Method for Estimating the Internal Permittivity of Proteins Using Dielectric Spectroscopy

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · FEBRUARY 2011

Impact Factor: 3.3 · DOI: 10.1021/jp1111873 · Source: PubMed

THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

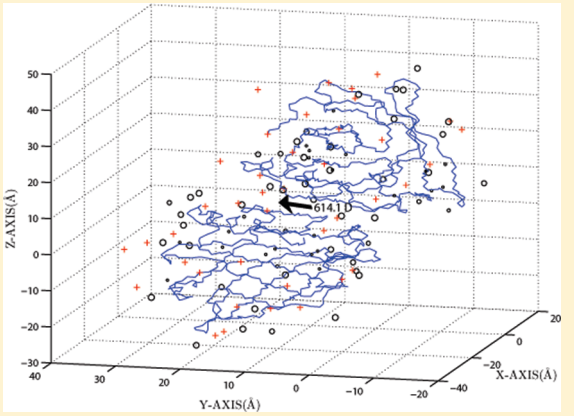
pubs.acs.org/JPCB

Method for Estimating the Internal Permittivity of Proteins Using Dielectric Spectroscopy

Brett L. Mellor,*[,†] Efrén Cruz Cortés,[†] David D. Busath,[‡] and Brian A. Mazzeo[†]

[†]Department of Electrical and Computer Engineering and [‡]Department of Physiology and Developmental Biology, Brigham Young University, Provo, Utah, United States

ABSTRACT: Protein charge organization is dependent on the low-permittivity region in the hydrophobic core of the molecule. We suggest a novel approach to estimate the dielectric constant of this region by comparing measured and simulated first- and second-order charge moments. Here, the dipole moment is measured as a function of pH using dielectric spectroscopy. The results are compared to dipole moments based on Poisson−Boltzmann estimates of pK_a shifts calculated from structures in the Protein Data Bank. Structures are additionally refined using CHARMM molecular dynamics simulations. The best estimate for the internal permittivity is found by minimizing the root-mean-square residual between measured and predicted charge moments. Using the protein β-lactoglobulin, a core dielectric constant in the range of 6−7 is estimated.



INTRODUCTION

The electrical properties of proteins are dependent on the permittivity of the protein interior and the surrounding medium. High-permittivity environments reduce Coulombic forces. Low-permittivity environments where Coulombic forces are greater, such as the hydrophobic core of a protein, are critical to structure and function. The simplest continuum dielectric model approximates the permittivity of the protein, $\varepsilon(\mathbf{r})$, by treating the molecule as a low-permittivity medium with relative permittivity ε_{low} surrounded by a high-permittivity solvent with relative permittivity ε_{high}. This model has been used to improve estimates of energies of interactions within[1,2] and between proteins,[3] electrons and protons,[4−6] charged solutes,[7] and ions moving through proteins.[8−11] Mostly, the continuum dielectric model has been used to estimate titratable amino acid pK_a shifts based on local electric fields in proteins,[12,13] although comparison of these estimates with direct measurements with NMR have revealed examples where other factors are also involved.[14,15] Dielectric spectroscopy (DS) measures overall electrical properties and we postulate that a first-order continuum electrostatic model is sufficiently accurate to compute overall charge moments.[16] Here we use a model protein, β-lactoglobulin (β-Lg), to explore the impact of continuum-electrostatics-derived pK_a shifts on the predicted protein dipole moments and the use of DS to determine the effective dielectric constant. In this first approach, we depend on the simple version of the model implemented in MEAD.[17]

Unlike ε_{high}, which can and has been measured directly,[18] ε_{low} has been found to be the most critical and least-agreed-upon parameter in calculations of electrostatic effects in proteins.[19,20] Though the majority of theoretical calculations and experimental determinations have placed the value of ε_{low} between 1 and 6,[21−30] a number of contradictory studies have emerged. A theoretical study by Nakamura et al. using electronic polarization of atoms and the polarizations of local dipoles predicted internal dielectric constants from 1 to 20.[31] Pennock and Schwan proposed a lower bound on ε_{low} of 2.5 based on series and parallel combinations of the dielectrics of water and protein powders.[32,33] Antosiewicz et al., in calculating pK_a values for ionizable residues in proteins, were only able to predict satisfactory values using $\varepsilon_{low} \geq 20$.[34] Demchuk and Wade additionally improved pK_a predictions using a hybrid technique where an ε_{low} close to aqueous solvent (∼80) was used for solvent-exposed residues and an ε_{low} in the range 10−20 for buried residues.[35] García-Moreno et al. measured ε_{low} in terms of the permittivity that needed to be used in the Born equation to reproduce pK_a shifts measured by difference potentiometry.[36] Measured values were in the range of 10−12. Similarly, Dwyer et al. measured the ΔpK_a of a single buried glutamic acid which reflected a dielectric constant of 12, interpreting this value as a small amount of solvent penetrating into the protein which could contribute significantly to a higher value of ε_{low}.[37] Other interpretations of high dielectric constants include effects of conformational reorganization,[38] fluctuations of charged side chains,[39] and the reaction field of bulk water.[40]

Dielectric spectroscopy (DS) has been used to probe electrical properties of solutions as one of the first biophysical

Received: November 23, 2010
Revised: January 7, 2011
Published: February 23, 2011

characterization techniques stemming from the pioneering work of Debye, Kirkwood, Onsager, Wyman, and Oncley.[41−45] DS is used to probe solution properties in protein studies for two major reasons. First, the magnitude of the measured dielectric increment is related to the second charge moment, the dipole moment, for individual molecules.[43] Second, the shape of the relaxation versus frequency yields size and shape information about the protein rotating in the viscous solvent.[46] As a consequence, applications of dielectric spectroscopy of proteins are diverse, emphasizing the estimation of structural and conformational changes triggered by variations in temperature, pH, and solvent conditions.[47−50]

This paper presents an alternative approach to estimate the low-permittivity core of proteins that compares calculations to measurements made using DS. The electric dipole moment of molecules is a useful calculated parameter because it is both directly related to ε_{low} and can be measured experimentally at various pHs, allowing theoretical and experimental agreement for multiple charge configurations to validate any selection of ε_{low}. The question we aim to answer here is selection of the most physically significant value of ε_{low} inside a specific protein, β-lactoglobulin (β-Lg), when the continuum electrostatics model is used. The process of answering this question will demonstrate the accuracy of the continuum model in charge moment calculations.

β-Lg is an ideal candidate for this study because it possesses a large permanent dipole moment and has interesting binding properties.[51,52] Several genetic variants of β-Lg exist. The two most common, A and B, have similar molecular weights and differ only by two amino acids (one of which is charged) and therefore can be expected to exhibit similar dielectric dispersions.[53] Previous treatments on the dielectric relaxation of β-Lg were limited and did not address the protein's dynamic electrical properties in acidic and basic solutions. The most noteworthy of these was the pioneering work done by Ferry and Oncley who measured the dielectric increment of β-Lg in 0.25 and 0.50 M glycine, at presumably neutral pH.[54,55] Here, the charge moments of β-Lg are measured from pH 3 to 10 and compared to theoretically derived charge moments.

CALCULATIONS

The dipole moment of β-Lg was calculated from multiple models from the Protein Data Bank (PDB), the most notable being 1beb, which contains structural data of β-Lg crystals grown at room temperature and at physiological pH. The method employed here is based on modeling the complex biomolecule as an array of point charges immersed in a two-region dielectric medium, i.e., a medium represented by an internal dielectric inside the solvent accessible surface of the protein and an external dielectric outside.

By convention, the dipole moment vector points from the direction of negative charge toward the direction of positive charge. The general dipole moment equation for a continuous charge density $\rho(\mathbf{r})$ is

$$\mu = \int_V \rho(\mathbf{r})(\mathbf{r} - \mathbf{r}_r)\, dV \quad (1)$$

where $\mathbf{r}_r$ is the vector from the origin to the reference point. For an array of point charges, this is simplified to

$$\mu = \sum_i q_i(\mathbf{r}_i - \mathbf{r}_r) \quad (2)$$

where $\mathbf{r}_i$ is the vector from the origin to point charge q_i. To approximate the center of rotation for the molecule, the center of mass, expressed as

$$\mathbf{r}_{cm} = \frac{\sum_i m_i \mathbf{r}_i}{\sum_j m_j} \quad (3)$$

is used for the reference point $\mathbf{r}_r$.

By using the general dipole moment equation with the center of mass as the reference point, we focus on the torque about that point produced by applying an electric field to a molecule, and disregard the translational movement of the molecule. An alternative choice of reference point is the center of charge, which has been used by other investigators at pH close to the isoelectric point.[56] A better reference might be the center of diffusion,[57,58] but this would require additional simulations and is not expected to improve the calculation beyond the error of the measurements.

Calculation of pK_a Shifts. Partial charges that exist at side chains of ionizable residues are the main contributers to the charge moments of proteins. These charges are governed by the Henderson−Hasselbalch equations,[59] i.e.

$$Q_+ = \frac{1}{1 + 10^{pH - pK_a}} \quad (4)$$

for positively charged residues such as arginine, histidine, lysine, and N-terminal, and

$$Q_- = \frac{10^{pH - pK_a}}{1 + 10^{pH - pK_a}} \quad (5)$$

for negatively charged residues such as aspartic acid, glutamic acid, tyrosine and C-terminal, where Q_+ and Q_- are the percent positively and negatively ionized, respectively. When pH = pK_a for a specific residue, it becomes 50% ionized. Electrostatic interactions among charged sites cause pK_a shifts of polar groups from their intrinsic (bulk aqueous solution amino acid) values, which can be estimated by calculation and compared to measured values.

Among the current methods used to compute the pK_as of residues, the continuum dielectric approach has proven to give a reasonable balance between speed and accuracy.[12,35,60] The model relates the pK_a of a chemical group M in a proton-exchange reaction with water

$$HM + H_2O \leftrightarrow M^- + H_3O^+ \quad (6)$$

to the standard Gibbs free energy ΔG for proton exchange through a Boltzmann distribution of the reaction species

$$pK_a^{calc} = pK_a^m + \frac{1}{2.303 k_b T}[\Delta G(M_pH^+ \rightarrow M_p) - \Delta G(M_sH^+ \rightarrow M_s)] \quad (7)$$

where M_p and M_s are the chemical group M in the protein and solution environments, pK_a^{calc} is the shifted pK_a, and pK_a^m is the model pK_a of the chemical group M. For a continuous charge distribution, electrostatic contributions to the free energy are deduced from the Poisson−Boltzmann equation

$$\nabla\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r}) - \varepsilon(\mathbf{r})\kappa(\mathbf{r})^2 \sinh[\phi(\mathbf{r})] + 4\pi\rho(\mathbf{r})/k_bT = 0 \quad (8)$$

where $\phi(\mathbf{r})$ is the electrostatic potential, $\rho(\mathbf{r})$ is the charge density and $k(\mathbf{r})^2$ is related to the Debye length, accounting for the ionic strength of the bulk solution.[61]

The permittivity $\varepsilon(\mathbf{r})$ is defined as a constant ε_{low} inside the protein and a constant ε_{high} outside the protein. The assumption of a uniform permittivity throughout the protein clearly is an approximation.[31] The actual permittivity is likely to be a function of several factors, including density of packed residues, molecular cavities, and solvent penetration. The ε_{low} approximation, however, has the advantage of being simple, reducing the computational expense.[12] A pK_a shift is calculated from the difference in electrostatic energy of an amino acid residue in its charged and neutral state.

To aid in this computation, several web servers are now available. The H++ server, used primarily for estimating pK_as and adding missing hydrogen atoms to molecules from the PDB, utilizes the continuum dielectric model to predict protonation states of ionizable residues in macromolecules.[62] The H++ calculation is based on the "Macromolecular Electrostatics with Atomic Detail" (MEAD) software package.[17] The server sets up the finite-difference and Poisson−Boltzmann equations, defining the boundary between the internal and external dielectric regions as the surface 1.4 Å outside the protein, and taking as user input, ε_{low}, ε_{high}, and the salt concentration which influences the calculation through the $k(\mathbf{r})^2$ term.

Core Dipole Moment. Due to the difference in electronegativity between different atoms, permanent dipoles exist at covalent bonds throughout the molecular structure. Although individually weak, in a macromolecule their contributions can become significant. We refer to their net dipole as the core dipole moment.

The core dipole moment is calculated from the locations of N−H, C═O, and C−N bonds in both side chains and protein backbone, using dipole moments of 1.31, 2.31, and 0.2 D respectively.[63] For a system of symmetric dipole moments with an arbitrary reference point, the general dipole moment equation (eq 2) can be rewritten

$$\mu_{core} = \sum_{i=1}^{N} \mu_i \tag{9}$$

where N is the total number of N−H, C═O, and C−N bonds in the structure.

Side-Chain Dipole Moment. pK_a values were calculated by the H++ server for each titratable group of the protein and eqs 4 and 5 were used to assign a partial charge value ranging from −1e to +1e to each residue. The locations of partial charges were approximated using the nitrogen atom location for amine groups and the oxygen atom location for carboxyl groups.

The total dipole moment is the vector sum of the core and side chain dipole moments. Figure 1 shows the point charges and the total dipole moment for β-Lg at pH 4.9 with pK_as calculated with H++.

The overall charge of the protein

$$Q = \sum_{i=1}^{M} q_i \tag{10}$$

like the dipole moment, is both pH and ε_{low} dependent (because of pK_a shifts), making it an additional tool to evaluate ε_{low} of the protein.

MEASUREMENTS

Measurement of the Dipole Moment. Bovine β-Lg lyophilized powder (L3908) containing a combination of variants A and B was obtained from Sigma. Measurements were made with a

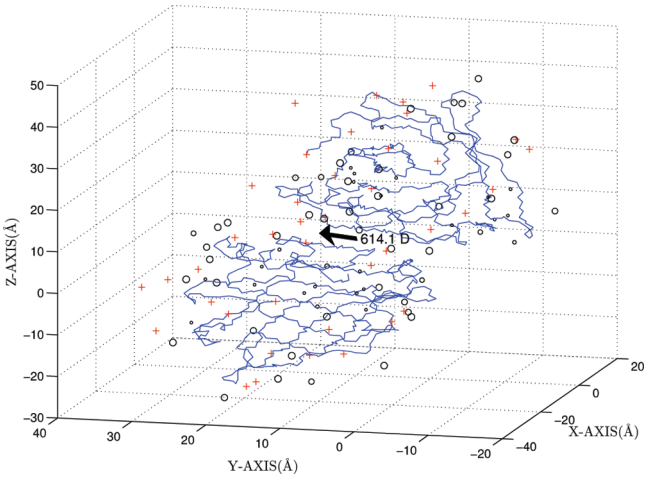


Figure 1. β-Lg (PDB 1beb) partial charges at pH 4.9, near the isoelectric point of the protein, pH 5.1. The vector dipole of 614.1 is shown. Positive charges are displayed as + and negative charges are displayed as circles, where size is proportional to strength. pK_as were calculated with the Poisson−Boltzmann equation using the H++ server.

dielectric cell, described previously,[52] connected to an Agilent 4294A impedance analyzer. The cell was rinsed with ethanol and DI water and dried between measurements. The water bath was allowed 15 min to stabilize after arriving at a temperature. The cell constant, k, and the parasitic capacitance of the cell, C_p, were found through the formula

$$C = k\varepsilon + C_p \tag{11}$$

using the permittivities of air and water (ε = 1 and ∼80[18]). Sweeps were made over 401 logarithmically spaced points from 40 Hz to 110 MHz with settings of a bandwidth of 5 and an oscillator strength of 500 mV. Measurements of permittivity were fitted to the real part of a single relaxation Cole−Cole curve

$$\varepsilon = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + j(\tau\omega)^{(1-\alpha)}} \tag{12}$$

where ε_∞ is the high-frequency permittivity, $\Delta\varepsilon$ is the change in permittivity, α is the Cole parameter describing the spread of the relaxation,[64] τ is the protein rotational relaxation time, and $j = (-1)^{1/2}$. When α = 0, the real part of this equation is given by

$$\varepsilon' = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + \tau^2\omega^2} \tag{13}$$

The electric dipole moment of the protein is related to $\Delta\varepsilon$ through the Oncley formula

$$\mu = \sqrt{\frac{2Mk_bT\varepsilon_0\delta}{Ng}} \tag{14}$$

where μ is the dipole moment, M is the protein molecular weight in kilodaltons, k_b is the Boltzmann constant, T is the temperature in kelvin, ε_0 is the permittivity of free space, N is Avogadro's number, g is the correlation parameter assumed to be 1 for dilute protein solutions,[65] and $\delta = \lim_{c\to 0}\Delta\varepsilon/c$ is the dielectric increment where c is the molar concentration of the protein. For the dimeric β-Lg, eq 14 can be reduced, assuming g = 1, and

expressed in Debye units as

$$\mu = 36.3\sqrt{T\delta} \tag{15}$$

To evaluate the dielectric increment as solute concentration approaches zero, measurements of $\Delta\varepsilon$ and c were fitted to a second-order interpolation equation of the form

$$\Delta\varepsilon = k_2c^2 + k_1c + k_0 \tag{16}$$

where k_1 is the value used for $\Delta\varepsilon/c$ extrapolated to infinite dilution (for eq 14). The k_2c^2 term accounts for the decrease in permittivity at high concentrations and is a measure of the protein−protein interaction caused by interfering electric fields of neighboring β-Lg molecules.[54] Thus, k_2 typically takes on a negative value, as interference of molecular rotation reduces permittivity.

Assuming the protein is roughly spherical, the effective hydrodynamic radius of the protein is estimated from the relaxation time taken from the fitted permittivity spectrum

$$\tau = \frac{4\pi\eta a^3}{k_bT} \tag{17}$$

where a is the effective hydrodynamic radius and η is the viscosity of the solvent.[41]

To ensure accurate parametrization of both the dipole moment and relaxation time, the small dispersion (δ-dispersion) around 100 MHz was accounted for by applying an upper cutoff frequency of 15 MHz to the fitting function. The δ-dispersion is in reality composed of 2^{66} or 3^{67} small dispersions, the lowest of which is around 100 MHz. By applying an upper cutoff frequency well below each of these, the dielectric spectrum considered contains only the dispersion of interest which reflects protein tumbling (β-dispersion). Additionally, when the upper cutoff frequency was moved from 15 to 110 MHz, the fitted dipole moments were nearly unaffected (<3% difference at pH 7). This small difference between fits suggests that the magnitude of the δ-dispersion is small compared to the large β-dispersion of β-Lg and any errors resulting from overlapping modes are negligible.

The pH of the solution was measured by inserting a 3 mm diameter Accumet MicroProbe into the cell cavity immediately after data collection. The pH of the solution was modified by pipeting small volumes of 1 M HCl and 1 M NaOH, thus maintaining the concentration of protein essentially constant. As electrode polarization is always an important consideration in electrical measurements of liquids,[68] the ionic strength of the solution was minimized by reconstituting the protein in DI water and subsequently adding either HCl or NaOH instead of using buffer solutions to achieve a desired pH.

RESULTS

Relaxation data for β-Lg is shown in Figure 2 along with least-squares fits of $\Delta\varepsilon$ (eq 12) to the equilibrium frequency sweeps at each of 3 β-Lg concentrations. This data was recorded using the titration capabilities of our apparatus. Every 20 min, 80 μL of 20 mg/mL β-Lg was pipeted into the 0.8 mL volume cell, which was initially filled with deionized water. Before each addition, 80 μL was also removed to keep the overall volume constant. Visible at each addition are permittivity contributions from the protein and effects of electrode polarization at low frequencies.

Relaxation Data of β-Lg in 0.1 mM HCl. Figures 3, 4, and 5 show measurements obtained in a 0.1 mM HCl solution. In

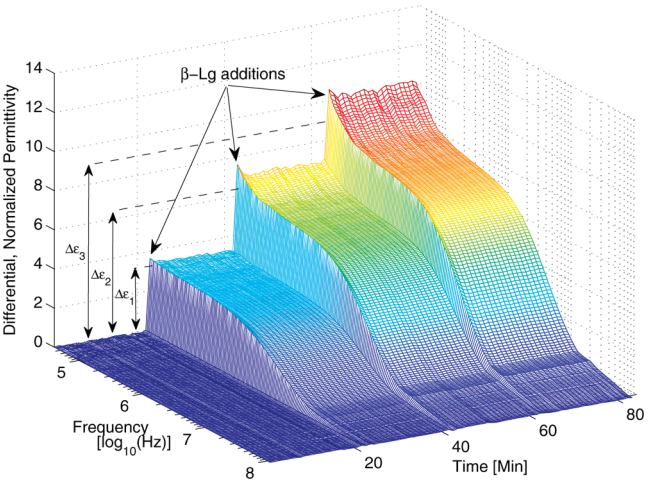


Figure 2. Plot of dielectric titration of β-Lg at 25 °C measured by dielectric spectroscopy. Differential, normalized permittivity was taken at multiple frequencies and frequency sweeps were taken at regular time intervals. $\Delta\varepsilon_1$, $\Delta\varepsilon_2$, and $\Delta\varepsilon_3$ represent least-squares fits of $\Delta\varepsilon$ (eq 12) at 3 concentrations of β-Lg: 2 mg/mL, ~4 mg/mL, and ~6 mg/mL. The pH of the solution was ~6.5. The relaxation time τ was ~40 ns, which increased slightly at each β-Lg addition.

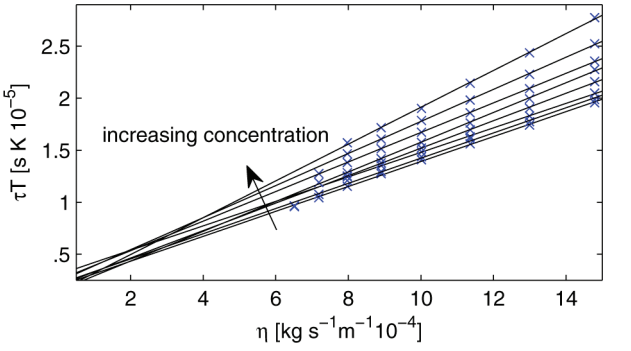


Figure 3. Relaxation time τ (eq 17) multiplied by temperature T against estimated solvent viscosity η for β-Lg in 0.1 mM HCl. Measured values (×) are plotted with linear least-squares fits (solid lines). Concentrations of β-Lg used were 1, 2, 3, 6, 8, 10, 15, and 20 mg/mL.

Figure 3 the relationship between (relaxation time × temperature) and solvent viscosity was linear for all concentrations as predicted by eq 17. Values for solvent viscosity were taken from ref 69. Figure 4 likewise shows a linear relationship between the change in permittivity and temperature as predicted by eq 14. The increasing slope of the linear fits reflects protein−protein effects at increasing concentration.

Figure 5 shows the nonlinear relationship between change in permittivity and concentration at various temperatures. Using eq 16, the value for k_2 of β-Lg in 0.1 mM HCl at 25 °C was found to be −0.010. This value, along with the value found for k_0 (~0), was substituted into eq 16 and rearranged to form the equation

$$\delta = k_1 = \frac{\Delta\varepsilon}{c} + 0.01c \tag{18}$$

Combining eqs 15 and 18 and assuming a constant temperature of 25 °C in the dielectric cell yields

$$\mu = 626.8\sqrt{\frac{\Delta\varepsilon}{c} + 0.01c} \tag{19}$$

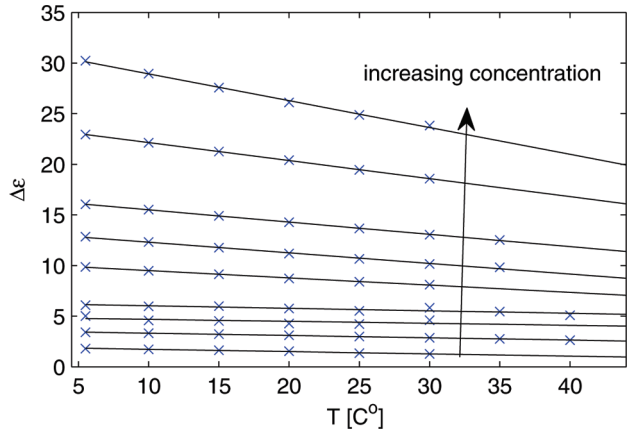


Figure 4. Change in permittivity $\Delta\varepsilon$ (eq 14) against temperature T for β-Lg in 0.1 mM HCl. Measured values ($\times$) are plotted with linear least-squares fits (solid lines). Concentrations of β-Lg used were 1, 2, 3, 4, 6, 8, 10, 15, and 20 mg/mL.

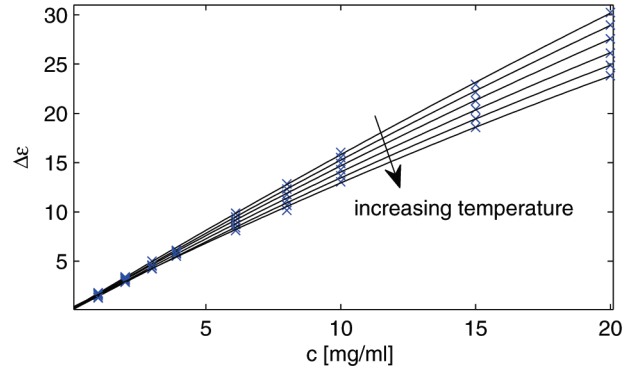


Figure 5. Change in permittivity $\Delta\varepsilon$ (eq 14) against concentration c for β-Lg in 0.1 mM HCl. Measured values ($\times$) are plotted with least-squares fits of second-order polynomials (solid curves, eq 16). Temperature was adjusted from 10 to 35 °C in 5 °C increments.

In this fashion, a single spectral measurement of β-Lg at a known concentration was sufficient to compute an electrical dipole moment μ at a given pH.

Relaxation Data of β-Lg at Variable pH. Estimates for a at various pH were made using eq 17 and are shown in Figure 6a. Values are in the range of ~24−30 Å. The changes in apparent radius at different pHs may be due to structural fluctuations and changes in the thin layer of bound water associated with the macromolecule. Larger fluctuations in radii of ±5 Å were observed by Bonincontro et al., who measured the relaxation parameters of lysozyme at various pH.[49] Using the equivalent volume sphere method described by Grant et al.,[70] the predicted radius of β-Lg from 1beb is ~24.6 Å, in good agreement with measured radii.

The β-Lg dimer is a prolate ellipsoid with an axial ratio of approximately 2. Thus, two different relaxation times with a ratio of 1.3 between them as predicted by Perrin's equations should exist.[70] The apparent radius of the molecule depends on the alignment of the dipole moment with each axis of rotation, e.g., if the dipole moment vector is parallel to the longer axis, the apparent radius would be larger than the apparent radius if it were aligned with the shorter axis. To determine if the observed fluctuations in radius were a result of the dipole moment vector changing alignment, the angle θ formed by the dipole moment vector and the long axis of β-Lg was calculated (Figure 6b). At all pHs, the dipole moment remains relatively aligned with the short axis (within 10°), and the changes in θ are unable to account for changes in a greater than 1 Å. Thus, the observed fluctuations in hydrodynamic radii in Figure 6a must additionally be influenced by a more complex process than changes in dipole alignment.

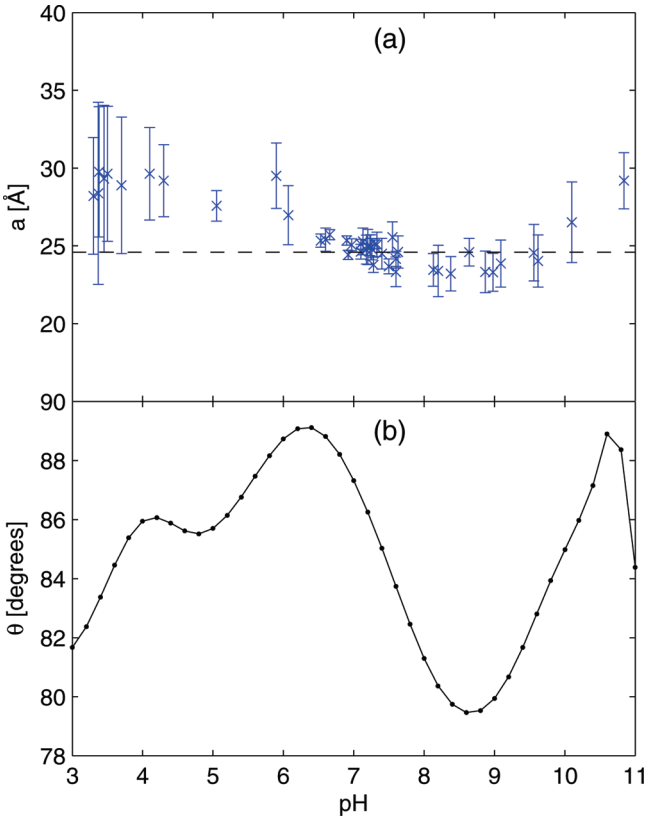


Figure 6. (a) Hydrodynamic radii a (eq 17) of β-Lg at various pH. The error bars represent the minimum and maximum values of several different least-squares fits of the measured data. The horizontal dashed line represents the predicted radius of ~24.6 Å using the equivalent volume sphere method described by Grant et al.[70] (b) Calculated angle θ formed by the dipole moment vector and the long axis of β-Lg. pK_as were calculated from the H++ server. The dipole moment vector remains relatively aligned with the shorter axis at all pHs considered.

Calculated dipole moments of β-Lg comparing Poisson−Boltzmann predictions based on 1beb and measurements based on eq 19 are shown in Figure 7. Predictions vary dramatically depending on ε_{low} as shown by the dashed curves. The agreement is quite good at nearly all pH for the highest values of ε_{low} considered. For subsequent fits, only data from pH 4 to 10 is used as β-Lg dissociates to a monomer from pH 4 to 3.[71]

Estimation of ε_{low} for β-Lg. To quantitate the goodness of fit of the measurements to the Poisson−Boltzmann model, we use the rms residual

$$R_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mu_{calc_i} - \mu_{meas_i})^2} \qquad (20)$$

where N is the number of comparison points. The vector μ_{meas} is obtained using a polynomial interpolation fitting of the measured data. We minimize R_{rms} to identify the value of ε_{low} used in the dielectric continuum model that best approximates the actual dielectric properties of the protein interior.

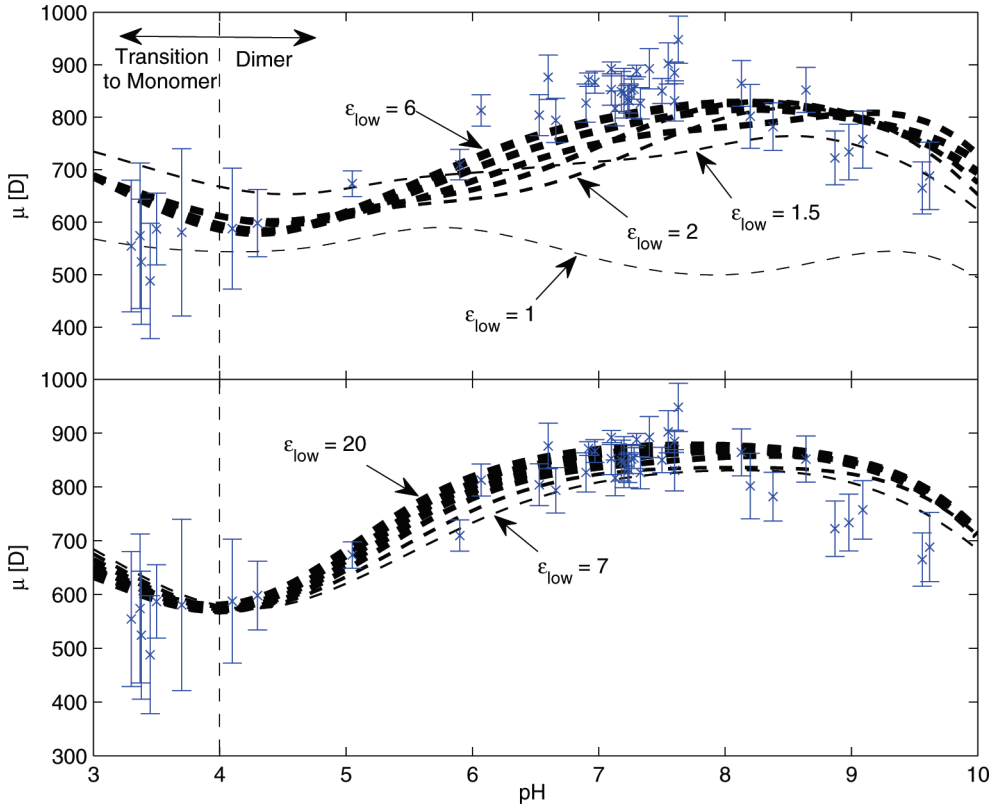


Figure 7. Measured and calculated dimer dipole moments μ of β-Lg at various pH for integral values of ε_{low} from 1 to 20 (and 1.5) using 1beb. Calculated μ are split into two plots (a, upper) and (b, lower) for easier viewing. Error bars represent the minimum and maximum values of several different least-squares fittings of the measured data. Between pH 3 and 4, β-Lg dissociates to a monomer.[71] Calculations appear to converge to the measurements as ε_{low} approaches 6 from below.

In addition to 1beb, two PDB models of dimeric β-Lg, 2akq and 2q39, and two monomeric structures, 1cj5 and 1dv9, are used for comparison. 2akq is the U′ lattice grown from β-Lg crystals at very low ionic strength. This model exhibits the slight structural changes that occur in the almost complete absence of salt. 2q39 is a β-Lg model grown from crystals at low humidity and exhibits an asymmetric dimer. None of these four models were expected to resemble the protein's conformation in the experimental conditions, but were included to display differences in charge moments of structural variations.

The rms residuals (eq 20) for Poisson−Boltzmann predictions for values of ε_{low} from 2 to 20 using the 5 PDB models of β-Lg in relation to the measured data are shown in Figure 8, a and b. The agreement for 1beb to measurements is quite remarkable for all $\varepsilon_{low} > 4$. An R_{rms} value of 35 D corresponds to an average error of 5%. Using the absolute minimum of the curve in Figure 8b, the best estimate of ε_{low} for β-Lg based on 1beb is between 6 and 7.

A similar estimation of ε_{low} is made by comparing the computed and measured isoelectric point of the protein, which was previously measured to be pH 5.1.[76] The calculated value is the pH at which the overall charge Q from eq 10 is 0. The residual value used here, R, is the difference between the measured and calculated pH values for the isoelectric point

$$R = |pI_{calc} - pI_{meas}| \quad (21)$$

Values of R for the five PDB models of β-Lg are plotted in Figure 8, c and d. Unlike the dipole moment function, from which a confident rejection and selection of PDB models could be made, the isoelectric point residual, R, is more useful as a confirmation of the estimate for ε_{low}. This is because the total charge (at least for the dimeric structure) is not as unique as an array of pH-dependent dipole moments; many proteins share the same isoelectric point. Figure 8, b and d, shows good agreement between the two methods for estimating ε_{low} for β-Lg.

DISCUSSION

The accuracy to which the pH-varying dipole moment of β-Lg was predicted in Figure 7 appears to be a significant step in the literature. To our knowledge, such agreement is not encountered elsewhere, for any molecule. South and Grant made predictions for the pH-varying dipole moments of horse and whale myoglobin from pH 5 to 8.[50] In that work, the error between measurements and calculations could likely have been decreased if pK_as were calculated and the dielectric increment δ was fitted to measurements using a wider range of protein concentrations.

The apparent convergence of calculations and measurements in Figure 7 as ε_{low} approaches 6 both helps to validate the electrostatic model used and elucidates the dielectric properties of the protein interior. The curves representing $\varepsilon_{low} < 4$ are well separated from measured values and indicate that the dielectric constant of the region must be higher.

Solvent Penetration as a Source of ε_{low}. Soon after the discovery of internal water molecules in solved crystal structures, it was suggested that water penetration could be a mechanism by which the effective dielectric constant of the protein could increase.[77] If indeed this is the case, the dielectric constant would be sensitive to breathing motions of the molecule, and the

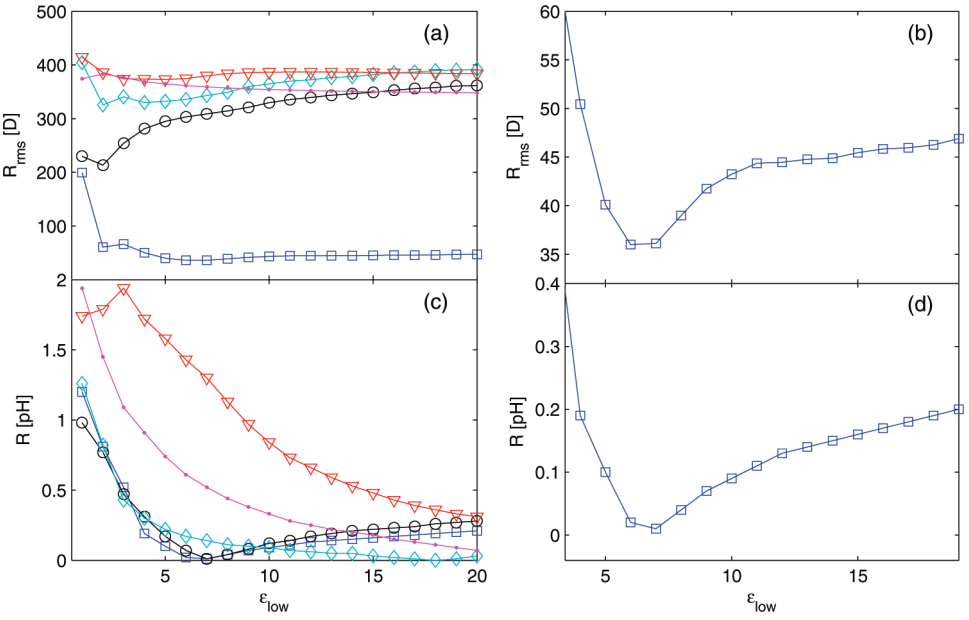


Figure 8. Dipole moment (a,b) and isoelectric point (c,d) residuals for 5 pdb models of β-Lg. PDB structures used are dimers (□) 1beb,[72] (○) 2akq,[73] (◇) 2q39,[74] and monomers (△) 1cj5,[75] and (•)1dv9.[71] Data in (a) shows good compability of the measurements with 1beb only, evident by the small R_{rms} for that structure alone. (b) and (d) are expanded views of 1beb showing good agreement between minimum function values at approximately ε_{low} = 6 to 7 for both R_{rms} and R.

variations in hydrodynamic radius estimates with pH (Figure 6a) would argue against the assumption of a constant ε_{low} at all pH. Calculations using ε_{low} estimates would thus be restricted to a pH range where stable structural conditions are found. This may explain the greater deviations in Figure 7 at pH 8−10.

If ε_{low} were a function of solvent penetration, as was previously suggested,[37] we could expect to see some correlation between a and ε_{low} for a given molecule. From Figure 7, ε_{low} is best approximated by low values ($\varepsilon_{low} < 6$) at high pH (pH > 8) and high values ($\varepsilon_{low} > 6$) at physiological pH (6 < pH < 8), opposite to the expectation from the hydrodynamic radius a which is large at high pH and small at physiological pH. Thus, the data appears to contradict the solvent penetration hypothesis. However, the radius a itself may not be strongly correlated with solvent penetration. Another factor such as the hydrophobicity of penetrable cavities may have a larger impact.

Implications for the Mechanism of the β-Dispersion in Protein Solutions. The β-dispersion in the dielectric spectrum of proteins in solution is the dispersion characterized by change in solution permittivity, $\Delta\varepsilon$, and the relaxation time, τ, typically occurring in the high kHz to MHz frequency range. Researchers have been measuring this dispersion for over 70 years. In the 1930s and 1940s, Oncley made dielectric measurements on several proteins, including β-Lg, and interpreted these results on the basis of the rotation of molecules having a permanent dipole moment.[45,46] Since then, other theories have been proposed to account for this dispersion, such as proton fluctuation,[78,79] Maxwell−Wagner mechanism,[80,81] and the ion-mobility model.[82] The two theories that are still given attention today are permanent dipole rotation and proton fluctuation.

The proton fluctuation theory holds that, when the pH of the solution is close to the pK_a of any group, protons continually bind and disassociate, causing the dipole moment to fluctuate about an equilibrium. Because the change in solution permittivity $\Delta\varepsilon$ is proportional to μ^2, it would be possible to obtain a dispersion even when the average dipole moment is zero. The theory predicts that the dipole moment is a monotonically decreasing function of pH, dropping off from a maximum at pH 2 to effectively zero at pH 10.[78] Figure 7 is in obvious contradiction to this. The proton fluctuation model was not accounted for explicitly in this work, neither did we consider Maxwell−Wagner and ion-mobility models. Nonetheless, the agreement in Figure 7 was achieved solely with the permanent dipole theory and eq 14; this indicates that to account for $\Delta\varepsilon$ of β-Lg with 5% error, the permanent dipole model is sufficient.

Refinement of PDB Models Using Molecular Dynamics. Analysis of protein structures obtained through X-ray crystallography is complicated by possible structural differences between frozen protein crystal and dissolved liquid states.[16] Additionally, the influence of temperature and ionic strength on the protein structure may not be negligible. Therefore, we employed MD simulations in CHARMM to equilibrate structures in aqueous solution.

MD simulations were performed by placing a water cube with side lengths of 10 nm around the molecule using default side-chain charges (corresponding to the intrinsic pK_a at pH 7), deleting overlapping solvent molecules, and heating the system to 25 °C. Before heating, energy minimizations were performed to remove bad contacts between atoms which could pose problems during the heating phase. The heating phase consisted of 3 ps of linear heating from −275 to 25 °C, followed by an equilibration phase lasting 5 ns at constant pressure and temperature. After heating and equilibration, the resulting protein structure was extracted for charge moment calculation using pK_as calculated from the H++ server.

Figure 9 shows computed and measured values of μ comparing R_{rms} for refined and unrefined β-Lg structures. In this case, use of the refined PDB model improved agreement with theory by ~2%. This improvement emphasizes the important role MD simulations could play in developing accurate electrostatic

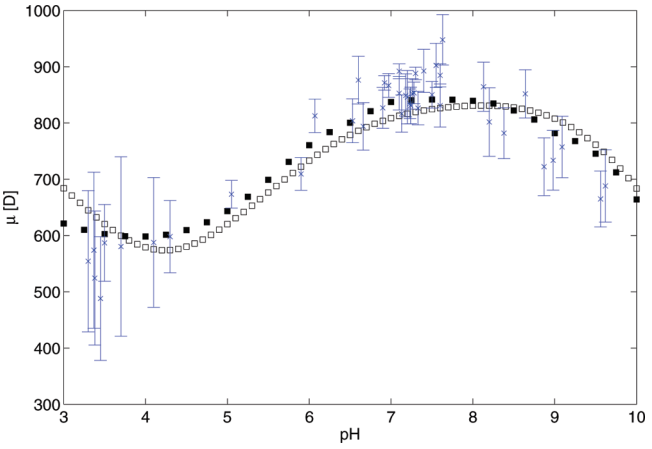


Figure 9. Measured and calculated dipole moments of refined (■) and unrefined (□) β-Lg structures at various pH. Both structures were calculated using an ε_{low} of 6. Use of the refined PDB model improved agreement with theory by ∼2%.

models of proteins from PDB models. Additionally, MD of the PDB model reduced differences between theory and measurements at nearly all pH considered. We note that it is possible the improved agreement is because the refined structure achieved via MD simply represents a better average structure, and that the improvement has nothing to do with MD itself. Further relaxations of the structure with pH-appropriate titration states could be tested for additional agreement of calculations with measurements, however this was not done in this work.

CONCLUSIONS

- Calculated charge moments using the continuum dielectric model were combined with measurements from DS to estimate ε_{low} of β-Lg to be in the range of 6−7. This value is higher than measurements of protein powders ($\varepsilon \approx 2.5$[33]), but lower than ε_{eff} of buried ionizable residues ($\varepsilon \approx 10-12$[36]), and is likely due to a combination of electronic polarization and orientational relaxation of a small amount of penetrating solvent.
- The pH-dependent β-dispersion of β-Lg was predicted with an error of 5% over the range of pH 4−10. This result validates the mechanism of permanent dipolar relaxation in proteins, charge models using Henderson−Hasselbalch equations, and pK_a shifts calculated with the Poisson−Boltzmann equation, such as with the H++ server.
- Refinement of the 1beb structure using CHARMM MD simulations increased the consistency of dipole moment measurements with theory by ∼2%. This result raises new possibilites of using MD to supplement theoretically derived charge moments of molecules in solution.

AUTHOR INFORMATION

Corresponding Author

*E-mail: brettmellor@byu.edu.

ACKNOWLEDGMENT

Thanks to the Brigham Young University Fulton College of Engineering for funds to support this work.

REFERENCES

(1) Hassan, S. A.; Mehler, E. L. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 45–61.
(2) Li, X. F.; Hassan, S. A.; Mehler, E. L. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 464–484.
(3) Mandell, J. G.; Roberts, V. A.; Pique, M. E.; Kotlovyi, V.; Mitchell, J. C.; Nelson, E.; Tsigelny, I.; Ten Eyck, L. F. *Protein Eng.* **2001**, *14*, 105–113.
(4) Hofinger, S.; Simonson, T. *J. Comput. Chem.* **2001**, *22*, 290–305.
(5) Gunner, M. R.; Mao, J. J.; Song, Y. F.; Kim, J. *Biochim. Biophys. Acta: Bioenerg.* **2006**, *1757*, 942–968.
(6) Simonson, T. *Photosynth. Res.* **2008**, *97*, 21–32.
(7) Hao, Y.; Pear, M. R.; Busath, D. D. *Biophys. J.* **1997**, *73*, 1699–1716.
(8) Parsegian, A. *Nature* **1969**, *221*, 844–846.
(9) Boda, D.; Busath, D. D.; Henderson, D.; Sokolowski, S. *J. Phys. Chem. B* **2000**, *104*, 8903–8910.
(10) Vora, T.; Corry, B.; Chung, S. H. *Eur. Biophys. J. Biophys. Lett.* **2008**, *38*, 45–52.
(11) Song, C.; Corry, B. *Biophys. J.* **2010**, *98*, 404–411.
(12) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219–10225.
(13) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. *Biophys. J.* **2002**, *83*, 1731–1748.
(14) Baran, K. L.; Chimenti, M. S.; Schlessman, J. L.; Fitch, C. A.; Herbst, K. J.; Garcia-Moreno, B. E. *J. Mol. Biol.* **2008**, *379*, 1045–1062.
(15) Castaneda, C. A.; Fitch, C. A.; Majumdar, A.; Khangulov, V.; Schlessman, J. L.; Garcia-Moreno, B. E. *Proteins: Struct., Funct., Bioinf.* **2009**, *77*, 570–588.
(16) Takashima, S. *Biopolymers* **2001**, *58*, 398–409.
(17) Bashford, D. *Sci. Comput. Object-Oriented Parallel Environ.* **1997**, *1343*, 233–240.
(18) Ellison, W. J.; Lamkaouchi, K.; Moreau, J. M. *J. Mol. Liq.* **1996**, *68*, 171–279.
(19) Matthew, J. B.; Gurd, F. R. N.; Garciamoreno, E. B.; Flanagan, M. A.; March, K. L.; Shire, S. J. *CRC Crit. Rev. Biochem.* **1985**, *18*, 91–197.
(20) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301–332.
(21) Tredgold, R. H.; Hole, P. N. *Biochim. Biophys. Acta* **1976**, *443*, 137–142.
(22) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097–2119.
(23) Jean-Charles, A.; Nicholls, A.; Sharp, K.; Honig, B.; Tempczyk, A.; Hendrickson, T. F.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 1454–1455.
(24) Rupley, J. A.; Careri, G. *Adv. Protein Chem.* **1991**, *41*, 37–172.
(25) Simonson, T.; Perahia, D.; Brunger, A. T. *Biophys. J.* **1991**, *59*, 670–690.
(26) Simonson, T.; Perahia, D.; Bricogne, G. *J. Mol. Biol.* **1991**, *218*, 859–886.
(27) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671–679.
(28) Yang, A. S.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. *Proteins: Struct., Funct., Genet.* **1993**, *15*, 252–265.
(29) Oberoi, H.; Allewell, N. M. *Biophys. J.* **1993**, *65*, 48–55.
(30) Kundu, S.; Gupta-Bhaya, P. *J. Mol. Struct.—THEOCHEM* **2003**, *639*, 21–26.
(31) Nakamura, H.; Sakamoto, T.; Wada, A. *Protein Eng.* **1988**, *2*, 177–183.
(32) Pennock, B.; Schwan, H. *J. Phys. Chem.* **1969**, *73*, 2600–2610.
(33) Takashima, S.; Schwan, H. P. *J. Phys. Chem.* **1965**, *69*, 4176–4182.
(34) Antosiewicz, J.; Mccammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415–436.
(35) Demchuk, E.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 17373–17387.
(36) GarciaMoreno, B.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E. *Biophys. Chem.* **1997**, *64*, 211–224.
(37) Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S.; Stites, W. E.; Garcia-Moreno, B. *Biophys. J.* **2000**, *79*, 1610–1620.

(38) Karp, D. A.; Gittis, A. G.; Stahley, M. R.; Fitch, C. A.; Stites, W. E.; Garcia-Moreno, B. *Biophys. J.* **2007**, *92*, 2041–2053.
(39) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452–8458.
(40) Loffler, G.; Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1997**, *270*, 520–534.
(41) Debye, P. *Polar Molecules*; Chemical Catalog Co., Inc.: New York, 1929.
(42) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351–361.
(43) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.
(44) Wyman, J. *J. Am. Chem. Soc.* **1936**, *58*, 1482–1486.
(45) Oncley, J. L. *J. Phys. Chem.* **1940**, *44*, 1103–1113.
(46) Oncley, J. L. In *The Electric Moments and the Relaxation Times of Proteins as Measured from Their In_uence upon the Dielectric Constants of Solutions*; Edsall, C., Ed.; Reinhold Publishing Corp.: New York, 1943; Chapter 22, pp 543−567.
(47) Bonincontro, A.; Cinelli, S.; Onori, G.; Stravato, A. *Biophys. J.* **2004**, *86*, 1118–1123.
(48) Keefe, S. E.; Grant, E. H. *Phys. Med. Biol.* **1974**, *19*, 701–707.
(49) Bonincontro, A.; De Francesco, A.; Onori, G. *Colloids Surf. B—Biointerfaces* **1998**, *12*, 1–5.
(50) South, G. P.; Grant, E. H. *Proc. R. Soc. London. Ser. A, Math. Phys. Sci.* **1972**, *328*, 371–387.
(51) Nigen, M.; Croguennec, T.; Renard, D.; Bouhallab, S. *Biochemistry* **2007**, *46*, 1248–1255.
(52) Mazzeo, B. A.; Chandra, S.; Mellor, B. L.; Arellano, J. *Rev. Sci. Instrum.* **2010**, *81*, 125103.
(53) Valkonen, K. H.; Marttinen, N.; Alatossava, T. *Bioseparation* **2002**, *10*, 145–152.
(54) Ferry, J. D.; Oncley, J. L. *J. Am. Chem. Soc.* **1941**, *63*, 272–278.
(55) Rogers, S. S.; Venema, P.; Van der Ploeg, J. P. M.; Van der Linden, E.; Sagis, L. M. C.; Donald, A. M. *Biopolymers* **2006**, *82*, 241–252.
(56) Takashima, S.; Asami, K. *Biopolymers* **1993**, *33*, 59–68.
(57) Harvey, S. C.; Garcia de la Torre, J. *Macromolecules* **1980**, *13*, 960–964.
(58) Antosiewicz, J. *Biophys. J.* **1995**, *69*, 1344–1354.
(59) Hasselbalch, K. A. *Biochemistry* **1916**, *78*, 112–144.
(60) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473–486.
(61) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144–1149.
(62) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368–W371.
(63) Smyth, C. P. *Dielectric behavior and structure: dielectric constant and loss, dipole moment and molecular structure*; McGraw-Hill: New York, 1955.
(64) Cole, K. S.; Cole, R. H. *J. Chem. Phys.* **1941**, *9*, 341–351.
(65) Pethig, R. *Dielectric and Electronic Properties of Biological Materials*; John Wiley & Sons Ltd.: New York, 1979.
(66) Pethig, R. *Annu. Rev. Phys. Chem.* **1992**, *43*, 177–205.
(67) Oleinikova, A.; Sasisanker, P.; Weingartner, H. *J. Phys. Chem. B* **2004**, *108*, 8467–8474.
(68) Scheider, W. *J. Phys. Chem.* **1975**, *79*, 127–136.
(69) Korson, L.; Drost-Hansen, W.; Millero, F. J. *J. Phys. Chem.* **1969**, *73*, 34–39.
(70) Grant, E. H.; Sheppard, R. J.; South, G. P. *Dielectric Behaviour of Biological Molecules in Solution*; Oxford University Press: Oxford, UK, 1978.
(71) Uhrinova, S.; Smith, M. H.; Jameson, G. B.; Uhrin, D.; Sawyer, L.; Barlow, P. N. *Biochemistry* **2000**, *39*, 3565–3574.
(72) Brownlow, S.; Cabral, J. H. M.; Cooper, R.; Flower, D. R.; Yewdall, S. J.; Polikarpov, I.; North, A. C. T.; Sawyer, L. *Structure* **1997**, *5*, 481–495.
(73) Adams, J. J.; Anderson, B. F.; Norris, G. E.; Creamer, L. K.; Jameson, G. B. *J. Struct. Biol.* **2006**, *154*, 246–254.
(74) Vijayalakshmi, L.; Krishna, R.; Sankaranarayanan, R.; Vijayan, M. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 241–249.
(75) Kuwata, K.; Hoshino, M.; Forge, V.; Era, S.; Batt, C. A.; Goto, Y. *Protein Sci.* **1999**, *8*, 2541–2545.
(76) MacKenzie, H. *Milk proteins: Chemistry and molecular biology*; Academic Press: New York, 1971; Vol. 2.
(77) Birktoft, J. J.; Blow, D. M. *J. Mol. Biol.* **1972**, *68*, 187–240.
(78) Kirkwood, J. G.; Shumaker, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **1952**, *38*, 855–862.
(79) Takashima, S. *J. Phys. Chem.* **1965**, *69*, 2281–2286.
(80) Wagner, K. W. *Arch. Elektrotech.* **1914**, *2*, 371–387.
(81) Fricke, H. *J. Phys. Chem.* **1953**, *57*, 934–937.
(82) Schwarz, G. *J. Phys. Chem.* **1962**, *66*, 2636–2642.