Dielectric Constant of Cytochrome c from Simulations in a Water Droplet Including All Electrostatic Interactions

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MAY 1998

Impact Factor: 12.11 · DOI: 10.1021/ja980071m

1 AUTHOR:

Thomas Simonson
École Polytechnique

116 PUBLICATIONS **20,641** CITATIONS

SEE PROFILE

Dielectric Constant of Cytochrome *c* from Simulations in a Water Droplet Including All Electrostatic Interactions

Thomas Simonson

Laboratoire de Biologie Structurale, Institut de Génétique et Biologie Moléculaire et Cellulaire (C.N.R.S.), 1 rue Laurent Fries 67404 Strasbourg-Illkirch, France

Received January 7, 1998
Revised Manuscript Received March 23, 1998

The dielectric constant of a protein measures its average polarizability and is the appropriate macroscopic quantity to characterize its relaxation in response to charged perturbations, such as electron or proton transfer, photoexcitation of a bound chromophore, or ion or proton binding. We report here a calculation of the dielectric constant of cytochrome *c* from a 1 ns molecular dynamics simulation in a water droplet surrounded by a vacuum, including all electrostatic interactions. This is the first such calculation for a protein of this size. Results are in remarkably good agreement with a previous droplet calculation that used an electrostatic cutoff[1−3] and with calculations for other proteins using periodic boundary conditions with a cutoff.[4,5] Together, these data provide a general picture of the dielectric constant of globular proteins. The bulk of cytochrome *c* has a very low dielectric constant of 3 ± 1, close to the experimental value for dry protein powders,[6,7] ensuring a low reorganization energy for electron transfer.[8] The overall dielectric constant is 25 ± 10, arising almost entirely from charged side chains at the protein surface and typical of all proteins studied so far. This suggests that the protein contribution to dielectric relaxation processes near the protein surface (e.g., proton self-energies) can be considerably larger than that of a simple apolar medium, as argued in the past by Warshel and co-workers.[9,10]

A calculation for an 18-residue zinc-finger peptide was described recently that also included all electrostatic interactions, combining periodic boundary conditions with Ewald summation.[11] For spatially homogeneous systems with these boundary conditions, fluctuation formulas for the dielectric constant were provided by the seminal work of Neumann and Steinhauser.[12] However, for an inhomogeneous protein−solvent mixture, no completely rigorous formula connects the protein fluctuations and its dielectric constant. An approximate formula was derived in ref 11 in a way that depends on protein concentration. Heuristic arguments suggest that this introduces a significant systematic error. However, a simple, generally applicable correction scheme for the concentration-dependence is derived below. When applied to the data of ref 11, the correction reduces the protein dielectric constant from 15 to 11.

We emphasize, as previously,[4] that the dielectric constant calculated here is a linear response coefficient, which characterizes the protein *polarizability*, i.e., its relaxation in response to perturbing charges. It is not related in a simple way to the equilibrium *polarity* of the protein. Indeed, a polar but rigid medium can have a weak polarizability. Thus, the heme region of cytochrome *c* contains several charged and polar groups, yet has a very low Fröhlich−Kirkwood dielectric constant. Enzyme active sites are very polar,[9] yet moderately polarizable.[10] The distinction between polarity and polarizability has important implications for Poisson−Boltzmann calculations on proteins.[2,4,13−18] Thus, while a dielectric constant of 1 or 2 is usually thought to be optimal for calculations of the equilibrium field or potential in proteins (in combination with molecular mechanics charge distributions),[17] the present results indicate that a larger value may be appropriate for the calculation of relaxation properties.[2,8,9] Examples of biochemically important relaxation free energies are the reorganization free energy in electron transfer theory[8] and the proton self-energy in pK_a calculations.[2,4]

To apply the Fröhlich−Kirwood theory of dielectrics,[19] we proceed as previously,[1,4] viewing the protein as made up of two distinct, concentric, spherical regions; an inner region of radius r_1 and dielectric constant ϵ_1, and an outer region of radius r_2 and dielectric constant ϵ_2. In this geometry,

$$\frac{\langle \Delta M_1^2 \rangle}{kTr_1^3} = \frac{f(\epsilon_1, \epsilon_2, \epsilon_w)(\epsilon_1 - 1)}{f(1, \epsilon_2, \epsilon_w)} \qquad (1)$$

where ΔM_1 is the instantaneous deviation from its mean of the dipole moment of the inner region 1, ϵ_w is the solvent dielectric constant, r_w is the outer radius of the solvent sphere, and $f(\epsilon_1, \epsilon_2, \epsilon_w)$ is the ratio between the cavity field inside the inner region 1 and an applied field. Though f is easy to obtain from elementary electrostatics, it has a complicated form and is not reported here. The dimensionless term on the left of the equation is referred to as the *G*-factor, by analogy to the Kirkwood *g*-factor.

To estimate the *G*-factor, a 1 ns molecular dynamics simulation of ferro-cytochrome *c* in a 24 Å radius water sphere was performed at 295 K as described previously,[1] except that no electrostatic cutoff was used. Electrostatic interactions at distances greater than 13 Å were treated efficiently by use of a multipole approximation[20] with the CHARMM program.[21] Overall rotation−translation of the protein was substracted from the trajectory, and the last 900 ps were used for analysis.

The protein remained significantly closer to the starting X-ray structure[22] than in the earlier cutoff simulation: the rms deviation increased gradually to 1.6 Å for heavy atoms, compared to 2.2 Å previously. Convergence of the *G*-factor is satisfactory (not shown), similar to previous work.[3,4] The variance of the protein dipole moment is 148.8 ± 7.0 (eÅ)^2; the variances of its three Cartesian components are 94.5, 22.2, and 32.1 (eÅ)^2. If the charged portions of the charged protein side chains are omitted from the calculation, the variance is 14.6 ± 2.3 (eÅ)^2, with

(1) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1082−1086.
(2) Simonson, T.; Perahia, D. *J. Am. Chem. Soc.* **1995**, *117*, 7987−8000.
(3) Simonson, T.; Perahia, D. *Faraday Disc.* **1996**, *103*, 71−90.
(4) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452−8458.
(5) Smith, P.; Brunne, R.; Mark, A.; van Gunsteren, W. *J. Phys. Chem.* **1993**, *97*, 2009−2014.
(6) Bone, S.; Pethig, R. *J. Mol. Biol.* **1982**, *157*, 571−575.
(7) Bone, S.; Pethig, R. *J. Mol. Biol.* **1985**, *181*, 323−326.
(8) Marcus, R. *Ann. Rev. Phys. Chem.* **1964**, *15*, 155.
(9) Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*; John Wiley: New York, 1991.
(10) King, G.; Lee, F.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366−4377.
(11) Löffler, G.; Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1997**, *270*, 520−534.
(12) Neumann, M.; Steinhauser, O. *Chem. Phys. Lett.* **1983**, *95*, 417−422.
(13) Gilson, M.; Honig, B. *Biopolymers* **1986**, *25*, 2097−2119.
(14) Krishtalik, L.; Kuznetsov, A.; Mertz, E. *Proteins* **1997**, *28*, 174−182.
(15) Sham, Y.; Chu, Z.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 4458−4472.
(16) Gunner, M.; Alexov, E.; Torres, E.; Lipovaca, S. *J. Biol. Inorg. Chem.* **1997**, *2*, 126−134.
(17) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(18) Gilson, M.; Rashin, A.; Fine, R.; Honig, B. *J. Mol. Biol.* **1985**, *184*, 503.
(19) Fröhlich, H. *Theory of Dielectrics*; Clarendon Press: 1949.
(20) Stote, R.; Karplus, M. *Proteins* **1995**, *23*, 12−31.
(21) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(22) Berghuis, A. M.; Brayer, G. D. *J. Mol. Biol.* **1992**, *223*, 959−976.

S0002-7863(98)00071-7 CCC: $15.00
Published on Web 05/05/1998

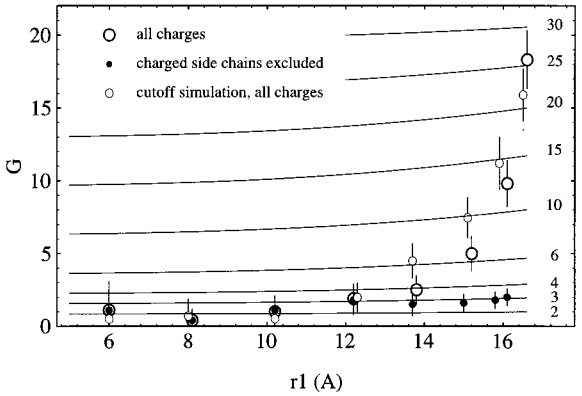

Figure 1. Radial variation of the *G*-factor of the inner protein region 1 as a function of its radius r_1. Theoretical curves (solid lines) are labeled on the right by the value of the protein dielectric constant ϵ_1. Error bars are twice the standard error, estimated from the autocorrelation function of *G*.[30]

Cartesian components 5.5, 4.3, and 4.8. Thus, almost all of the dipole fluctuations (and their anisotropy) arise from the charged side chains. The probability distributions (not shown) of ΔM_1 for different radii r_1 agree reasonably with those predicted by continuum theory,[19,23] as observed previously.

Figure 1 compares the observed *G*-factors to the values predicted by eq 1, for a series of values of ϵ_1. The assumption $\epsilon_1 = \epsilon_2$ is made in order to apply eq 1; other reasonable assumptions, such as $\epsilon_2 = 25$, lead to similar results (not shown). The variations throughout the protein bulk are consistent with a dielectric constant of 2−3, while the much greater fluctuations of charged groups near the protein surface lead to an overall dielectric constant of 25. If charged portions of the charged side chains are omitted, the *G*-factor is consistent with $\epsilon_1 = 2-3$ throughout the protein. Our earlier uncertainty analysis applies,[1] indicating that with all charged side chains the dielectric constant has an uncertainty of ±10, mainly due to uncertainty in the protein radius r_2, which is not uniquely defined. When charged side chains are excluded, the uncertainty is reduced to ±1. The low dielectric constant of the protein bulk agrees with early *in vacuo* calculations.[13,24,25] It should noted that local protein regions can presumably deviate from this average behavior, e.g., the active site of trypsin has polar fluctuations consistent with a dielectric constant of 10,[10] even when ionized groups are not considered.

The observed *G*-factors agree within statistical uncertainty with those from the earlier simulation (see the figure), despite the lack of an electrostatic cutoff here. Qualitative agreement was also observed earlier for pure water droplets simulated with and without a cutoff.[1,26] For a modified TIP3P water model,[27,28] dielectric constants of 110 and 82 were obtained with and without cutoff, respectively. This is in contrast to bulk water, where the neglect of long-range interactions underestimates the dielectric constant by more than a factor of 3.[29] The smaller cutoff artifacts in water droplets must result from the absence of long-range interactions in finite droplets. However in the protein case, other factors must play a role as well. Indeed, the dielectric constants of six proteins calculated with periodic boundary conditions and cutoffs were in the same range as in the present no cutoff calculation.[4,5] This implies that for bulk protein−solvent systems, other factors are present that limit cutoff artifacts. One such factor is the insensitivity of the protein dielectric constant ϵ_1 in eq 1 to the dielectric constant ϵ_w of the solvent, as long as ϵ_w is ≥50.[1] Thus, even if the water fluctuations, and hence ϵ_w, are significantly modified by the use of a cutoff, the impact on ϵ_1 will be limited, being "damped" by the mathematical form of eq 1. More generally, the insensitivity of the protein dipolar fluctuations to the use of a cutoff appears to result from the dielectric heterogeneity of the system, formed of a low-dielectric protein surrounded by a high-dielectric solvent. A detailed analysis will be presented elsewhere.

Steinhauser and co-workers recently proposed a formulation of dielectric theory for a protein−solvent mixture with periodic boundary conditions and Ewald summation.[11] In Fröhlich−Kirkwood theory, the dielectric constant ϵ_p of a material (e.g., protein) is defined by the ratio between the Maxwell field averaged over a large volume of the material, E_p, and the polarization averaged over the same volume, P_p:[19]

$$4\pi P_p = (\epsilon_p - 1)E_p \tag{2}$$

However, for an inhomogeneous protein−solvent system with these boundary conditions, no simple fluctuation formula for the protein dielectric constant could be obtained starting from this definition. Therefore, Steinhauser et al. started from an alternate definition of the protein dielectric constant, based on the ratio between the Maxwell field averaged over *the entire simulation cell* (protein + water), E, and the protein polarization P_p

$$4\pi P_p = (\hat{\epsilon}_p - 1)E \tag{3}$$

where $\hat{\epsilon}_p$ is a new quantity. The $\hat{\epsilon}_p$ thus defined will depend in principle on system geometry and composition and differ from the usual ϵ_p. This can be seen for a spherical protein in a very large volume of water, for example. If the cell volume is almost entirely filled with water, the Maxwell field averaged over the entire cell, E, will be essentially identical to the applied field E_0. This differs from the Maxwell field averaged over the protein volume, which is exactly the cavity field of Fröhlich theory: $E_p = 3\epsilon_w/(2\epsilon_w + \epsilon_p)E_0$.[19] The two differ by a factor $3\epsilon_w/(2\epsilon_w + \epsilon_p) \approx 1.4$. Hence, $\hat{\epsilon}_p - 1 \approx 1.4(\epsilon_p - 1)$. For the system studied by Steinhauser et al., the protein occupies only a few percent of the cell volume, so that the above argument should be quantitatively correct. Correcting the computed $\hat{\epsilon}_p = 15$ gives a peptide dielectric constant ϵ_p of 11, at the low end of the range observed for other proteins.[4] In the general case where the protein volume is not negligible compared to the cell volume, an additional factor $V_w/(V_w + V_p)$ must be applied, where V_w and V_p are the solvent and protein volumes.

In summary, the present study validates earlier studies on seven proteins, including cytochrome *c*, that used electrostatic cutoffs. The lack of significant cutoff artifacts for the protein dipole fluctuations is itself an interesting result, related to the dielectric heterogeneity of the protein−solvent systems. More importantly, the unified view of protein dielectric constants suggested by the earlier studies is established. The overall protein dielectric constants are high, ranging from about 11 to 35. This indicates that the protein contribution to biochemically important relaxation free energies can be significant. As pointed out before,[1] the dominance of a few charged side chains at the protein surface is inconsistent with a homogeneous continuum model for treating protein dielectric relaxation. A more consistent picture is obtained if one views the charged portions of the charged side chains as part of the outer, solvent, medium. In this picture, anisotropy and sensitivity to model parameters are reduced, and the calculated dielectric constants are low, close to those of dry protein powders.[6,7]

JA980071M

(23) Kusalik, P. *Mol. Phys.* **1993**, *80*, 225−231.
(24) Simonson, T.; Perahia, D.; Bricogne, G. *J. Mol. Biol.* **1991**, *218*, 859−886.
(25) Simonson, T.; Perahia, D.; Brünger, A. T. *Biophys. J.* **1991**, *59*, 670−690.
(26) Simonson, T. *Chem. Phys. Lett.* **1996**, *250*, 450−454.
(27) Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys* **1983**, *79*, 926−935.
(28) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(29) Steinhauser, O. *Mol. Phys.* **1982**, *45*, 335−348.
(30) Straatsma, T.; Berendsen, H. Stam, A. *Mol. Phys.* **1986**, *57*, 80−95.