PHYSICAL REVIEW E **75**, 021916 (2007)

Cubic equation governing the outer-region dielectric constant of globular proteins

Hwangseo Park[1,*] and Young Ho Jeon[2,†]
[1]*Department of Bioscience and Biotechnology, Sejong University, 98 Kunja-Dong, Kwangjin-Ku, Seoul 143-747, Korea*
[2]*The Magnetic Resonance Team, Korea Basic Science Institute, 804-1 Yangchung-Ri, Ochang, Chungbuk 363-883, Korea*
(Received 22 June 2006; published 26 February 2007)

Based on the Fröhlich-Kirkwood theory of dielectrics and homogenization theory, a cubic equation is proposed for the outer-region dielectric constant of a globular protein in aqueous solution. The results for simple test cases confirm that a soluble globular protein may be considered to consist of a hydrophobic core surrounded by a hydrophilic external surface. Considering the widespread debate and ambiguity about the dielectric constant of proteins, the equation is expected to provide a straightforward way to determine the critical parameter for characterizing the relaxation of a protein in response to a charged perturbation.

DOI: 10.1103/PhysRevE.75.021916 PACS number(s): 87.15.Aa, 82.37.Rs

An accurate calculation of the electrostatic energy term is necessary for understanding the intramolecular and intermolecular interactions pertinent to the structure and function of macromolecules, and therefore it has been subject to intense experimental and theoretical scrutiny [1–3]. The electrostatic interaction is often computed using a screened Coulomb potential with varying dielectric constants to represent the bulk electrostatic properties of the macroscopic environment [4–6]. A direct determination of the dielectric constant from atomistic simulation has proven difficult because it measures a slowly convergent ensemble property. However, it was shown that the calculation of the dielectric constant of a bulk solvent based on microscopic simulations could be successful with accuracy enough to account for the macroscopic properties of water [7–9].

In the case of proteins, the calculation of the dielectric constant becomes even more difficult due to the structural inhomogeneity which makes it difficult to cast inherently microscopic properties into a macroscopic framework [10]. Considerable efforts have nonetheless been made to determine the dielectric constants of proteins based on atomistic simulations [11–13], which showed that it would be more informative to divide a protein into a hydrophobic core and an outer region with different dielectric constants. The dielectric constant in the outer region is of greater importance than that of the inner region because the former is directly involved in a variety of biochemical processes including protein-protein interaction and ligand binding. In the present study, a cubic equation is proposed that can be applied to determine the dielectric constant of the outer region of globular proteins in solution.

The dielectric constant is a fundamental electrostatic property of bulk materials and measures the polarization response of a material to the applied electric field. A globular protein molecule in aqueous solution may be viewed as a spherical dielectric medium with a dielectric constant ε_p and radius r_p in a continuum solvent with dielectric constant ε_w that is 78.4 at room temperature. Within the framework of the Fröhlich-Kirkwood theory of dielectrics, ε_p is related to the deviation of the protein dipole moment from its mean (ΔM_p) as follows [14]:

$$\frac{\langle \Delta M_p^2 \rangle}{kTr_p^3} = \frac{(2\varepsilon_w + 1)(\varepsilon_p - 1)}{2\varepsilon_w + \varepsilon_p}. \tag{1}$$

Here, the angular brackets represent an ensemble average. ε_p can be determined because the total dipole moment fluctuation is available from atomistic simulations in aqueous solution.

However, it is well appreciated that the dielectric constant changes to a significant extent in going from the center of a protein to the outer region, due mainly to the variation of hydrophilicity [11]. As shown in Fig. 1, therefore, it seems to be more reasonable to represent a globular protein as a union of two distinct spherical regions: the inner region with radius r_1 and dielectric constant ε_1, and the outer region with radius $r_2 = r_p$ and dielectric constant ε_2. Then, a fluctuation formula similar to Eq. (1) can be obtained in a straightforward way [15]:

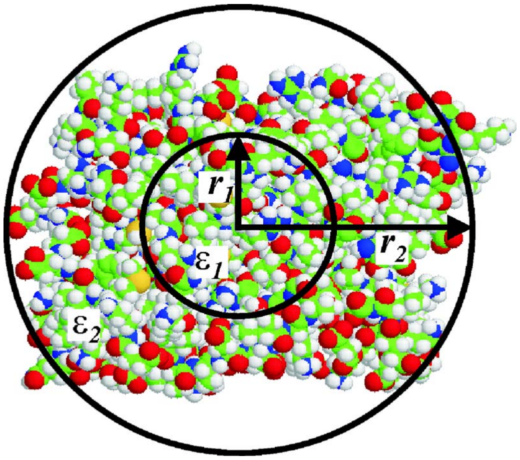

FIG. 1. (Color online) An approximate globular protein comprising two concentric dielectric regions.

---

[*] Email address: hspark@sejong.ac.kr

[†] Email address: yhjeon@kbsi.re.kr

$$\frac{\langle \Delta M_1^2\rangle}{kTr_1^3}=\frac{(\varepsilon_1-1)\left[(1+2\varepsilon_2)(2\varepsilon_w+\varepsilon_2)-2\left(\frac{r_1}{r_2}\right)^3(\varepsilon_w-\varepsilon_2)(1-\varepsilon_2)\right]}{(\varepsilon_1+2\varepsilon_2)(2\varepsilon_w+\varepsilon_2)-2\left(\frac{r_1}{r_2}\right)^3(\varepsilon_w-\varepsilon_2)(\varepsilon_1-\varepsilon_2)}. \tag{2}$$

Here, ΔM_1 represents the deviation of the dipole moment of the inner region from its average value. Similar to $\langle \Delta M_p^2\rangle$ in Eq. (1), $\langle \Delta M_1^2\rangle$ can also be computed by taking the ensemble average over the trajectory sampled by microscopic simulations. It is worth mentioning that either ε_1 or ε_2 can be selected as the active site dielectric constant depending on whether the active site is located in the center or in the periphery of a protein. Except in a few cases, ε_1 is known to have trivially small values ranging from 1 to 4 due mainly to the rigidity of the nominally hydrophobic inner region of proteins [16,17]. Therefore, the interest is focused on ε_2 in this study on the grounds of the dominant contribution of the polar outer region to the intrinsic polarizability of proteins [18].

In order for ε_1 and ε_2 to be determined, another relation between the two variables should be provided in addition to Eq. (2). For this reason, the overall protein dielectric constant ε_p is assumed to be approximated by a volume-weighted average of ε_1 and ε_2. Then, we have

$$r_2^3\varepsilon_p=r_1^3\varepsilon_1+(r_2^3-r_1^3)\varepsilon_2. \tag{3}$$

This approximation is based on the homogenization theory [19] which has been widely applied to solve a complex two-phase problem in liquid and solid phases. For example, the volume-weighted average of the solubility parameters of various ions in solution was found to be a useful quantity for estimating the overall activity coefficient [20]. It was also shown that the solubility parameter of a complex ion could be taken as the volume-weighted average of those of its constituents [21]. Most relevantly, the volume-weighted average dielectric constant has proved to be a good approximation for the overall dielectric constant of periodic dielectric structures to the extent to be a critical factor governing their spectral properties [22], further supporting the validity of the assumption made in Eq. (3).

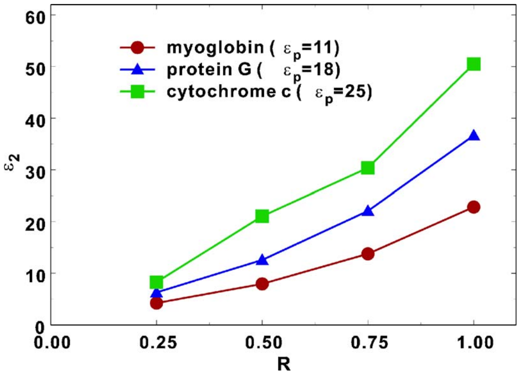

FIG. 2. (Color online) Radial variations of the outer-region dielectric constants of myoglobin, protein G, and cytochrome *c*.

Equations (1)–(3) constitute a linear simultaneous equation for ε_p, ε_1, and ε_2, which can be solved by the elimination method. If the variable ε_1 is eliminated by using Eqs. (2) and (3), we can obtain a cubic equation for ε_2 that has the following form (see the Appendix for details of the derivation):

$$\begin{aligned}\varepsilon_2^3&+\frac{R}{2-R}[(R+4)\varepsilon_w+2R-2\varepsilon_p-K+1]\varepsilon_2^2\\&+R\biggl(1-\frac{R}{(2-R)^2}[2(R+4)\varepsilon_w\varepsilon_p+2(R+1)(K-1)\varepsilon_p\\&+4K(2R-1)\varepsilon_w+R(R+4)\varepsilon_w+R(R+1)]\biggr)\varepsilon_1\\&+\frac{R\varepsilon_w}{2-R}[2(K-1)\varepsilon_p+R]=0,\end{aligned} \tag{4}$$

where $k=\langle \Delta M_1^2\rangle/kTr_1^3$ and $R=2(r_1/r_2)^3$.

Here, the parameter R is proportional to the cube of the ratio of the radius of the inner region to the radius of the whole protein. Therefore, it increases with increase in the thickness of the inner region, and with decrease in that of the outer shell. For a protein that has a hydrophobic core with negligibly small dipole fluctuation ($K\approx 0$), it follows immediately that the quadratic coefficient is positive in Eq. (4) while the linear and constant coefficients are negative. In this case, only one acceptable positive root is obtained, with the remaining two roots being negative or imaginary.

The practical application of Eq. (4) can be illustrated through a simple case study of calculating the outer-region dielectric constant. For simplicity, let us assume that the fluctuation of the inner-region dipole moment is negligibly small such that the parameter K may be set equal to zero. Equation (4) can then be solved with respect to the two parameters ε_p and R. Myoglobin, protein G, and cytochrome *c* are selected in this case study because their whole dielectric constants were reported in an earlier study [11]. Equation (4) is thus solved for the three proteins to obtain their outer-region dielectric constants with respect to four values of R, and the results are illustrated in Fig. 2.

It is seen that at all values of R the calculated ε_2's increase in the same order as the overall dielectric constants of the three proteins under study. This is not surprising because the protein dipole fluctuations are usually dominated by the motions of polar residues in the outer region rather than those of the hydrophobic core. More importantly, ε_2 appears to increase with increasing R in all three cases, indicating that ε_2

gets higher as the boundary between the inner and outer regions approaches the protein surface on which the dielectric constant should be close to that of water. This radial dependence of the protein dielectric constant was also observed in earlier studies [23,24], and has been attributed to the relative abundance of polar or charged residues with high polarizability in the outer region [25,26].

In conclusion, we have developed a cubic equation that can be useful for obtaining the dielectric constant of the outer region of a globular protein in solution. The equation is expected to provide a straightforward way to determine the critical parameter for describing the electrostatic interactions involving protein atoms and the relaxation of a protein in response to a charged perturbation.

This work was supported by the fund from the Bio-MR Research Program (granted to Y.H.J.) of the Korea Basic Science Institute.

APPENDIX

In this Appendix, we will detail how to arrive at Eq. (4) in the text. Let $K = \langle \Delta M_1^2 \rangle / k_B T r_1^3$ and $R = 2(r_1/r_2)^3$ as in the text. Then, Eq. (2) in the text becomes

$$K = \frac{(\varepsilon_1 - 1)[(1 + 2\varepsilon_2)(2\varepsilon_w + \varepsilon_2) - R(\varepsilon_w - \varepsilon_2)(1 - \varepsilon_2)]}{(\varepsilon_1 + 2\varepsilon_2)(2\varepsilon_w + \varepsilon_2) - R(\varepsilon_w - \varepsilon_2)(\varepsilon_1 - \varepsilon_2)}.$$

This relation can also be written as

$$\begin{aligned} &K[(2-R)\varepsilon_2^2 + [\varepsilon_1 + 4\varepsilon_w + R(\varepsilon_w + \varepsilon_1)]\varepsilon_2 + (2-R)\varepsilon_w\varepsilon_1] \\ &\quad = (\varepsilon_1 - 1)[(2-R)\varepsilon_2^2 + \{(R+4)\varepsilon_w + R + 1\}\varepsilon_2 + (2-R)\varepsilon_w]. \end{aligned} \tag{A1}$$

Equation (3) in the text can be rearranged as follows:

$$\begin{aligned} \varepsilon_1 &= \frac{r_2^3\varepsilon_p - (r_2^3 - r_1^3)\varepsilon_2}{r_1^3} = \left(\frac{r_2}{r_1}\right)^3 \varepsilon_p - \left[\left(\frac{r_2}{r_1}\right)^3 - 1\right]\varepsilon_2 \\ &= \frac{2\varepsilon_p}{R} - \left(\frac{2}{R} - 1\right)\varepsilon_2 \end{aligned}$$

Using this relation, the left-hand side of Eq. (A1) can be rearranged as

$$\begin{aligned} &K[(2-R)\varepsilon_2^2 + \{(R+1)\varepsilon_1 + (R+4)\varepsilon_w\}\varepsilon_2 + (2-R)\varepsilon_w\varepsilon_1] \\ &\quad = K\left[(2-R)\varepsilon_2^2 + \left\{(R+1)\frac{2\varepsilon_p}{R} - (R+1)\left(\frac{2}{R} - 1\right)\varepsilon_2 + (R+4)\varepsilon_w\right\}\varepsilon_2 + (2-R)\frac{2\varepsilon_w\varepsilon_p}{R} - (2-R)\left(\frac{2}{R} - 1\right)\varepsilon_w\varepsilon_2\right] \\ &\quad = K\left[\left\{(2-R) - (R+1)\left(\frac{2}{R} - 1\right)\right\}\varepsilon_2^2 + \left\{(R+1)\frac{2\varepsilon_p}{R} + (R+4)\varepsilon_w - (2-R)\left(\frac{2}{R} - 1\right)\varepsilon_w\right\}\varepsilon_2 + 2\left(\frac{2}{R} - 1\right)\varepsilon_w\varepsilon_p\right] \\ &\quad = K\left[\left(1 - \frac{2}{R}\right)\varepsilon_2^2 + \left\{2\left(1 + \frac{1}{R}\right)\varepsilon_p + 4\left(2 - \frac{1}{R}\right)\varepsilon_w\right\}\varepsilon_2 + 2\left(\frac{2}{R} - 1\right)\varepsilon_w\varepsilon_p\right]. \end{aligned}$$

Then, Eq. (A1) becomes

$$K\left[\left(1 - \frac{2}{R}\right)\varepsilon_2^2 + \left\{2\left(1 + \frac{1}{R}\right)\varepsilon_p + 4\left(2 - \frac{1}{R}\right)\varepsilon_w\right\}\varepsilon_2 + 2\left(\frac{2}{R} - 1\right)\varepsilon_w\varepsilon_p\right] = (\varepsilon_1 - 1)[(2-R)\varepsilon_2^2 + \{(R+4)\varepsilon_w + R + 1\}\varepsilon_2 + (2-R)\varepsilon_w].$$

Further rearrangement of this equation gives

$$\begin{aligned} &\left\{K\left(1 - \frac{2}{R}\right) + 2 - R\right\}\varepsilon_2^2 + \left\{2K\left(1 + \frac{1}{R}\right)\varepsilon_p + 4K\left(2 - \frac{1}{R}\right)\varepsilon_w + (R+4)\varepsilon_w + R + 1\right\}\varepsilon_2 + 2K\left(\frac{2}{R} - 1\right)\varepsilon_w\omega_p + (2-R)\varepsilon_w \\ &\quad = \left\{\frac{2\varepsilon_p}{R} - \left(\frac{2}{R} - 1\right)\varepsilon_2\right\}[(2-R)\varepsilon_2^2 + \{(R+4)\varepsilon_w + R + 1\}\varepsilon_2 + (2-R)\varepsilon_w]. \end{aligned}$$

Rearranging terms, we get

$$\varepsilon_2^3 + \frac{R}{2-R}\{(R+4)\varepsilon_w + 2R - 2\varepsilon_p - K + 1\}\varepsilon_2^2 + \left\{R - \frac{2R(R+4)\varepsilon_w\varepsilon_p}{(2-R)^2} + \frac{2R(R+1)(K-1)\varepsilon_p}{(2-R)^2} + \frac{4KR(2R-1)\varepsilon_w}{(2-R)^2} + \frac{R^2(R+4)\varepsilon_w}{(2-R)^2} + \frac{R^2(R+1)}{(2-R)^2}\right\}\varepsilon_2 + \frac{2R(R-1)\varepsilon_w\varepsilon_p}{2-R} + \frac{R^2\varepsilon_w}{2-R} = 0,$$

or in a more simplified form

$$\varepsilon_2^3 + \frac{R}{2-R}\{(R+4)\varepsilon_w + 2R - 2\varepsilon_p - K + 1\}\varepsilon_2^2 + R\left[1 - \frac{R}{(2-R)^2}\{2(R+4)\varepsilon_w\varepsilon_p + 2(R+1)(K-1)\varepsilon_p + 4K(2R-1)\varepsilon_w + R(R+4)\varepsilon_w + R(R+1)\}\right]\varepsilon_2 + \frac{R\varepsilon_w}{2-R}\{2(K-1)\varepsilon_p + R\} = 0.$$

This is Eq. (4) in the text.

[1] B. Honig and A. Nicholls, Science **268**, 1144 (1995).

[2] B. E. Cohen, T. B. McAnaney, E. S. Park, Y. N. Jan, S. G. Boxer, and L. Y. Jan, Science **296**, 1700 (2002).

[3] T. Simonson, J. Phys. Chem. B **104**, 6509 (2000).

[4] E. L. Mehler and T. Solmajer, Protein Eng. **4**, 903 (1991).

[5] R. Luo, J. Moult, and M. K. Gilson, J. Phys. Chem. B **101**, 11226 (1997).

[6] P. Werner and A. Caflisch, J. Am. Chem. Soc. **125**, 4600 (2003).

[7] X. Ni and R. M. Fine, J. Phys. Chem. **96**, 2718 (1992).

[8] S. Senapati and A. Chandra, J. Phys. Chem. B **105**, 5106 (2001).

[9] M. S. Skaf, J. Phys. Chem. A **103**, 10719 (1999).

[10] A. Warshel and A. Papazyan, Curr. Opin. Struct. Biol. **8**, 211 (1998).

[11] T. Simonson and C. L. Brooks III, J. Am. Chem. Soc. **118**, 8452 (1996).

[12] T. Simonson, J. Am. Chem. Soc. **120**, 4875 (1998).

[13] J. W. Pitera, M. Falta, and W. F. van Gunsteren, Biophys. J. **80**, 2546 (2001).

[14] H. Fröhlich, *Theory of Dielectrics* (Clarendon, Oxford, 1949).

[15] T. Simonson and D. Perahia, Proc. Natl. Acad. Sci. U.S.A. **92**, 1082 (1995).

[16] M. K. Gilson and B. Honig, Biopolymers **25**, 2097 (1986).

[17] K. A. Sharp, Biophys. J. **74**, 1241 (1998).

[18] J. Antosiewicz, J. A. McCammon, and M. K. Gilson, J. Mol. Biol. **238**, 415 (1994).

[19] M. J. Beran, *Statistical Continuum Theories* (Wiley Interscience, New York, 1968).

[20] C. F. Bases, Jr, and B. A. Moyer, J. Phys. Chem. B **101**, 6566 (1997).

[21] T. G. Levitskaia, J. C. Bryan, R. A. Sachleben, J. D. Lamb, and B. A. Moyer, J. Am. Chem. Soc. **122**, 554 (2000).

[22] Z. Jian and D. M. Mittleman, Appl. Phys. Lett. **87**, 191113 (2005).

[23] C. H. Paul, J. Mol. Biol. **155**, 53 (1982).

[24] I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, Proteins **1**, 47 (1986).

[25] J. B. Matthew, Annu. Rev. Biophys. Biophys. Chem. **14**, 387 (1985).

[26] P. E. Smith, R. M. Brunne, A. E. Mark, and W. F. van Gunsteren, J. Phys. Chem. **97**, 2009 (1993).