Low-Frequency Modes in Proteins: Use of the Effective-Medium Approximation to Interpret the Fractal Dimension Observed in Electron-Spin Relaxation Measurements

R. Elber and M. Karplus

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

(Received 29 October 1985)

A method based on the effective-medium approximation is developed for calculations of the low-frequency density of vibrational states, $g(\omega)$, in proteins. The calculated values of the fractal exponent γ [$g(\omega) \sim \omega^{\gamma}$] are anomalously low ($0 < \gamma < 1$) in accord with estimates from measurements of the electron-spin relaxation time in low-spin ferric proteins. The essential element leading to the low γ value is the limited cross-chain connectivity in proteins, rather than the geometric arrangement of the polypeptide chain.

PACS numbers: 87.15.-v, 36.20.-r

The internal motions of protein molecules are being examined theoretically and experimentally by a variety of techniques.[1,2] Of particular interest, because of their possible role in protein function, are the low-frequency modes that extend over the dimensions of the protein. Evidence concerning such modes has been obtained from molecular dynamics[3] and normal-mode[4-6] calculations. Measurements providing direct information about the nature of these modes are difficult. However, there are experimental approaches to the number or density of modes in the low-frequency region. These include inelastic neutron scattering,[7,8] heat capacity,[9-11] and electron-spin relaxation[12-14] measurements. It is the latter that are the primary concern of this Letter.

Stapleton and co-workers[12-14] have found that in the temperature range between about 4 and 20 K the electron-spin relaxation rate ($1/T_1$) of low-spin ferric iron in a number of heme and iron-sulfur proteins is dominated by a two-phonon (Raman) process with a temperature dependence that deviates significantly from the T^9 power law expected[15,16] and found experimentally[17,18] in ordinary three-dimensional solids. For a Debye temperature much higher than the temperature of interest and for vibrations distributed uniformly over the molecule,[19] the contribution of the Raman process can be shown to be given by the integral

$$\frac{1}{T_1} = \int_0^{\infty} \frac{\omega^4 [g(\omega)]^2 \exp(h\omega/kT)}{[\exp(h\omega/kT) - 1]^2} d\omega. \qquad (1)$$

There is evidence from detailed vibrational calculations for small proteins[4-6] that in the relevant frequency range the normal modes are indeed delocalized. Equation (1) leads to a temperature dependence of the form $T^{5+2\gamma}$ if the density of states of the contribution modes can be expressed as a simple power law with $g(\omega)$ proportional to ω^{γ}. For an ordinary solid, $\gamma = D - 1 = 2$, where D is the physical dimension, so that $1/T_1$ varies as T^9. Stapleton and co-workers[12-14] have found temperature dependencies in the range $T^{5.6}$ to $T^{6.3}$ for ferric-iron–containing heme and nonheme proteins. This corresponds to values of $\gamma = 0.3$ to 0.65 or $D_{\text{eff}} = 1.3$ to 1.65, which they have interpreted in terms of estimates of the geometric fractal dimension of the α-carbons of the polypeptide chain. However, Alexander and Orbach[20] have pointed out that for a single chain, D_{eff} is at most equal to 1, irrespective of the geometric shape. A possible way to overcome this objection was suggested by Helman, Coniglio, and Tsallis,[21] who proved that D_{eff} approaches the fractal dimension if there is a high probability of "connections" between different portions of the chain. The existence of a connection requires that the "cross chain" interactions be of the same order as those along the chain. As we point out below, such a three-dimensional network does not, in fact, occur in globular proteins.

In this Letter we take an approach to the fractal dimension, D_{eff}, that is based on the essential interaction between the amino-acid residues of which a protein is composed. Each amino acid has the possibility of four strong interactions, two with its bonded neighbors along the chain and two with other amino acids involving its hydrogen-bonding (C=O and NH) groups. These four interactions, if saturated, would lead to a protein that behaves as a two-dimensional object ($D_{\text{eff}} = 2$). However, the actual value of D_{eff} is expected to be less than 2 ($1 < D_{\text{eff}} < 2$), in agreement with experiment, because many of the hydrogen bonds are short range (e.g., those in α helices) and so they affect only the interactions along the chain rather than connecting different portions of the chain. It is this fact which leads to the low connectivity that violates the Stapleton-Herman model. Table I shows the fraction of the hydrogen bonds involved in connecting different portions of the chain for proteins that have been studied experimentally. The α-helical proteins (myoglobin and cytochrome c 551) have particu-

larly low values, as expected; ferredoxin, which has no α helices, has a connectivity that is significantly higher.

To demonstrate that the qualitative picture based on chemical connectivity gives quantitatively correct results, we have developed an approximate method based on the effective-medium approximation (EMA)[22] for calculating $g(\omega)$ in the low-frequency region (0 to 75 cm^{-1}) that makes the dominant contributions to $1/T_1$ and have used the results to evaluate the integral in Eq. (1).

We represent the protein motions in terms of the two main-chain dihedral angles ϕ and ψ for each amino acid and assume that its interactions can be approximated by those with the four "nearest" neighbors; two along the chain and two that are involved in strong main-chain hydrogen bonds (see above) or, if there are none, are closest to the amino acid being considered, but are at least six residues removed along the chain; the latter restriction eliminates the α-helical hydrogen bonds and corresponds approximately to the wavelength limit of the T_1 process. Such a reduced set of interactions leads to the following expression for the Laplace transform of the normal-mode equations for the protein vibrations:

$$(M_n^m\omega^2+\Lambda_n^m)X_n^m+\chi_{n,n+1}^m(X_n^m-X_{n+1}^m)+\chi_{n,n-1}^m(X_m^m-X_{n-1}^m)$$
$$+\Gamma_n^{m,m+1}(X_n^m-X_n^{m+1})+\Gamma_n^{m,m-1}(X_n^m-X_n^{m-1})=M_n^m\delta_{m0}\delta_{n0}, \quad (2)$$

where $\chi_{n,n\pm1}^m$ and $\Gamma_n^{m,m\pm1}$ are the nearest-neighbor force-constant matrices, Λ_n^m is the internal force-constant matrix for an amino acid, M_n^m is the "mass" matrix, and X_n^m is the Laplace transform of the dihedral angle coordinates; the right-hand side of Eq. (2) arises from the choice of initial condition [i.e., $x_n^m(t=0)=\delta_{m0}\delta_{n0}$ and $\dot{x}_n^m(t=0)=0$]. The index n runs over the chain units while m runs over the cross links. The EMA replaces the distribution of variables by a single, frequency-dependent variable.[22] The implicit assumption made here is that the protein may be described by an infinite two-dimensional array of bonds on the scale of the acoustic wavelength which is relevant to the measurement; this is on the order of 8 to 25 Å. To simplify the calculations, we treat each direction independently within the EMA approximation. For the present problem the most convenient frequency-dependent variable is the so-called Dyson variable; that is, $\xi_n^+=M_n^{-1}A_{n,n+1}(X_n-X_{n+1})X_n^{-1}$. Since we are treating the two dimensions separately, we have dropped the superscript and n refers to the dimension of interest; the force-constant matrix element $A_{n,n+1}$ depends on which dimension is being treated. It can be shown[22] that ξ_n^+ satisfies the following form of iterative equation:

$$\xi_n^+=M_n^{-1}[A_{n,n+1}^{-1}+(M_{n+1}\omega^2+M_{n+1}\xi_{n+1}^+)^{-1}]^{-1}; \quad (3)$$

there is a corresponding equation for ξ_n^-. The self-correlation function $Y_0(\omega^2)$ is given by

$$Y_0(\omega^2)=(\omega^2+\xi_0^++\xi_0^-)^{-1} \quad (4)$$

and the density of states is obtained from $Y_0(\omega^2)$ averaged over all possible values of $\xi_0^\pm$ by standard methods.[22] The total density of frequencies $N(\omega^2)$ is then given by the convolution

$$N(\omega^2)=\int_0^{\omega^2}d\omega'^2\,N_n(\omega'^2)N_m(\omega^2-\omega'^2),$$
$$g(\omega)=2\omega N(\omega^2), \quad (5)$$

where N_n and N_m are the eigenvalue densities for the two directions.

To implement Eqs. (1)–(3) we need values of the "mass" matrix elements M_n^m and the force-constant matrix elements $A_{n,n+1}$. The force-constant matrices were obtained for each amino acid n by numerical differentiation of the potential energy with respect to ϕ_n, ψ_n and the ϕ, ψ of the four nearest neighbors. Thus, the second-derivative matrices are divided into two by two blocks, one for n itself and four involving n and each of the neighbors; e.g., ϕ_n, ψ_n and ϕ_{n+1}, ψ_{n+1}. The "mass" matrices are determined by use of the standard formulas for torsions.[23]

To test the EMA in the present form for proteins, 625×625 random matrices were generated with force-constant distributions similar to these actually calculated for the proteins studied. In Fig. 1, we present a comparison between EMA calculations and actual diagonalization of the random matrix. Several tests were

TABLE I. Analysis of iron-containing proteins.

Protein[a]	Fraction of H bonds[b]	Exponent γ[c] Expt.	Theory
Myoglobin	0.14	0.62	0.6
Cytochrome c 551	0.30	0.43	0.4
Ferredoxin	0.60	0.34	0.4

[a]Structures obtained from the Brookhaven Protein Data Bank where the original references can be found.

[b]A relevant hydrogen bond is considered to exist when a donor (N of NH) and acceptor (O of CO=O) are at least six residues apart with the distance between the two less than 6 Å; such a large cutoff is used to obtain an upper limit to the connectivity.

[c]The distribution function $g(\omega)$ is taken to have the form ω^γ.

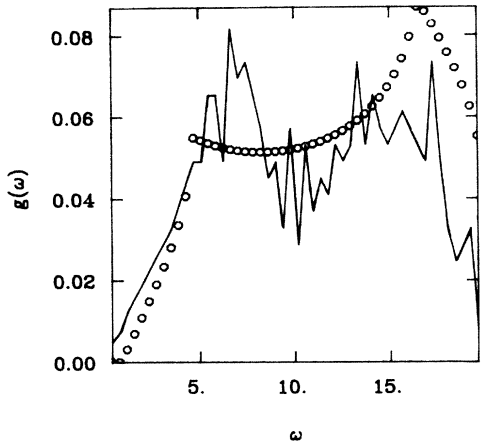
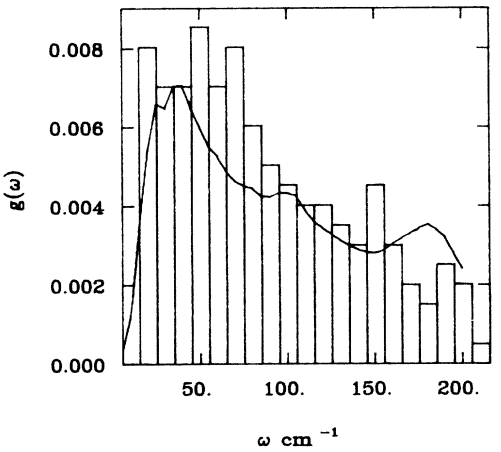


FIG. 1. Comparison of $g(\omega)$ obtained from EMA and exact diagonalization of a 25-residue effective-dihedral-angle chain: solid line, exact; circles, EMA. Along the chain, the nearest-neighbor force-constant distribution has the form $\rho(\chi)=95^{-1}$ for $5<\chi<100$ and zero otherwise; across the chain the distribution is $\rho(\Gamma)=\{(1-\alpha)/\Gamma_c^{1-\alpha}\}\Gamma^{-\alpha}$ for $\Gamma<\Gamma_c$ and zero otherwise with $\alpha=0.7$ and $\Gamma_c=7.5$; $M=1$.

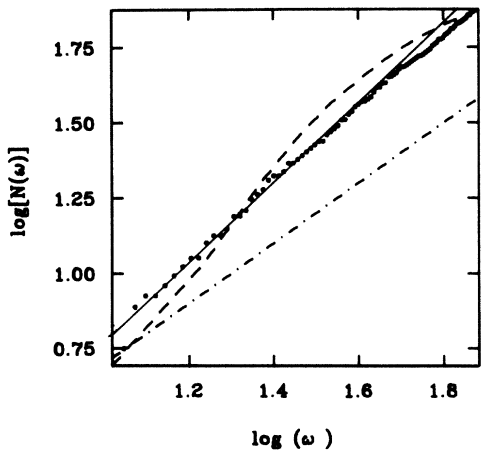
run with different values for the parameters of the distribution function given in the caption to Fig. 1. In all cases the agreement corresponded to that shown in the figure. Thus, the EMA appears to yield satisfactory results for $g(\omega)$ in the low-frequency region for proteinlike force-constant distributions. Further, separation into two independent directions is justified when the EMA is valid for each one-dimensional problem.

Figure 2 shows a comparison between the EMA results and a full normal-mode calculation in (ϕ,ψ) space for the bovine pancreatic trypsin inhibitor, a protein composed of 58 residues; the latter was done by evaluating the second-derivative matrix exactly by use of the CHARMM potential function[24] and then obtaining the normal modes corresponding to that matrix.[5] Figure 2(a) shows $g(\omega)$, which is reproduced quite well by the EMA in the region up to 200 cm^{-1}; in Fig. 2(b) the integrated distribution functions are compared. Although the EMA and full normal-mode results are not identical, they are sufficiently close to suggest that the former method, which is much simpler to use, can be employed for the present problem.

To apply the EMA to the proteins listed in Table I for which electron-spin-relaxation data are available, the force-constant and mass matrices were evaluated from their known structures by use of the program CHARMM and the model described above. These were substituted into Eq. (5) which was solved for $\xi_{\rm eff}(\omega^2)$ over the desired frequency range (0 to 70 cm^{-1}) at 0.5 cm^{-1} intervals. To calculate $1/T_1$, $g(\omega)$ was substituted into Eq. (1) and the integral evaluated as a function of temperature. By a plot of log $(1/T_1)$ vs log T (Fig. 3), the effective frequency exponents valid



(a)



(b)

FIG. 2. Comparison of EMA results with full normal-mode treatment for the bovine pancreatic trypsin inhibitor. (a) $g(\omega)$ vs ω: histogram is from full normal-mode results in (ϕ,ψ) space; solid line is the EMA result. (b) $\log N(\omega)$ vs $\log\omega$ where $N(\omega)$ is the integrated distribution function: dotted line, full normal-mode calculation; dashed line, the EMA result; also shown are the results for $N(\omega)$ proportional to $\omega^{1+\gamma}$ with the best fit $\gamma=0.35$ (solid line) and $\gamma=0$ (dot-dashed line).

between 1 and 10 K were obtained for comparison with the experimental fits. The results given in Table I show that the observed behavior is reproduced, with the calculated values deviating from experiment by less than the error bounds.

The essential element in the above analysis of the observed fractal behavior of the density of states is the difference in the connectivity along the chain and across the chain. For the main peptide chain the interaction is always strong and the force-constant values are never close to zero. This is not the case for the cross links, where the force-constant distribution is peaked at the origin. These two types of distribution

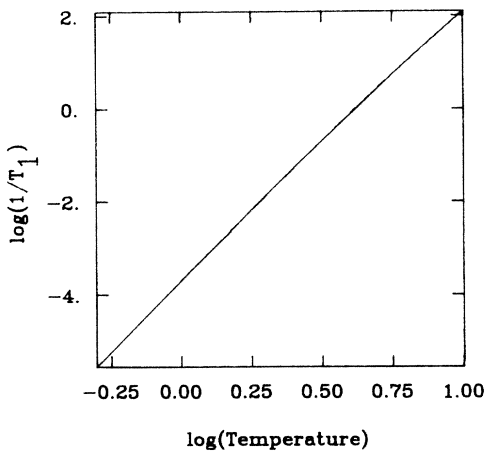


FIG. 3. Calculated electron-spin-resonance relaxation time $(1/T_1)$ plotted vs temperature for cytochrome c 551.

have been considered in previous applications of the EMA to a variety of problems. They can be used to introduce an analytical model for the low-frequency probability density.[22] For cytochrome c 551, which we use as an example, the main chain is indeed strongly connected. Thus, the density of states $g_n(\omega)$ asymptotically approaches a constant as $\omega \rightarrow 0$.[22] The distribution of crosslink force constants can be related to a probability density for which an analytical solution exists at $\omega \rightarrow 0$;[21] i.e., a probability density of the form $\rho(\Gamma) \sim \Gamma^{-\alpha}$ used in Fig. 1. Using the numerical distribution of force constants, we obtain the value $\alpha = 0.6 \pm 0.1$. Performing the convolution of the two distributions, we find for this protein model that $\gamma = 2(1-\alpha)/(2-\alpha)$. The frequency exponent extracted from this model is $\gamma = 0.57$, as compared with the experimental value of 0.43 and the full EMA value of 0.4 (see Table I). The difference between the analytic model and full calculation is reasonable since the former considers only the limit as $\omega \rightarrow 0$.

The success of the EMA and simplified analytic model for the density of the low-frequency vibrations of proteins supports the validity of the explanation for the observed "anomalous" fractal dimension based on the physical strength of the interactions between amino acids rather than on geometrical effects. It is likely that the apparent dimensionality of the specific heat measured for solid polypeptide chains at low temperatures (1D for α-helical poly-alanine, 2D for β-sheet poly-alanine)[9] has a corresponding origin.

This work was supported in part by the National Science Foundation.

[1]M. Karplus and J. A. McCammon, Annu. Rev. Biochem. **53**, 263 (1983).
[2]P. G. Debrunner and H. Frauenfelder, Annu. Rev. Phys. Chem. **33**, 283 (1982).
[3]S. Swaminathan, T. Ichiye, W. F. van Gunsteren, and M. Karplus, Biochemistry **21**, 5230 (1982).
[4]N. Go, T. Noguti, and T. Nishikawa, Proc. Nat. Acad. Sci. U.S.A. **80**, 5671 (1983).
[5]B. Brooks and M. Karplus, Proc. Nat. Acad. Sci. U.S.A. **80**, 6571 (1983).
[6]M. Levitt, C. Sanders, and P. S. Stern, J. Mol. Biol. **181**, 423 (1985).
[7]B. Jacrot, S. Cusack, A. J. Dianoux, and D. E. Engleman, Nature **300**, 84 (1982).
[8]J. Smith, S. Cusack, U. Pezzeca, B. Brooks, and M. Karplus, J. Chem. Phys. (to be published).
[9]L. Finegold and J. L. Cude, Nature **238**, 38 (1972).
[10]V. I. Goldanskii, Y. F. Krupyanskii, and N. V. Flerov, Dokl. Akad. Nauk SSSR **272**, 978 (1983).
[11]G. P. Singh, H. J. Schink, H. v. Lohnysen, P. Parak, and S. Hunklinger, Z. Phys. B **55**, 23 (1984).
[12]R. C. Herrick and H. J. Stapleton, J. Chem. Phys. **65**, 4778 (1976).
[13]J. P. Allen, J. T. Colvin, D. G. Stinson, C. P. Flynn, and H. J. Stapleton, Biophys. J. **38**, 299 (1982).
[14]G. C. Wagner, J. T. Colvin, J. P. Allen, and H. J. Stapleton, J. Am. Chem. Soc. **107**, 5589 (1985).
[15]K. W. H. Stevens, Rep. Prog. Phys. **30**, 189 (1967).
[16]R. Orbach and H. J. Stapleton, in *Electron Paramagnetic Resonance,* edited by S. Geschwind (Plenum, New York, 1972).
[17]T. G. C. Bray, J. Brown, and A. Kiel, Phys. Rev. **127**, 730 (1962).
[18]A. Rannestad and P. E. Wagner, Phys. Rev. **131**, 1953 (1963).
[19]S. Alexander, O. Entin-Wohlman, and R. Orbach, J. Phys. (Paris), Lett. **46**, L555 (1985).
[20]S. Alexander and R. Orbach, J. Phys. (Paris), Lett. **43**, L625 (1982).
[21]J. S. Helman, A. Coniglio, and C. Tsallis, Phys. Rev. Lett. **53**, 1195 (1984).
[22]S. Alexander, J. Bernasconi, W. R. Schneider, and R. Orbach, Rev. Mod. Phys. **53**, 175 (1981).
[23]J. C. Decius, J. Chem. Phys. **16**, 1025 (1948).
[24]B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, No. 2, 187 (1983).