using an airtight dielectric cell that was filled with pure nitrogen. At least 5 h were allowed before the beginning of the measurement. From the curves in these figures, we calculate specific dielectric increments $\delta = (\epsilon_0 - \epsilon_\infty)/g$, where ϵ_0 and ϵ_∞ are low and high frequency limiting dielectric constants and g is the weight concentration of proteins in grams per liter. The values of δ are plotted against the concentration and the intrinsic dielectric increment was obtained by extrapolating the curves to zero concentration (see Fig. 3). Although the intrinsic increment of oxyhemoglobin is not shown in this figure, its δ value is slightly lower than that of deoxy Hb (14). It should be noted that the dielectric increment of HbS is considerably smaller than that of HbA. The values of intrinsic δ are 0.42 and 0.45 for oxy and deoxy HbA and 0.147 for deoxy HbS. The dipole moments can be calculated using the following equation (reference 12, also see reference 15):

$$\mu^2 = \frac{9{,}000kTM}{4\pi Nh}\,\delta, \quad (2)$$

where M is molecular weight, N is Avogadro's number, and h is a correction factor for internal fields; its numeri-

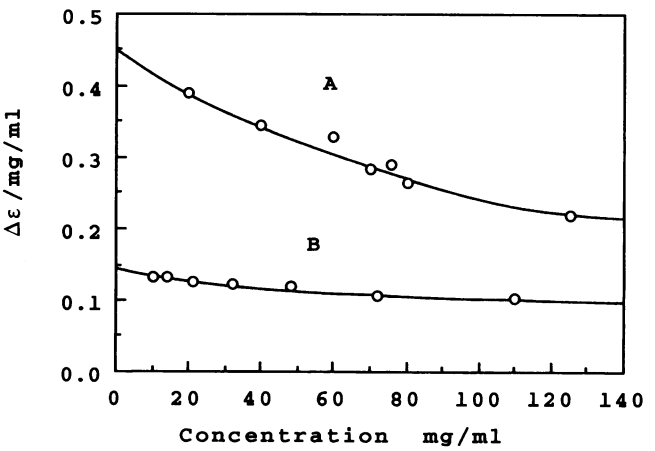

FIGURE 3 The specific dielectric increments of deoxy HbA (*curve A*) and HbS (*curve B*). Extrapolation of these plots to zero concentration gives an intrinsic increment δ.

cal value was determined by Oncley (12) by applying Eq. 2 to a simple amino acid, glycine, solution with a dipole moment that had been well established. A value of 5.8 was obtained. δ is intrinsic dielectric increment. The dielectric dispersion curve of oxyhemoglobin reaches a plateau at 100 kHz and no further increase was observed below this frequency. However, the dielectric dispersion of deoxy Hb is somewhat equivocated by the presence of an additional small dispersion below 100 kHz (not shown in the figure). However, this additional increment does not seem to have a direct relevance to the permanent dipole moment of Hb and, therefore, was disregarded.

Method of computation

The three-dimensional coordinates of amino acid residues of Hbs were obtained from the Brookhaven Protein Databank (Brookhaven, NY). The data on human deoxy Hb and oxy Hb were based on the works by Perutz et al. (16), Ten Eyck and Arnone (17), and Fermi (18). The data on sickle cell anemia deoxy HbS were obtained from Padlan and Love (19, 20). Of these, oxy Hb data were listed only as a pair of α- and β-chains instead of a tetramer. On the other hand, HbSs are listed as an octamer, i.e., two tetrameric HbS molecules as a dimer.

The positive and negative charges included in the calculation of fixed charge dipole moments are listed in Appendix A. Of 10 histidine residues in α-chains, only 6 residues are titratable (21). Likewise, four of nine histidine residues in β-chains are masked. The detailed acid–base titration data are not, to the author's knowledge, available for HbS and all the charged amino acid residues are incorporated in the calculation. Thus the results with HbS should be construed as approximate. Although some tyrosine residues are untitratable, the masking of tyrosines has no effect on the calculation of dipole moment because of their pK being far from the isoelectric point (where all the computations were carried out) and thus they are all protonated.

The dipole moment produced by a group of positive and negative charges is defined by the following equation:

$$\mu = (\sum n_j L_j) \cdot e \cdot R_{N-P}, \quad (3)$$

where e is elementary charge and R_{N-P} is the distance between positive and negative charge centers and is given by the following equation

$$R_{N-P} = \sqrt{(X_+ - X_-)^2 + (Y_+ - Y_-)^2 + (Z_+ - Z_-)^2}, \quad (4)$$

where X_+, X_-, Y_+, Y_-, Z_+, and Z_- are given by

$$X_+ = \sum \{L_{j+} \cdot \sum X_j\} \quad X_- = \sum \{L_{j-} \cdot \sum X_j\} \quad (5)$$

$$Y_+ = \sum \{L_{j+} \cdot \sum Y_j\} \quad Y_- = \sum \{L_{j-} \cdot \sum Y_j\} \quad (6)$$

$$Z_+ = \sum \{L_{j+} \cdot \sum Z_j\} \quad Z_- = \sum \{L_{j-} \cdot \sum Z_j\}. \quad (7)$$

The coordinates X_j, Y_j, and Z_j are found in the protein data base (11). In these equations, L_j is given by Henderson–Hasselbach equations 8 and 9 (see reference 21).

$$L_{j+} = \frac{1}{1+B} \quad \text{for Lys, Arg, His, and } NH_2\text{-terminal} \quad (8)$$

$$L_{j-} = \frac{B}{1+B}, \quad \text{for Asp, Glu, Tyr, Pro, and COOH-terminal} \quad (9)$$

where B is defined as $B = 10^{pH-pK}$. All the computations were carried out at the isoelectric point where the effective positive and negative charges are equal, i.e.,

$$\sum n_{j+} \cdot L_{j+} = \sum n_{j-} \cdot L_{j-}, \quad (10)$$

where n_j is the number of jth amino acid residues. No attempt was made to calculate the dipole moment at pHs other than isoelectric point.

The magnitude of the dipole moment due to fixed surface charges depends on the pK of polar amino acid residues. The pK's of amino acids are shifted in protein molecules from the intrinsic pK's mainly because of the electrostatic interactions between charged sites. The electrostatic interaction between ionizable groups in proteins was first investigated by Kirkwood (22). This theory was then applied by Tanford and Kirkwood (23) for the calculation of the pK shifts in spherical molecules with reasonable agreement with experimental results. In the present work, the pK shifts were computed using the theory by Tanford et al. because of its relative ease of the computation and also because of the fact that the Hb molecule can be approximated by a sphere and is applicable to Tanford and Kirkwood's theory. Tanford and Kirkwood's theory was used in a previous work (10) and was proven to give acceptable results for the computation of dipole moments of a variety of small proteins. There are modifications of the original theory by Matthew et al. (24, 25) and by Warshell et al. (26). In addition, the theory by Bashford et al. (27) has a novel feature, enabling the computation of electrostatic energies for molecules with an arbitrary shape. Of these, Warshell and Russell's theory was used, with reasonable success, by Antosiewicz and Porschke (8) for chymotrypsin. It was pointed out by these authors, however, that Warshell and Russell's theory requires a large value for the dielectric constant of protein interior in contrast to a small dielectric constant for protein core used by Tanford et al. In spite of the cumbersome and time-consuming calculation, the incorporation of electrostatic effects has only a small effect on the value of calculated dipole moments.

The pK shift of ionizable groups is correlated to the free energy of proton binding ΔW by the following equation (8):

$$\Delta W = \sum_{k=1}^{m} 2.303\chi kT(\mathrm{pH} - \mathrm{pK}), \quad (11)$$

where χ is the occupation number and is 0 for protonated sites and 1 or -1 for fully charged sites; m is the total number of charged sites. The energy of charging a sphere having an internal dielectric constant ϵ_i was calculated by Tanford and Kirkwood (23) as shown below.

$$W = (e^2/2b) \sum_{k=1}^{m} \sum_{l=1}^{m} \xi_k \xi_l (A_{kl} - B_{kl}) - (e^2/2a) \sum\sum \xi_k \xi_l C_{kl}, \quad (12)$$

where ξ_k and ξ_l are charges at kth and lth sites $+1$, 0, or -1. The term A_{kl} represents the energy of charging in an unbounded medium and the factor involving B_{kl} represents the modification of the energy arising from the protein molecule being a bounded cavity within a medium of a dielectric constant ϵ. The term C_{kl} represents the interaction with the salt ions dissolved in the medium and should reduce to zero for zero ionic strength of the medium. Since the measurements of dipole moment of protein molecules are carried out with very low ionic strengths, pK shifts were also calculated under a condition where C_{kl} is almost zero. Simplification of the theory can be achieved by introducing an assumption that all the charges are distributed on the surface of a spherical cavity and thus the distances between charges and the center of the sphere are all equal. The detailed expressions for A_{kl}, B_{kl}, and C_{kl} are given in Appendix B.

The pK shifts were calculated for individual ionizable amino acids considering every possible configuration of surface charges, i.e., attractive and repulsive interactions among charged sites with opposite and equal signs and also the repulsive interactions within the same type of amino acid residues. The number of charge configurations depends on the pK value relative to the isoelectric pH where all the calculations were carried out. If the pK of an amino acid is far removed from the isoelectric pH, the sites are either all charged or discharged and thus there is only one charge configuration. Only when the pK of an amino acid residue is near the isoelectric pH are the sites partially charged, and this produces a number of charge configurations. In hemoglobin, the isoelectric point is 6.7 and the only amino acid residue having a pK near this pH is histidine. Since the intrinsic pK of histidine is 6.0 and the total number of titratable histidines in Hb is 22, the Henderson factor suggests the number of charge configurations to be about

$$W = (22!)/(15!)(7!) \approx 1.7 \times 10^5. \quad (13)$$

To avoid this excessive number of iterations, we performed the calculation of pK shifts separately for individual subunits $\alpha 1$, $\alpha 2$, $\beta 1$, and $\beta 2$. We use symbols A, B, C, and D for these subunits in accordance with the notation found in the databank. The calculation of proton binding energy ΔW was repeated for every charge configuration and used the lowest ΔW value to determine the pK shift. The pK shifts for each subunit are shown in Appendix A. By calculating the pK shifts for individual subunits, however, only the intrachain interactions are considered, ignoring all the interchain interactions. Although this will produce some errors, the effect of this error on the calculated dipole moment can be deemed small and insignificant.

Calculation of dipole moment

Dipole moment due to fixed surface charges

The calculation of the dipole moment of Hb is somewhat simplified because the pK's of all the amino acid residues except histidine are far removed from the isoelectric pH. Because of this, acidic as well as basic groups are nearly fully charged except histidine. Therefore, only the partial ionization of histidine residues generates multiple charge configurations. Even with this simplification, the number of charge configurations is too large and virtually prevents one-step calculation of the dipole moment of the tetramer as a whole. To facilitate the calculation, the dipole moments of the subunits A, B, C, and D were calculated separately using Eqs. 3–9 and the overall dipole moment of the tetramer was synthesized by vectorially summing up the moments of four subunits.

There are two possible methods for the calculation of the dipole moments. In the first method, the Henderson–Hasselbach factor is considered to represent the effective charge of ionizable groups. When a group of amino acid residues are fully ionized or completely protonated, the effective charge is 1 (or -1) and 0. With the change in pH the group becomes partially ionized or protonated and the effective charge of the group decreases or increases accordingly. Thus the effective charge can be calculated by multiplying the total number of charges by the Henderson factor. In this method, the charge distribution within the same group remains unaltered during the protonation. Simply, the effective charges of a site either diminishes or increases uniformly with pH changes. This method was used for the calculation with HbS molecule and a reasonable result was obtained as discussed later.

The second method considers the Henderson factor as the fraction of the charged sites. The first step of the calculation is to compute the Henderson factor using Eqs. 8 and 9, as in Method I. This in turn enables one to obtain the value of L_j for individual residues. Multiplying the number of histidines by this factor, we obtain the number of ionized histidine residues at the isoelectric pH where the dipole moment is computed. Inserting this in Eq. 1, we can calculate the dipole moment of subunits.

Since the number of histidine residues is small in each subunit, only a small number of iterations are needed. Repeating the same procedure for every combination of charge configurations, we can calculate the mean dipole moment and its variance, $\Delta\mu$,[1] i.e.,

$$\Delta\mu^2 = \langle\mu^2\rangle - \langle\mu\rangle^2. \quad (14)$$

This method was used for HbA and excellent results were obtained. In general, the second method yields better agreements between calculated and measured dipole moments than the first method. However, as mentioned earlier, this method is often impossible to use for large proteins where the number of charged sites is excessive and their pK's close to the isoelectric point, as for Hb. This forced us to resort to the computation by parts and vectorial summation of subunit moments in order to calculate the moment of the whole molecule.

Core dipole moment

In addition to the fixed charge dipole moment, the core dipole moment due to polar groups must be computed. The major contribution to the core moment is the dipole moment of carbonyl groups (2.7 DU). In addition, the moment of N—H bond is substantial (1.31 DU). It is debatable whether we should use only C=O bonds with well-known coordinates for the core moment or use C=O—N—H with a dipole moment of 3.4 D (1). The dipole moment of amide group is well known only in α-helix where C=O and N—H bonds are almost colinear because of an intrachain hydrogen bonding. In proteins, peptide chains are often in a random coil configuration and C=O and N—H are not colinear. In addition, virtually all the protein databases are based on x-ray crystallographic analyses and do not provide the coordinates of H atoms. Under these circumstances, the use of C=O—N—H groups as the core moment requires a great deal of inferences as to the site of H atoms. Under these circumstances, it was decided to use only the dipole moment of C=O groups as the core moment. The computation of the core moment was repeated by incrementing the C=O moment from 2.7 to 3.4 D. The increase in the core moment (at most 2–3% of the core moment) was found to have only a minute effect on the net dipole moment of the whole molecule. In this calculation, all the C=O moments in the main chain as well as those in side chains were used, i.e., glutamine, glutamic acid, asparagine, and aspartic acid. These group moments were vectorially summed for each subunit and added to that of fixed surface charges. It was found that the core moment is much smaller than those due to fixed surface charges.

[1] This $\Delta\mu$ should not be confused with the induced dipole moment due to mobile protons as discussed later (28).

Induced dipole moment due to proton fluctuation

In general, the dipole moment of protein consists of the mean permanent dipole moment (fixed charge and core moments) and a small induced dipole moment associated with the excess fluctuation of mobile protons driven by an electrical field (28). Whereas the calculations of fixed and core dipole moments are exact in nature, there are some ambiguities as to the calculation of proton fluctuation moment. In addition, South and Grant (6) reported that inclusion of the proton fluctuation moment did not have a tangible effect on the overall dipole moment of myoglobin at isoelectric pH. In view of these observations, the proton fluctuation moment was disregarded in this work.

RESULTS

Deoxy HbA

As discussed earlier, the dipole moment of the subunits of normal HbA was computed with and without the electrostatic interaction among charged sites. The pK shifts due to this interaction are tabulated in Appendix A. As shown, electrostatic interaction causes wider spreads of pK values of ionizable groups. However, these pK shifts have no tangible effect on the calculated dipole moment unless the amino acid residues have a pK that is close enough to the isoelectric point. In Hb, only histidine has a pK value close enough to the isoelectric point to have some effect on the calculation.

Table 1 shows the summary of the calculation of dipole moments of HbA with inclusion of the electrostatic interaction. The table shows *a*) the mean fixed charge dipole moments of subunits and the root mean square fluctuation and *b*) the net dipole moment of tetramer as the sum of fixed charge and core moments. The last row shows the coordinates of the dipole moment of tetrameric HbA and the separation of positive and negative charge centers. The calculated net dipole moment 481.04 D is in good agreement with the observed moment 495.4 D. The same calculation was repeated without the electrostatic interaction and the overall dipole moment was found to be 487.33 D (see Table 2). Surprisingly, inclusion of the electrostatic interaction did not necessarily improve the agreement. In any case, the calculated overall moment is sufficiently close to the observed value. It should also be noted that some asymmetry exists between the dipole moments of subunits. In particular, the difference between B and D is far beyond the computational error. Since the amino acid compositions and their sequences are identical, this difference must arise from some distortion of the conformation of either one of these subunits. It should be emphasized, however, that in order to produce these differences, only

TABLE 1 **Calculated and measured dipole moments of HbA with electrostatic interactions**

Fixed charge moments of subunits

	μ_x	μ_y	μ_z	μ^*	$\Delta\mu$
HbA	−94.331	28.822	28.2	102.587	52.412
HbB	9.403	−244.52	−178.12	302.663	52.345
HbC	64.478	18.498	−36.546	76.388	42.326
HbD	−38.818	−257.27	433.245	505.367	53.408

Net dipole moments of tetramer as the sum of fixed charge and core moments

	μ_x	μ_y	μ_z	μ
μ (charge)	−59.273	−454.47	246.772	520.52
μ (core)	0.247	−48.79	4.787	49.01 ($\angle^{**}$ = 164°)
μ (net)	−59.52	−405.681	251.559	*481.041 (cal)*
				495.40 + 10 D(obs)

Coordinates of positive and negative charge centers in tetrameric deoxy HbA

		$R_{N\text{-}P}$
HbA	$X+, Y+, SZ+ = -0.2670, -4.6756, 2.7437$	4.7417 A
	$X-, Y-, Z-, = 0.3787, 1.2476, -0.1453$	

* $\mu = (\mu_x^2 + \mu_y^2 + \mu_z^2)^{1/2}$.

** $\angle$ is the angle between fixed charge and core dipole moments. Experimental values for α-chain is 170 D and >270 D for β-chain (29). The average of A and C is ~149.2 D and average of B and D is 498.39 D.

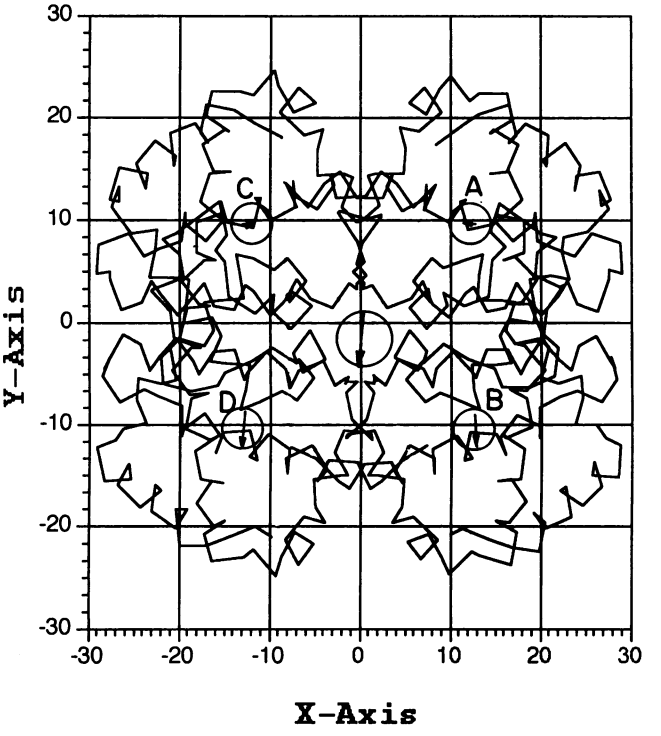

FIGURE 4 The projection of deoxy HbA molecule (C—C bonds only) on the x-y plane. The dipole moments of subunits are illustrated by arrows. The arrow near the center is the net dipole moment of HbA, which was calculated by summing subunit dipole moments vectorially.

small distortions of peptide chain configurations would be sufficient.

Fig. 4 shows the x-y plane projection of the deoxy HbA. The dipole moments due to fixed charges of each subunits are shown by arrows in circles. It should be noted that the dipole moment vectors are quite small compared with the size of the molecule. This indicates the high degree of symmetry of charge distribution resulting in a very small separation of the positive and negative charge centers. The dipole moment vectors of subunits A and C are antiparallel and those of B and D are nearly parallel. It was found that the x-component of the net moment in tetrameric Hb molecule is nearly zero because of the cancellation between A and C.

TABLE 2 **Calculated and measured dipole moment of HbA without electrostatic interactions**

Calculated fixed charge moment

Subunit	μ_x	μ_y	μ_z	μ
A	−94.081	28.325	27.278	101.967
B	15.813	−247.346	−177.548	304.882
C	67.254	15.964	−38.216	78.983
D	−48.512	−262.175	429.657	505.662

Sum of fixed charge and core dipole moments of HbA

μ (charge)	−59.526	−465.232	241.171	527.397
m (core)	−0.247	48.79	4.787	49.01
m (net)	−59.773	−416.442	245.958	487.331
				495.4 (Obs)

Oxy Hb and deoxy Hb

Fig. 2, demonstrates that the dielectric increment of deoxy Hb is slightly larger than that of oxy Hb. However, the difference may be too small to be detected by the computational technique used in this work.

Dipole moment of HbS

The major difference between HbA and HbS is the replacement of glutamic acid residues with valines at the sixth positions of B and D chains, resulting in the loss of two negative charges. X-ray analyses by Padlan and Love (19) indicate considerable differences between the internal structures of HbA and HbS. As discussed in Appendix A, as many as 16 histidine groups are masked in HbA molecules. These masked histidine residues are eliminated for the computation of dipole moments. However, no such detailed information on the acid–base titration could be found for HbS. Under these circumstances, it was decided not to perform a detailed study on

TABLE 3 **Calculated and measured dipole moment of HbS**

Net dipole moment of tetrameric HbS

	μ_x	μ_y	μ_z	net μ
HbS-1				
μ (charge)	−163.05	−105.35	−274.98	336.82
μ (core)	10.108	0.774	85.241	85.84 ($\angle$ = 142°)
μ (net)	−155.94	−109.57	−195.74	*273.20 (cal)*
				286.07 + 10 (obs)
HbS-2				
μ (charge)	−82.052	198.622	248.496	328.532
μ (core)	22.968	−74.876	−27.625	83.048 ($\angle$ = 149°)
μ (net)	−59.084	123.746	220.871	*259.976 (cal)*
				286.07 (obs)

Coordinates of positive and negative charge centers

		$R_{N\text{-}P}$
HbS-1	$X+, Y+, Z+$ = 5.759, 25.595, 31.482	0.986 A
	$X-, Y-, Z-$ = 5.277, 25.278, 30.677	
HbS-2	$X+, Y+, Z+$ = 37.69, 68.170, 48.641	0.961 A
	$X-, Y-, Z-$ = 37.450, 68.751, 49.369	

The angle between two dipole vectors was graphically determined to be ~135°.

the dipole moment of individual subunits. Instead of calculating by parts, the dipole moment of the whole HbS molecule was computed using Method I (see earlier discussion) with all 38 histidine residues, assuming that they were not masked.

The database provides us with the coordinates of two deoxy HbS molecules relative to a common origin, thus enabling us to study not only the dipole moment of each molecule but also the relative orientation of the dipole moments in the dimer. The magnitude and the coordinates of the dipole moments of HbS 1 and HbS 2 molecules are calculated separately using the same method discussed earlier. The experimental data shown in Fig. 3 and Table 2 indicate that the overall dipole moment of HbS molecule is ~260–270 D and is considerably smaller than that of HbA, i.e., 496 D (the calculated value is 481.04 and 487.3 D). The results of the computations are summarized in Table 3.

Table 3 shows the calculated dipole moment of 273.01 D for HbS 1 and 259.9 D for HbS 2. The slight difference between these values may be the limit of the resolution of the computational technique. These are both in good agreement with the observed value. The relative orientation of two dipole vectors can be graphically determined and the dipole angle was found to be ~130°. This means that the dipole moments of two HbS molecules in the dimer are nearly anti-parallel.

DISCUSSION

The computation of the dipole moment of globular proteins had been considered nearly impossible until the successful application of x-ray crystallography to small globular proteins. With the rapid progress of the x-ray crystallography, the information on the three-dimensional coordinates of individual amino acid residues in proteins became available, and the attempt to compute the dipole moment using the three-dimensional databank emerged as early as 1969 (4).

The calculation of the dipole moments of proteins is, in principle, not difficult as long as the x-ray crystallography provides us with an accurate database. For small proteins, such as myoglobin, the resolution is of the order of 1.5 A, which is sufficient for the calculation of the dipole moment. Because of limited computer facilities, early calculations were restricted to small proteins. The resolution of the database for medium-sized proteins such as Hb has been improved greatly from 6.0 to 2.5 A. Because of the improved database combined with rapid improvements of computer facilities, we are now able to compute the dipole moment of protein molecules having a molecular weight of the order of 10^5. With the advent of these technological developments, the calculations have been extended to many other proteins of small to even moderately large proteins (7–10). These investigations clearly demonstrate that the calculation of the dipole moment of globular proteins can be carried out with an acceptable accuracy. A detailed and quantitative discussion on the accuracy of the computation is, however, not feasible. Only when an extensive error analysis is performed in which the coordinates of individual amino acid residues are allowed to fluctuate within the resolution of x-ray crystallography, i.e., 2.5 A of HbA and 3.5 A for HbS, can we more specifically state the error limits of the computation. The excellent agreement between the observed and calculated dipole moments, nevertheless, is indicative of the reliability of the computation.

This work is an application of the computational technique discussed in our previous paper in attempts to study the dipolar structure of Hb molecules in detail. It was pointed out before that the computation of dipole moment was less successful if the proteins consist of several subunits. In addition to this, in Hb, there are 38 histidine residues, if they are all exposed, that have pK's sufficiently close to the isoelectric pH. Under these conditions, the partial dissociation of histidine residues generates an exceedingly large number of charge configurations, which made the computation virtually impossible. To circumvent this difficulty, it was decided to calculate the dipole moment of subunits individually and later sum these dipole moments vectorially. Using this method, we found the calculated dipole moment sufficiently close to the observed value. Moreover, we are able to find the relative orientation of dipole moments of each subunit and that in a tetramer. The orientations of some of these dipole moment vectors are such that the

net dipole moment of HbA is dominated by the moments of subunits B and D.

The interest with the calculation for HbS, in addition to the comparison between calculated and observed dipole moments, is the relative orientation of the dipole moments of HbS 1 and HbS 2 molecules. The calculated angle between μ_1 and μ_2 was found to be $\sim 130°$. This indicates that the vector sum of μ_1 and μ_2 results in a net dipole moment (~ 225 D) that is smaller than those of individual proteins.

APPENDIX A

TABLE A-1 Charged sites in hemoglobinA and -S

	HbA		HbS	
	α	β	α	β
Lysine	11	11	11	11
Arginine	3	3	3	3
Histidine*	6	5	6	5
NH_2-terminal	1	1	1	1
Aspartic acid	8	7	8	7
Glutamic acid	4	8	4	7
Tyrosine	3	3	3	3
Heme	2	2	2	2
COOH-terminal	1	1	1	1

* His-58α, 87α, 103α, 122α, 63β, 92β, 97β, and 116β are masked in deoxy HbA and are omitted for the computation. No such information, to the author's knowledge, was found for HbS. See text for this problem.

TABLE A-2 pK shifts in α and β chains

	Intrinsic	α	β
Lysine	10.53	9.842	9.873
Arginine	12.48	12.798	12.491
Histidine	6.00	5.950	5.825
NH_2-teminal	9.62	10.234	9.819
Aspartic acid	3.86	2.976	3.375
Glutamic acid	4.25	3.405	3.493
Tyrosine	10.07	9.816	9.901
Propionic acid	4.00	3.842	3.796
COOH-terminal	2.17 (α)	1.664	
	1.82 (β)		1.201

APPENDIX B

The electrostatic energy between two charged sites is given by Eq. B1.

$$W = (e^2/2b) \sum_{k=1}^{m} \sum_{l=1}^{m} \xi_k \xi_l (A_{kl} - B_{kl}) - (e^2/2a) \sum\sum \xi_k \xi_l C_{kl}, \quad \text{(B1)}$$

where a is the radius of the protein including the layer of solvent, b is the radius of bare protein molecule, ξ_k and ξ_l are charges on kth and lth sites, and e is charge $+1$ or -1.

$$A_{kl} = b/\epsilon_i r_{kl} \quad \text{(B2)}$$

$$B_{kl} = (1/\epsilon_i) \sum_{n=0}^{\infty} \frac{(n+1)(\epsilon - \epsilon_i)}{(n+1)\epsilon + n\epsilon_i} \rho_{kl}^n P_n(\cos\theta_{kl}), \quad \text{(B3)}$$

where ϵ and ϵ_i are dielectric constants of unbounded medium and that of the space bounded by protein surfaces. Generally $\epsilon_i \ll \epsilon$. $P_n(\cos\theta_{kl})$ is the Legendre polynomial.

$$C_{kl} = (1/\epsilon)\left\{\frac{x}{1+x} + \sum_{n=1}^{\infty}\left(\frac{2n+1}{2n-1}\right)\left(\frac{\epsilon}{(n+1)\epsilon + n\epsilon i}\right)^2 \times \left(\frac{x^2\sigma_{kl} n P_n(\cos\theta_{kl})}{Y}\right)\right\}, \quad \text{(B4)}$$

where

$$Y = \frac{K_{n+1}(x)}{K_{n-1}(x)} + \frac{n(\epsilon - \epsilon_i)}{(n+1)\epsilon + n\epsilon_i}\left(\frac{b}{a}\right)^{2n+1} \frac{x^2}{4n^2 - 1} \quad \text{(B5)}$$

$$x = \kappa a, \quad \rho_{kl} = r_k r_l / b^2, \quad \sigma_{kl} = r_k r_l / a^2.$$

a is the radius of the sphere including the circular area with a dielectric constant ϵ. b is the radius of protein core with a small dielectric constant ϵ_i. κ is the Debye screening constant. r_k and r_l are the distance between the points k and l on the surface. r_{kl} is the distance between points k and l.

The function $K_n(x)$ is given by

$$K_n(x) = \sum_{s=0}^{n} \frac{2^s n!(2n-s)!}{s!(2n)!(n-s)!} x^s. \quad \text{(B6)}$$

Because of the convoluted nature of the calculation, the results of our computation, in particular, the values of C_{kl} function, were compared with those by Tanford et al. The agreement between our calculation and those by the original authors is satisfactory. The value of the parameter x (κa) was assumed to be 0.1 and internal dielectric constant ϵ_i's were assumed to be 6. The radius b of subunits was assumed to be 25 A and a value of 0.85 was used for the ratio b/a.

We thank Drs. K. Adachi and H. Mizukami for their help with HbA and HbS purification, respectively.

This work was supported by the Olympus Corporation of America.

Received for publication 9 September 1992 and in final form 25 January 1993.

REFERENCES

1. Wada, A. 1962. Chain regularity and dielectric properties of poly-α-amino acids in solution. *In* Polyamino Acids, Polypeptides, and Proteins. M. A. Stahman, editor. University of Wisconsin Press, Madison. 131–146.
2. Kendrew, J. C., R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore. 1960. *Nature (Lond.)*. 185:422.
3. Perutz, M. F., M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North. 1960. *Nature (Lond.)*. 185:416.
4. Orttung, W. H. 1969. Calculation of the mean-square dipole moment and proton fluctuation. *J. Phys. Chem.* 73:418–423.
5. Schlecht, P. 1969. Dielectric properties of hemoglobin and myo-

globin. II. Dipole moment of sperm whale myoglobin, *Biopolymers.* 8:757–765.

6. 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.
7. Barlow, D. J., and J. M. Thornton. 1986. The distribution of charged groups in proteins. *Biopolymers,* 25:1717–1733.
8. Antosiewicz, J., and D. Porschke. 1989. The nature of protein dipole moments: experimental and calculated permanent dipole of α-chymotrypsin. *Biochemistry.* 28:10072–10078.
9. Takashima, S., and K. Asami. 1991. Measurement and calculation of the dipole moment of small proteins. *Int. Symp. Colloid Molec. Electro-Optics, Varna, Bulgaria.* 187–195.
10. 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.
11. Bernstein, F. C. 1977. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.
12. Oncley, J. L. 1943. *In* Proteins, Amino Acids and Peptides. E. J. Cohn and J. T. Edsall, editors. Reinhold, New York.
13. Schwan, H. P. 1962. Determination of biological impedances. *In* Physical Techniques in Biological Research. W. L. Nastuk, editor. Vol. 6. Academic Press, New York. 323.
14. Schlecht, P., H. Vogel, and A. Mayer. 1968. Effect of oxygen binding on the dielectric properties of hemoglobin. *Biopolymers.* 6:1717.
15. Takashima, S. 1989. *In* Electrical Properties of Biopolymers and Membrane. Adam Hilger, Bristol, United Kingdom. 91.
16. Perutz, M. F., S. S. Hasnain, P. J. Duke, J. E. Sessler, and J. E. Hahn. 1982. Sterochemistry of iron in deoxyhemoglobin. *Nature (Lond.).* 295:535–538.
17. Ten Eyck, L. F., and A. Arnone. 1976. Three dimensional Fourier synthesis of human deoxyhemoglobin at 2.5 A resolution. I. X-ray analysis, *J. Mol. Biol.* 100:3–11.
18. Fermi, G. 1975. Three dimensional Fourier synthesis of human deoxy-Hb at 2.5A resolution: refinement of the atomic model. *J. Mol. Biol.* 97:237–256.
19. Padlan, E. A., and W. E. Love. 1985. Refined crystal structure of deoxy Hb S. I. Restrained least-squares refinement at 3.0 angstroms resolution. *J. Biol. Chem.* 260:8272–8279.
20. Padlan, E. A., and W. E. Love. 1985. Refined crystal structure of deoxy Hb S. II. Molecular interactions in the crystals. *J. Biol. Chem.* 260:8280–8291.
21. Bray, H. G., and K. White. 1966. *In* Kinetics and Thermodynamics in Biochemistry. Academic Press, New York. 127.
22. Kirkwood, J. G. 1934. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Physics.* 2:351–361.
23. 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.
24. Matthew, J. B., and F. R. N. Gurd. 1986. Calculation of electrostatic interactions in proteins. *Methods Enzymol.* 130:413–436.
25. Matthew, J. B., G. I. H. Hanania, and F. R. N. Gurd. 1979. Electrostatic effects in hemoglobin: Hydrogen ion equilibria in deoxy- and oxyhemoglobin A. *Biochemistry.* 18:1919–1928.
26. Warshell, A., and S. T. Russell. 1984. Calculations of electrostatic interactions in biological systems and in solutions. *Q. Rev. Biophys.* 17:283–422.
27. Bashford, D., and M. Karplus. 1990. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry.* 29:10219–10225.
28. Kirkwood, J. G., and J. B. Shumaker. 1952. The influence of dipole moment fluctuations on the dielectric increment of proteins in solution. *Proc. Natl. Acad. Sci. USA.* 38:855–862.
29. Schlecht, P. 1971. Dielectric studies of aggregation phenomena with separated α and β chains of human hemoglobin. *Z. Naturwiss-Med. Grundlagenforsch.* 26b:453–457.