JOURNAL

OF THE AMERICAN CHEMICAL SOCIETY

VOLUME 107, NUMBER 20 OCTOBER 2, 1985

Fractal Models of Protein Structure, Dynamics, and Magnetic Relaxation

Gerald C. Wagner,*† J. Trevor Colvin,‡ James P. Allen,§ and Harvey J. Stapleton

Contribution from the Departments of Physics and Biochemistry, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801. *Received February 25, 1985*

Abstract: A geometric model is presented to interpret the anomalous T^{3+2m} temperature dependence of the Raman spin–lattice relaxation rates in heme and iron–sulfur proteins. Analysis of relaxation data is based on a modified Debye relationship between the spectral exponent m and the density of vibrational states $\rho(\nu) \propto \nu^{m-1}$, where $0 \leq \nu \leq \nu_{max}$. Magnetic relaxation measurements on cytochrome *c*-551 and putidaredoxin yield noninteger values of m that are influenced by changes in the ionic medium. The apparent physical significance of m is revealed, in part, by correlation to a protein's fractal geometry, which characterizes a repeating structural motif by a single parameter called the fractal dimension $\bar{d}$. Estimates of $\bar{d}$ for 70 proteins are computed by a method that identifies geometric and statistical self-similarities of α-carbon coordinates; values range within the limits ($1 \leq \bar{d} \leq 2$) of well-defined test structures and correlate principally with dominant elements of secondary structures. In six iron proteins, the highest values of m derived from relaxation data are approximated by the estimated values of $\bar{d}$ calculated from the covalent structure. The interrelationship between the fractal models of protein structure and molecular dynamics, i.e., $m = \bar{d}$, is also evident in the good agreement between the predicted $\rho(\nu) \propto \nu^{\bar{d}-1}$ and the reported distribution of low-frequency normal modes ($\nu \leq 75$ cm^{-1}) calculated for bovine pancreas trypsin inhibitor. The present findings indicate $\bar{d}$ defines a fundamental parameter that is inherent to both the structural and dynamic properties of a protein.

In earlier works we reported an anomalous temperature dependence of the Raman electron spin–lattice relaxation rates in various iron-containing proteins.[1,2] The well-known Raman mechanism is an inelastic, two-phonon-scattering process that balances the magnetic energy of a spin flip, $g\beta H$, of a paramagnetic ion with the change in vibrational energy $h(\nu_1 - \nu_2)$ between lattice modes. At sufficiently low temperatures, the integral expression for the Raman relaxation rate, $1/T_1$, reduces to a simple T^n power law in temperature. The value of n depends strongly upon the functional form of the density of vibrational states $\rho(\nu)$,[1,2] which is defined as the number of vibrational states $\rho(\nu)\,d\nu$ within the frequency interval ν and $\nu + d\nu$. We asserted previously that the general form of $\rho(\nu)$ is a power law relationship[2]

$$\rho(\nu) \propto \nu^{m-1} \qquad 0 \leq \nu \leq \nu_{max} \tag{1}$$

provided the velocity of sound v_s is independent of ν,[3] and the Debye cutoff frequency ν_{max} is chosen to make the total number of vibrational modes equal to the degrees of freedom within the system, i.e., $\rho(\nu) = 0$ for $\nu > \nu_{max}$. This form of $\rho(\nu)$ results in

$$1/T_1(\text{Raman}) \propto T^n \equiv T^{3+2m} \tag{2}$$

for ions with an odd number of unpaired electrons (Kramers' ions). The spectral exponent m corresponds normally to the Euclidean dimension d of the space filled by the periodic lattice structure of a paramagnetic salt. As $m = d = 3$, eq 1 and 2 reduce to the well-known expressions in solid-state theory of a ν^2 density of vibrational states and a T^9 Raman rate. In heme and iron–sulfur proteins,[1,2] the temperature dependence of the relaxation rates also follows a T^n relationship, which is consistent with a simple power law for $\rho(\nu)$. The temperature exponents n, however, are noninteger values, well below 9, that correspond to spectral exponents between one and two. The noninteger values of m, by analogy, imply lattice structures of non-Euclidean or fractional dimension.

Structures of fractional dimension are known. Mandelbrot[4,5] has pioneered the theoretical concepts and physical applications of this relatively new field of geometry[6,7] and has popularized the term *fractal* for a structure characterized by a fractional dimension. By definition, any structure possessing a self-similar or repeating motif that is invariant under a transformation of scale may be represented by a fractal dimension. Self-similarity is geometric in regular structures; in random or irregular objects, self-similarity is primarily statistical in nature. The average (root mean square, RMS) end-to-end length R of an unbranched polymer chain constitutes a statistically self-similar property. The

† Present address: Department of Physics, University of Illinois, Urbana.
‡ Present address: Hughes Research Laboratories, Malibu, CA 90265.
§ Present address: Department of Physics, University of California at San Diego, La Jolla, CA 92093.

(1) Stapleton, H. J.; Allen, J. P.; Flynn, C. P.; Stinson, D. G.; Kurtz, S. R. *Phys. Rev. Lett.* **1980**, *45*, 1456–1459.
(2) Allen, J. P.; Colvin, J. T.; Stinson, D. G.; Flynn, C. P.; Stapleton, H. J. *Biophys. J.* **1982**, *38*, 299–310.
(3) Kittel, G. "Introduction to Solid State Physics", 5th ed.; Wiley & Sons, Inc.: New York, 1976; pp 131–136.
(4) Mandelbrot, B. B. "Fractals: Form, Chance, and Dimension"; W. H. Freeman & Co.: San Francisco, 1977.
(5) Mandelbrot, B. B. "The Fractal Geometry of Nature"; W. H. Freeman & Co.: San Francisco, 1982.
(6) Gardner, M. *Sci. Am.* **1978**, *238*(*4*), 16–32.
(7) McDermott, J. *Smithsonian* **1983**, *14*, 110–117.

0002-7863/85/1507-5589$01.50/0

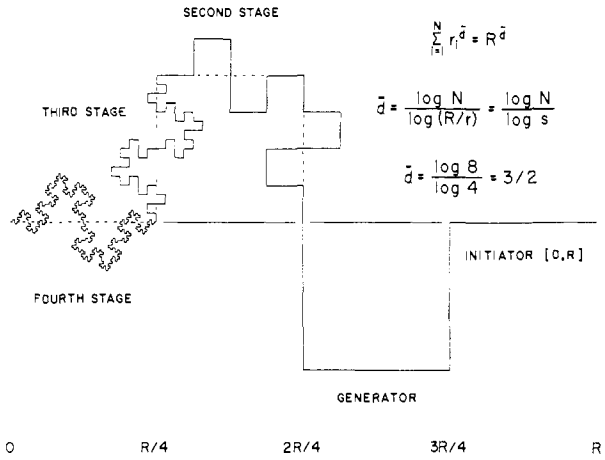

Figure 1. A fractal structure illustrating properties relevant to a polypeptide. Fractals with geometric self-similarity are comprised of two elements—an initiator, which in this example is a straight line over the interval $[O,R]$, and a generator, which is the largest N-sided structure that spans the interval $[O,R]$. The most general relation characterizing a fractal is given by the uppermost expression (see page 56 of ref 5). If all lengths r_i describing the generator are equal, which is a good approximation for C_α–C_α distances in proteins ($r \simeq 3.8$ Å), then the expression simplifies to the more common form in the second line (eq 3). A fractal dimension is calculated directly from the properties of the initiator and generator ($N = 8$, $r = R/4$). The ratio R/r also defines a scaling factor s of the generator that determines the maner of constructing the fractal in an infinite number of stages. Note that each stage of complete construction in the figure should extend over the entire interval of the initiator and that a fifth stage would be virtually imperceptible. The $\bar{d}$ of a structure, however, is independent of the number of constructions and the increasing contour length. Estimating $\bar{d}$ for an arbitrary polypeptide, without an explicit form for the initiator or generator, is based on identifying statistically the most common repeating motifs within the structure.

fundamental relationship in fractals predicts that the number of monomer segments N of length r is related to R by the fractal dimension $\bar{d}$

$$N = (R/r)^{1/\nu_F} \equiv (R/r)^{\bar{d}} \tag{3}$$

where the exponent is also equal to the inverse Flory constant ν_F in polymer theory.[8] Theoretical considerations provide limits of $1 \leq \bar{d} \leq 2$ in eq 3 that correspond to a linear structure, $R = rN$, and a structure represented by an unrestricted random walk, i.e., Brownian motion in $\bar{d} \geq 2$, where $R = rN^{1/2}$.[9]

Our previous studies applied these basic concepts of fractals to protein structures and, as a consequence, equated the noninteger spectral exponents directly to fractal dimensions. The fractal interpretation was enhanced further by developing a method to estimate the fractal dimension of protein structure from crystallographic coordinates of the α-carbon backbone atoms. An excellent correlation was found in two hemeproteins between the $\bar{d}$ estimated from the covalent structure and m deduced from dynamic relaxation data. The implications that relaxation measurements could probe both the structural and dynamic properties of a noncrystalline polypeptide are intriguing. The present study extends the critical tests of the proposed fractal models with complete data (m and $\bar{d}$) for four additional proteins. New factors that alter the spectral exponents in protein dynamics are identified and interpreted within available theoretical models. Finally, fractal dimension estimates are refined from our earlier method and correlated with the structural properties of 70 proteins.

Experimental Procedures

Putidaredoxin from *P. putida* is purified according to standard procedures[10] and concentrated to 8.3 mM in a buffer of 50 mM Tris-Cl pH 7.4 and 10 mM β-mercaptoethanol that contained zero or 1.0 M NaCl. Samples are reduced aerobically with a minimal excess of solid $Na_2S_2O_4$ and frozen immediately in liquid nitrogen. Cytochrome *c*-551 from *P. aeruginosa* is purified according to a modification[11] of the standard procedure.[12] Samples are concentrated to 6.4 mM in a buffer of 50 mM KP_i pH 7.0 that contained zero or 1.0 M NaCl.

Electron spin–lattice relaxation rates are measured to 10^4 s^{-1} with the pulse-saturation/recovery technique. A short pulse of microwave power, 30–50 dB above a low-power monitoring level, is generated by microwave diode switches to drive the EPR absorption signal to zero. The exponential recovery of the saturated signal to the thermal equilibrium value is monitored in real time with a signal averager. Temperatures between 1.5 and 25 K are measured and controlled to within ±4 mK by the output from a calibbrated germanium or carbon glass resistance thermometer that is mounted directly to the upper cavity and sample holder.

The observed relaxation rates $1/T_1$ are comprised of three terms—a direct, a Raman, and an Orbach (resonant Raman) process—with each mechanism characterized by a unique temperature dependence. In low-spin ferric hemeproteins and reduced Fe_2S_2 proteins, rate contributions from an Orbach process are negligible because of the isolated $S = {}^1/_2$ ground state. Analysis of the Raman rate, therefore, requires only an explicit determination of the direct rate process.

The direct relaxation mechanism is a one-phonon process that couples a narrow band of vibrational modes at the Larmor frequency ($g\beta H/h \sim$ 9 GHz) to the precessing electron spins. In comparison to the second-order Raman mechanism, which utilizes the entire vibrational frequency spectrum ($\nu \leq \nu_{max}$), the direct process is dominant only at very low temperatures (<3 K) and is frequently phonon limited or bottlenecked.[13-16] The complete expressions for $1/T_1$ are given by a sum of the direct and Raman (eq 2) processes

$$1/T_1 = A' \coth (h\nu/2k_BT) + CT^n \approx AT + CT^n \tag{4a}$$

$$1/T_1 = \frac{A' \coth (h\nu/2k_BT)}{1 + B' \tanh (h\nu/2k_BT)} + CT^n \tag{4b}$$

$$1/T_1 = (A'/B') \coth^2 (h\nu/2k_BT) + CT^n \approx BT^2 + CT^n \tag{4c}$$

where the first terms in each equation correspond to an expression for a normal (4a), a partially bottlenecked (4b), and fully bottlenecked (4c) direct process. Observed relaxation rates are analyzed according to one of the three forms of eq 4 with a weighted, nonlinear least-squares program. Weighting factors are chosen to minimize the unweighted relative error for all data points.

The basic properties of a fractal structure are reviewed in Figure 1. Fractal dimensions of proteins are estimated from the slope $(1/\bar{d})$ in a logarithmic plot of eq 3, where N is treated as the independent variable. End-to-end lengths (RMS) of $R(N \geq 2)$ and $R(N = 1) \equiv r$ of a covalent polypeptide chain are derived from the backbone represented by databank[17] α-carbon coordinates. To account for both random (statistical) and ordered elements (geometric, secondary structure) in proteins, $R(N)$ values are calculated as a weighted mean between the two limits, $R_S(N)$ and $R_G(N)$. Statistical self-similarity, in either an ensemble or a single chain of random polymers containing M residues,[8] is dependent on all possible combinations of the mean-square lengths $R_s^2(N)$

$$R_S^2(N) = (1/(M-N)) \sum_{i=1}^{M-N} R_{i,i+N}^2 \tag{5}$$

where $R_{i,i+N}$ is the through-space distance between residues i and $i + N$ separated by N sequential α-carbon segments. An algorithm for identifying geometric self-similarity in arbitrary structures is dependent on the mean-square lengths $R_G^2(N)$ from a consecutive sequence of segments, each comprised of N-adjacent α-carbon segments, i.e,

$$R_G^2(N,j) = (1/D) \sum_{i=1}^{D} R_{j+(i-1)N,j+in}^2 \tag{6}$$

where j is a residue index, $j = 1, 2, 3, \ldots, N-1$, and D is limited solely

(8) Havlin, S. Ben-Avraham, D. *J. Phys. A: Math. Gen.* **1982**, *15*, L311–L328.
(9) Tanford, C. "Physical Chemistry of Macromolecules"; Wiley & Sons, Inc.: New York, 1961; 152–154.
(10) Gunsalus, I. C.; Wagner, G. C. *Methods Enzymol.* **1978**, *52*, 166–188.
(11) Toscano, Jr., W. A. Ph.D. Thesis, University of Illinois, Urbana, 1978.
(12) Gudat, J. C.; Sing, J.; Wharton, D. C. *Biochem. Biophys. Acta* **1973**, *292*, 376–390.
(13) Van Vleck, J. M. *Phys. Rev.* **1941**, *59*, 724–729.
(14) Scott, P. L.; Jeffries, C. D. *Phys. Rev.* **1962**, *127*, 32–51.
(15) Faughnan, B. W.; Strandberg, M. W. P. *J. Phys. Chem. Solids* **1961**, *19*, 155–166.
(16) Mikkelson, R. C.; Stapleton, H. J. *Phys. Rev.* **1965**, *140*, A1968–A1982.
(17) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. D.; Vennard, O.; Shimanouchi, T.; Tasumu, M. *J. Mol. Biol.* **1977**, *112*, 535–542.

Table I. Relaxation Parameters of Iron Proteins

protein	source	ionic strength I	relaxation parameters[a] direct A	direct B	direct A'	direct B'	Raman[b] C	Raman[b] n	Raman[b] m
hemeproteins									
myoglobin · OH[c]	sperm whale	~0.25		2.79			0.007267	6.22 (9)	1.61 (5)
myoglobin · N_3[c]	sperm whale	0.6	3.11				0.00617	6.29 (8)	1.65 (4)
cytochrome P450[c]	P. putida	0.2	0.152				0.00910	6.27 (6)	1.64 (3)
cytochrome c[c]	horse	0.1, crystal	4.85				0.0221	6.34 (6)	1.67 (3)
cytochrome c-551	*P. aeruginosa*	1.1			4.14	8.90	0.064	5.86 (18)	1.43 (9)
cytochrome c-551	*P. aeruginosa*	0.1			8.43	8.26	0.256	4.55 (26)	0.88 (14)[c]
iron sulfur proteins									
ferredoxin[c]	*S. maxima*	~0.005		0.861			0.00359	5.67 (11)	1.34 (6)
putidaredoxin	*P. putida*	1.04		0.439			0.000371	5.67 (6)	1.34 (3)
putidaredoxin	*P. putida*	0.04		0.777			0.00156	5.22 (6)	1.11 (3)

[a] Parameters have units that yield rates in s^{-1} for temperatures given in K. [b] Values in parentheses correspond to magnitude of uncertainty in n and m, e.g., (4) ≡ ±0.04. [c] Allen et al., 1982. [d] Value estimated from the average with a second analysis that excludes the lowest temperature datum in Figure 3: $A' = 35.6$, $B' = 62.0$, $C = 0.104$, $n = 4.96$ (50).

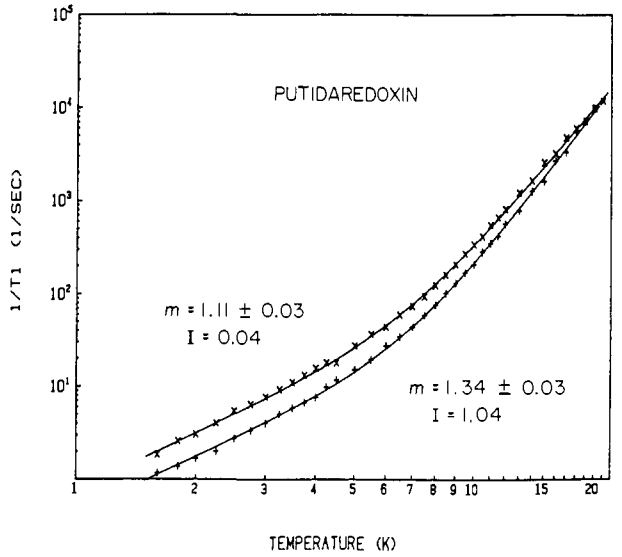

Figure 2. Relaxation data on two frozen solutions of putidaredoxin. The temperature dependence of the Raman relaxation rates (eq 2, Table I) implies a spectral exponent $m = 1.11 \pm 0.03$ for the sample in a low ionic strength solution ($I = 0.04$) and $m = 1.34 \pm 0.03$ for the sample in a 1.0 M NaCl solution ($I = 1.04$). Rates are monitored on the EPR absorption maximum at $g = 1.94$.

by M. For example, $R_G^{2\,1,8}$ would only contain the terms $R_{1,9}^2$, $R_{9,17}^2$, and $R_{17,25}^2$ if $26 \leq M \leq 34$. For each N, $R_G(N)$ is determined by the one value of j that yields a minimum in $\bar{d}$[18] according to eq 3. A best estimate of $\bar{d}$ is derived from a weighted, linear least-squares analysis of log N vs. log $(R(N)/r)$ data, where the intercept is constrained to the origin. Weighting factors originate from the inverse of the variances in $R_S^2(N)$ and $R_G^2(N)$ (from the one value of j) and compensate for the limiting sampling of $R(N)$ in proteins. The region of the data $2 \leq N \leq N_{max}$ that conforms best to a linear relationship is determined by the $\bar{d}$ values at the succession of local maxima in the linear correlation coefficient. A $\bar{d}$ is assigned from the largest range of N that yields a slope unchanged, within the standard deviation of its determination, from the preceeding estimate and uncertainty.

Results

Relaxation data from previous studies[1,2] on mygoblin, cytochrome c, cytochrome P450, and algal ferredoxin are summarized in Table I with the present findings on putidaredoxin and cytochrome c-551. Putidaredoxin is a redox-transfer protein (M = 106) that couples an NADH/FAD reductase to cytochrome P450 in the camphor monoxygenase system. The two high-spin irons

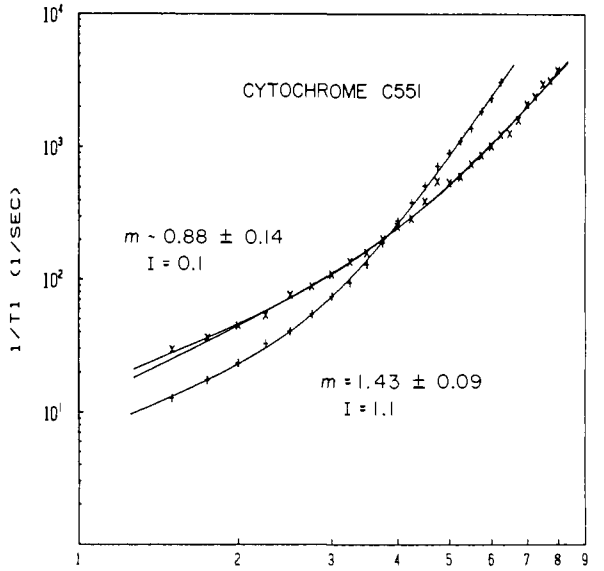

Figure 3. Relaxation data on two frozen solutions of cytochrome c-551. The temperature dependence of the Raman relaxation rates implies a spectral exponent near the theoretical lower limit of unity for the sample in a low ionic strength solution ($I = 0.1$). Data from the sample in a 1.0 M NaCl solution ($I = 1.1$) indicate $m = 1.43 \pm 0.09$. Rates are monitored at $g = 2.04$ ($I = 0.1$) and $g = 2.16$ ($I = 1.1$).[37]

in Pd, like all Fe_2S_2*Cys_4 plant-type ferredoxins, are antiferromagnetically coupled to form a diamagnetic $S = 0$ species in the oxidized state and a one-electron, paramagnetic $S = 1/2$ reduced state. Relaxation rates from two frozen solutions of reduced Pd, which differ only in solvent composition, are presented in Figure 2. Since the lowest excited state in Pd is sufficiently high, relative to the temperatures examined in Figure 2, a significant Orbach relaxation process is excluded from the data. Relaxation parameters in Table I indicate the direct process in both samples is strongly bottlenecked. More significant, however, is the dependence of the spectral exponent on the ionic strength of protein's aqueous solvent. This rather surprising result is the first experimental observation that m may vary with the solvent medium.[19] Remarkably, a similar set of results are obtained with cytochrome c-551, a small redox-transfer hemeprotein (M = 82). Relaxation

(18) Test calculations on the fractal structure in Figure 1 indicate this criterion leads to a proper assignment of the $\bar{d}$ value when considering contiguous subsections of the structure that start at an arbitrary point and extend for N = 50, 150, or 450 elemental segments. These N values were chosen to be comparable to the number of residues in globular proteins.

(19) Previous studies on myoglobin and cytochrome c indicated m was essentially independent of the solvent medium (ref 2), although a systematic investigation of ionic conditions was not performed. More recent measurements (Colvin, J. T.; Stapleton, H. J. *J. Chem. Phys.* **1985**, *82*, 4699–4706) have probed in greater detail the effect of solvent medium on the spectral exponent for myoglobin azide.

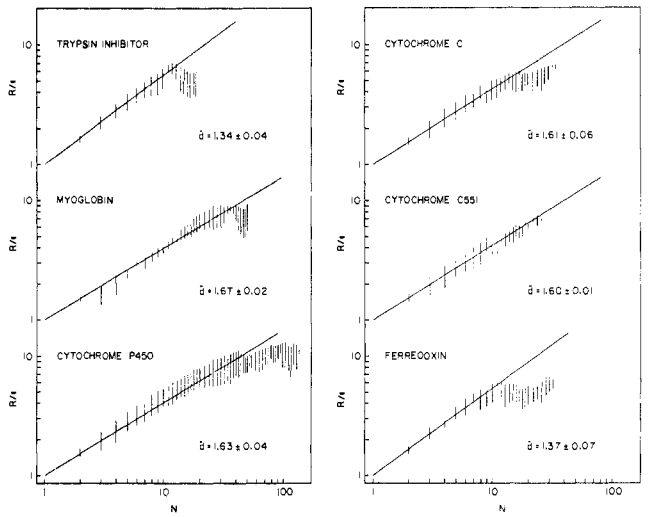

Figure 4. Graphic estimates of fractal dimensions in protein structures. Protein coordinates from data-bank files: bovine pancreas trypsin inhibitor, 3PTI; sperm whale myoglobin, 2MBN; albacore cytochrome *c*, 3CYT, outer chain; *P. aeruginosa* cytochrome *c*-551, 351C; *S. platensis* ferredoxin, 3FXC. Coordinates of *P. putida* cytochrome P450 are from the following unpublished preliminary data set (T. L. Poulos, B. C. Finzel, I. C. Gunsalus, G. C. Wagner, and J. Kraut). Vertical lines are twice the standard deviation in log (R/r). For clarity, some data points are symmetrically deleted for all plots.

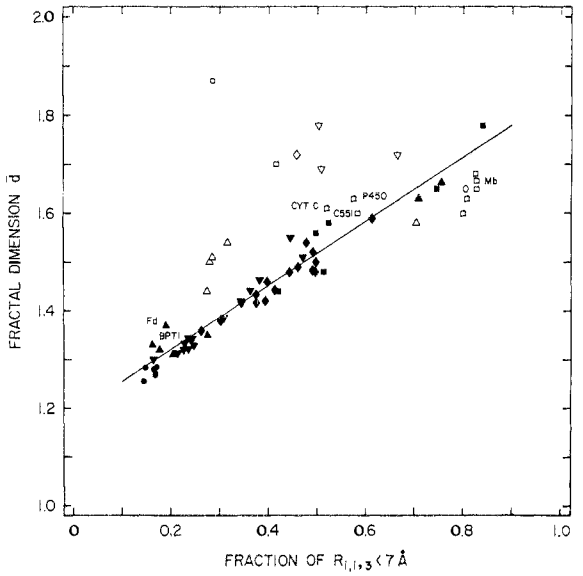

Figure 5. Correlation of fractal dimensions with 70 protein structures. The abcissa is assigned to a structural parameter that estimates the fraction of a protein's tetrapeptides in an α-helical or β-turn conformation, i.e., $R_{i,i+3} < 7$ Å. The least-squares line is drawn with data from the 50 proteins indicated by closed symbols; open symbols exceed the limits of the linear correlation. Symbols: ■, hemeproteins; ▲, iron–sulfur, nonheme iron and copper proteins; ♦, reductases, dehydrogenases, kinases, and related proteins; ▼, proteolytic, glycolytic, nucleolytic, and inorganic hydrolases; ●, immunoglobins, agglutinins, and unrelated structural proteins.

rates from two frozen solutions of low-spin ferric *c*-551, which again differ only in solvent composition, are presented in Figure 3. Qualitatively, from the slopes of the high-temperature data, the spectral exponents of *c*-551 increase with the ionic strength of the solvent medium. The relaxation data for both samples are fit readily to parameters in Table I for a partially bottlenecked direct process. Spectral exponents for *c*-551, although similar to Pd, are not as well-defined. In particular, the analysis of data at the lower ionic strength (see Table I) yields only an estimate of m that is, within the error, at the anticipated lower limit of unity (vide infra).

Fractional dimensions of the proteins analyzed in Table I (or of structures closely related) are estimated in Figure 4 by a refinement of our previous method[2] that enables a more direct correlation with structural detail. Inspection of each of the linear regions of Figure 4 reveals the fractal nature of a protein is determined principally by small values of N (<40). Each of the fractal dimensions derived in Figure 4 represents an index that is proportional to the extent the polypeptide chain fills the Euclidean space occupied by the protein. The smaller values of N that influence the estimate of $\bar{d}$ suggest secondary structures that modulate α-carbon distances $R_{i,i+N}$ should be correlated directly to $\bar{d}$. As one group, the compact structures of α-helices and chain-reversal β-turns are characterized by decreases in the separations between noneighboring α-carbons that should generate increases in $\bar{d}$. Conversely, as a remaining second group, the more extended structures of random chains and ordered β-sheets are characterized by increases in α-carbon distances that should produce decreases in $\bar{d}$. To support this simplified interpretation, some parameter or index that estimates the relative content of these compact and extended structures should provide a correlation with $\bar{d}$ values. Since secondary structures for the majority of proteins in the crystallographic database are not uniformly assigned and reported, a simple and unambiguous method is devised to screen the abundance of available data. The method is based on a stereochemical property of a tetrapeptide segment, namely that the $R_{i,i+3}$ α-carbon separation is nearly always less than 7 Å for both α-helical and β-turn conformations.[20] The fraction of all possible tetrapeptide segments that satisfies this criterion should provide an index that closely approximates the fractional number of residues in either an α-helical or β-turn conformation. Estimates of $\bar{d}$ for 70 globular proteins (single or protomeric chains) are correlated in Figure 5 with the fractional number of tetrapeptide segments with $R_{i,i+3} < 7$ Å. The apparent linear corelation for 50 proteins, within the precision of $\bar{d}$, indicates the fractal nature of polypeptide chains is determined primarily by the relative contributions of compact and extended secondary structures. The most significant deviations from the linear correlation correspond to changes in $\bar{d}$ that are attributable to the addition of repeating structural motifs at large N values. At these N values, the effect of decreasing (increasing) separation distances $R_{i,i+N}$ produces a shift in $\bar{d}$ above (below) the line in Figure 5. The most obvious case in Figure 5 is the example of wheat germ agglutinin (coordinates 0.286, 1.87). The unique structure of this protein (M = 164) is characterized by a conformation of four consecutive and apparently identical 41 residue domains.[21] The four domains are positioned at distances where large values of N (~40) increase the estimate of $\bar{d}$ above the value representing the more localized secondary structures.

Discussion

Earlier studies on two ferric hemeproteins[1,2] indicated the spectral exponents deduced from relaxation data are equivalent to the fractal dimensions estimated from corresponding crystallographic structures. Relaxation measurements on low-spin sperm whale MbOH, MbN_3, and horse cytochrome C yield similar m values (Table I) that correlate remarkably with the $\bar{d}$ estimates (Figure 4) for metMb and albacore cytochrome C, which contain 18 conservative amino acid substitutions from the horse. The magnitudes of m and $\bar{d}$ suggest globular proteins could be modeled by a structure similar to that represented by a self-avoiding random walk. The $\bar{d}$ of a self-avoiding random walk (SAW) in three dimensional space has been approximated by methods in renormalization group theory[22,23] and by numerical simulations.[24,25] The results of these calculations all lie within 2% of the $^5/_3$ value derived from excluded volume arguments.[26] (When restricted

(20) Chou, P. Y.; Fasman, G. D. *J. Mol. Biol.* **1977**, *115*, 135–175.

(21) Wright, C. S. *J. Mol. Biol.* **1977**, *111*, 439–457.
(22) De Gennes, P. G. *Phys. Lett. A* **1972**, *A38*, 339–340.
(23) Le Gillou, J. C.; Zinn-Justin, J. *Phys. Rev. Lett.* **1977**, *39*, 95–98,
(24) Kremer, K.; Baumgartner, A.; Binder, K. *Z. Phys. B: Condens. Matter* **1981**, *B40*, 331–341.
(25) Barr, R.; Brender, C.; Lax, M. *J. Chem. Phys.* **1980**, *72*, 2702–2707.

to two-dimensional space, an SAW $\bar{d}$ decreases further from an unrestricted random walk ($\bar{d} = 2$) to a value approaching $^4/_3$.) The spectral exponent (Table I) for the monoxygenase hemeprotein cytochrome P450 from *P. putida* was also consistent with this argument. A reanalysis[2] of the relaxation data for an Fe_2S_2*Cys_4 type plant ferredoxin from *S. maxima*[27], however, yielded an m value (Table I) that was distinctly lower than the values for hemeproteins. In the interim since our previous reports, crystallographic data of *P. putida* P450 (Poulos, T. L.; Finzel, B. C.; Gunsalus, I. C.; Wagner, G. C.; Kraut, J., unpublished preliminary data) and *S. platensis* Fd[28,29] became available to estimate $\bar{d}$ and to test further the empirical correlation with m. The microbial P450 from the camphor monoxygenase system is the largest protein investigated ($M = 414$) that yields a $\bar{d}$ estimate (Figure 4) in excellent agreement with the spectral exponent deduced from relaxation data (Tale I). The two algal sources provide ferredoxins ($M = 98$) that differ only by four conservative amino acid substitutions.[30] The excellent agreement between m derived from *S. maxima* Fd (Fd = algal ferredoxin) data (Table I) and $\bar{d}$ estimated from *S. Platenis* Fd coordinates (Figure 4) lends further support of the correlation between our fractal models of protein structure and magnetic relaxation. The wide range of fractal dimensions in Figure 5, spanning $1.27 \leq \bar{d} \leq 1.87$, indicates more specific structural information (than that of an arbitrary SAW) is embodied in the index and that it may be confirmed in relaxation measurements.

The relaxation data of Pd (putidaredoxin) and *c*-551 indicate, however, the solvent medium may exert an influence on the spectral exponent that complicates an unambiguous assignment of $\bar{d}$. Spectral exponents (Table I) deduced from Pd and *c*-551 samples in high ionic strength conditions ($I \geq 1$) provide reasonable approximations of the fractal dimensions (Figure 4) estimated for the same *c*-551 and the related Fd. The reduced magnitude in m values with lower ionic strength ($I \leq 0.1$) is unanticipated from the results of previous investigations; at present, the effect appears specific to individual proteins (see Table I). A comprehensive mechanism for the dependence of m on solvent medium will require further investigation. Partial unfolding of these small protein structures—producing a significant change in the electrostatic field at the iron, a lower value of $\bar{d}$, and a concomitant decrease in m—is an unlikely mechanism because the EPR spectra of Pd and *c*-551 did not vary with the solvent medium.[31] We conclude that any structural changes in $\bar{d}$, due to the ionic properties of the solvent, are too small to account for the large changes observed in m. The present findings do suggest that changes in the bound hydration shell of a protein may impart a considerable effect on the fractal model of protein dynamics.

Precise relationships between the spectral exponent and the fractal dimension have been considered theoretically by various groups; the results are always of the general form

$$m = \bar{d}/a \qquad 1 \leq m \leq \bar{d} \tag{7}$$

Rammal and Toulouse[32] assign m to a *spectral dimension* and interpret a as the low-frequency scaling exponent. Alexander and Orbach[33] refer to m as the *fracton dimension* and replace the coefficient a by $(1 + \bar{\delta}/2)$, where $\bar{\delta}$ is the diffusion constant scaling exponent. Helman et al.[34] replace a by $d_w/2$, where d_w is the fractal dimension of an unrestricted random walk on a fractal structure. According to the Helman et al. theory, a random walk on an unbranched polymer chain of dimension $\bar{d}$ is characterized by a d_w value of $2\bar{d}$. In this case the spectral exponent reduces to unity (eq 7) and is independent of $\bar{d}$. The obvious parallel is evident in the relaxation dynamics observed with Pd and *c*-551 in low ionic strength. If, on the contrary, the unbranched chain incorporates numerous massless bridges, such as hydrogen bonds or other noncovalent interactions that connect distant elements of the polypeptide chain, then a random walk over this structure becomes unrestricted and should approach a d_w value of 2 (vide supra). In this situation the spectral exponent reduces (eq 7) to the $\bar{d}$ of the backbone structure. Depending on this degree of off-chain connectivity, the correlation emerges between the highest value of m from relaxation data and the $\bar{d}$ estimate from a protein structure.

Aside from the structural elements that influence m in the relaxation process, the present findings indicate the dynamic properties of a protein merit equal consideration. The vibrational modes that are active in the Raman mechanism must be of high energy, in comparison to $g\beta H \sim 0.3$ cm^{-1}, and numerous to compensate for the weak, second-order coupling. The modes must also be delocalized, since the probability of the spin interacting with two nearby and localized vibrations with the proper frequency difference would be vanishingly small. An estimate of ν_{max} (eq 1) for the delocalized vibrations of a polypeptide backbone may be obtained from $\upsilon_s/2r$. With a velocity of sound of 2×10^5 cm/s, a crude approximation of ν_{max} for propagating backbone vibrations in proteins is $\sim$2600 GHz ($\sim$87 cm^{-1}). A more detailed account of the frequencies that contribute to the Raman rate is described by the complete integral expression,[1,2] where $\rho(\nu)$ is given by eq 1. An exact computation of the integral expression for $m = {}^5/_3$ leads to a Raman relaxation rate that decreases less than 1% from the predicted T^{3+2m} dependence as ν_{max}/T decreases to 333 GHz K^{-1}. For typical values of $m = 5/3$ and temperatures of 3–12 K, the Raman process probes the frequencies between 100 and 4000 GHz that are matched conveniently to low-frequency vibrational modes in proteins (3–120 cm^{-1}). Remembering that the temperature variation of the Raman rate (eq 2) follows directly from the functional form that describes $\rho(\nu)$, a T^n relationship implies a power law for $\rho(\nu)$. All available theoretical and experimental evidence indicate eq 1 is the proper form of $\rho(\nu)$ for the low-frequency backbone vibrations in proteins. The limits given by eq 7 indicate further that $\rho(\nu)$ varies between a constant and $\nu^{\bar{d}-1}$. In bovine pancreas trypsin inhibitor, for example, an estimated $\bar{d} = 1.34 \pm 0.04$ (Figure 4) predicts an upper limit of $\rho(\nu)$ that varies as $\nu^{0.34\pm0.04}$.

The prediction made by a fractal model of protein dynamics in BPTI (bovine pancreas trypsin inhibitor) agrees well with a reported distribution of calculated, low-frequency vibrational modes.[35] The essence of the Gō et al. calculation (in vacuo) considers bond lengths and angles fixed and treats only dihedral angles of the main chain and the side chains as dynamic variables. Since each α-carbon position of BPTI ($M = 58$) is described by two dihedral angles, the main chain constitutes 116 degrees of freedom that correlate with about 116 delocalized modes of the backbone. The results of their normal mode analysis indicate vibrations below 120 cm^{-1} (3600 GHz) are predominantly delocalized, i.e., they correspond to concerted movements of the entire polypeptide backbone (including side chains). Since the lowest 117 vibrational modes are below 75 cm^{-1}, we estimate a ν_{max} of 2250 GHz for the main chain of BPTI. A histogram plot of the distribution of normal modes within the range of $0 \leq \nu \leq 5100$ GHz, which corresponds to the lowest 201 modes out of a 241 total, is reproduced in Figure 6 from the Gō et al. study.[35] Superimposed on this plot are normalized density of states predicted by the fractal model, $\rho(\nu) \propto \nu^{0.34}$, and by the classical Debye model

(26) Flory, P. J. "Principles of Polymer Chemistry", Cornell University Press: Ithaca, 1953.
(27) Gayda, J. P.; Bertrand, P.; Deville, A.; More, C.; Roger, G.; Gibson, J. F.; *Cammack, R. Biochem. Biophys. Acta* **1979**, *581*, 15–26.
(28) Fukuyama, K.; Hase, T.; Matsumoto, S.; Tsukihara, T.; Katsube, Y.; Tanaka, N.; Kakudo, M.; Wada, K.; Matsubara, H. *Nature (London)* **1980**, *286*, 522–524.
(29) Tsukihara, T.; Fukuyama, K.; Nakamura, M.; Katsube, Y.; Tanaka, N.; Kakudo, M.; Wada, K.; Hase, T.; Matsubara, H. *J. Biochem. (Tokyo)* **1981**, *90*, 1763–1773.
(30) Matsubara, H.; Hase, T.; Wakabayashi, S.; Wada, K. In "The Evolution of Protein Structure and Function", Sijman, D. S.; Branzier, M. A. B., Eds.; Academic Press, Inc.: New York, 1980; pp 245–266.
(31) According to eq 3 for an $N = M - 1 = 81$, a reduction of $\bar{d}$ for cytochrome *c*-551 from 1.43 to unity would require that the protein increase a factor of 3.75 from its original size.
(32) Rammal, R.; Toulouse, G. *J. Phys., Lett.* **1983**, *44*, L13–L22.
(33) Alexander, S.; Orbach, R. *J. Phys. Lett.* **1982**, *43*, L625–L631.
(34) Helman, J. S.; Coniglio, A.; Tsallis, C. *Phys. Rev. Lett.* **1984**, *53*, 1195–1197.
(35) Gō, N.; Noguti, T.; Nishikawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3696–3700.

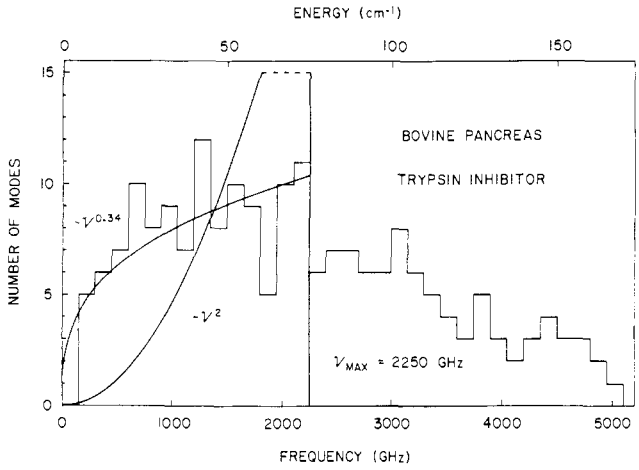

Figure 6. Histogram of the lowest 201 vibrational modes in bovine pancreas trypsin inhibitor. Data reproduced from Gō et al.[35] Superimposed are normalized densities of vibrational states (eq 1) varying as ν^2 (Debye model) and $\nu^{0.34}$ (fractal model). The latter $\rho(\nu)$ is derived from the estimated d for BPTI, assuming the spectral exponent is equal to the fractal dimension.

for a regular three-dimensional solid, $\rho(\nu) \propto \nu^2$. Inspection of Figure 6 demonstrates convincingly that the fractal model provides a more accurate description of the low-frequency variation of modes. (An unweighted nonlinear least-squares analysis of the data between $0 \leq \nu \leq \nu_{max}$ yields a best-fit normalized $\rho(\nu) \propto \nu^{0.24}$.) We conclude that over this magnitude of frequencies some of the details from a complex protein dynamics calculation may be anticipated from an estimate of the fractal dimension. At frequencies above the estimated ν_{max}, the results of Gō et al. and of a more detailed study by Brooks and Karplus[36] both show that the distribution of modes decreases toward a nearly constant value (see Figure 6). Since vibrations at these higher frequencies correspond to more localized modes, they are outside the scope of both the fractal model and the vibrations that contribute to the Raman relaxation mechanism.

In summary, fractal models of protein structure and dynamics provide the best explanation for the diversity observed in the temperature dependence of the Raman relaxation rates in paramagnetic proteins. The fractal nature of a protein is characterized by a fractal dimension—a conformational index sensitive to secondary structures and supra-secondary domains—that defines the distribution of vibrational states associated with the collective motions of the polypeptide backbone. Future studies will concentrate on additional paramagnetic proteins, either natural or spin-labeled, that extend the experimental verifications of the proposed fractal models.

Acknowledgment. This research was supported in part by the U.S. Public Health Service and NIH Grants GM-24488 and AM00562 (GCW). Computer facilities of the Materials Research Laboratory were used in this study.

Registry No. Reductase, 9037-80-3; dehydrogenase, 9035-82-9; kinase, 9031-44-1; hydrolase, 9027-41-2; cytochrome P450, 9035-51-2; cytochrome *c*, 9007-43-6; cytochrome *c*-555, 9048-77-5; trypsin inhibitor, 9035-81-8; pancreatic basic trypsin inhibitor, 9087-70-1.

Supplementary Material Available: Table listing protein fractal dimensions for the 70 proteins described in Figure 5 (3 pages). Ordering information is given on any current masthead page.

(36) Brooks, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6571-6575.

(37) Earlier work on *P. putida* cytochrome P450[38] indicated n and m are essentially independent of changes in the magnetic field position. The variation in n that was observed from measurements at three magnetic fields, which correspond to the three principal axes of the P450 **g** tensor (2.42, 2.25, 1.92), is comparable to the precision (±0.06) of a single determination of the temperature exponent of the Raman rate. Relatively small changes in the position of the magnetic field are thus inconsistent with the large differences in n and m that are observed in the present C551 data.

(38) Herrick, R. C., Stapleton, H. J. *J. Chem. Phys.* **1976**, *65*, 4778-4785.

Fourier Transform Infrared Spectra of Conjugated Diene and Cumulene Complexes with HF in Solid Argon

Kenneth O. Patten, Jr.,[†] and Lester Andrews*

Contribution from the Department of Chemistry, University of Virginia, Charlottesville, Virginia 22901. *Received April 1, 1985*

Abstract: Codeposition of conjugated diene and substituted conjugated diene samples with HF produced new absorptions due to ν_s(HF) stretching, ν_l(HF) librational, and ν^c(base) modes in the complexes. Comparison to similar alkene complexes suggests a structural model with HF on the C_2 axis of butadiene and the acid hydrogen oscillating between the regions of maximum π electron density of both π bonds. On the other hand, the degeneracy in the allene CH_2 wagging mode is lifted in the allene–HF complex, which provides evidence for hydrogen bonding of HF to one π bond in allene.

The hydrogen-bonding interaction has attracted considerable attention in recent years. Infrared studies of a wide variety of gas mixtures and matrix isolated samples have been performed.[1-4] Rotational spectra and structural data have been obtained by using pulsed expansion of gases from a supersonic nozzle and Fourier-transform microwave spectroscopy,[5,6] molecular beam electric resonance,[7,8] and direct observations upon gas mixtures for stronger complexes.[9,10] Predissociation and photodissociation spectra for HF complexes and polymers have been obtained,[11,12] as has rotational fine structure for these species using infrared laser

[†] Undergraduate research student.

(1) Couzi, M.; LeCalve, J.; Van Huong, P.; Lascombe, J. *J. Mol. Struct.* **1970**, *5*, 363.
(2) Thomas, R. K. *Proc. R. Soc. London, A* **1975**, *344*, 579.
(3) Barnes, A. J. *J. Mol. Struct.* **1983**, *100*, 259.
(4) Andrews, L. *J. Phys. Chem.* **1984**, *88*, 2940.
(5) Legon, A. C.; Aldrich, P. D.; Flygare, W. H. *J. Chem. Phys.* **1981**, *75*, 625.
(6) Legon, A. C. *Annu. Rev. Phys. Chem.* **1983**, *34*, 275.

0002-7863/85/1507-5594$01.50/0