An Approach to the Mapping of Internal Motions in Proteins. Analysis of ^{13}C NMR Relaxation in the Bovine Pancreatic Trypsin Inhibitor

Anthony A. Ribeiro, Roy King, Charles Restivo, and Oleg Jardetzky*

Contribution from the Stanford Magnetic Resonance Laboratory, Stanford University, Stanford, California 94305. *Received May 29, 1979*

Abstract: A general formalism for estimating molecular parameters characterizing the complex motions of proteins and other flexible macromolecules from NMR relaxation measurements is illustrated for the case of ^{13}C NMR relaxation in the bovine pancreatic trypsin inhibitor (BPTI) (mol wt 6500). Specifically, T_1, T_2, and NOE at 45 and 90 MHz have been measured for 40 assigned or partially identified protonated carbon resonances in the methyl, methylene, methine, and aromatic regions of the ^{13}C spectrum of BPTI. Accurate information on the protein motional frequencies and less precise information on the relative amplitudes of each motion are obtained from the general formalism based on Markov processes. A minimum of three motions at each carbon group are *required* to account for the six experimental parameters measured at two field strengths. Low-frequency components make a small but finite contribution to the relaxation of all resonances, suggesting a general low-frequency distortion of the backbone. Rotational diffusion of the protein makes a relatively minor contribution to the relaxation process. For aliphatic groups, rotation of side chains dominates the relaxation process.

Introduction

The study of internal motions in proteins is of considerable importance for the understanding of their function. Relatively rare is the case of a protein that could function as a rigid entity—enzyme inhibitors and some simple enzymes fall perhaps into this class. For allosteric proteins, antibodies, and proteins functioning as part of a control mechanism, structural rearrangement and hence at least some degree of flexibility and motion are an integral part of the function.

The existence of some type of fluctuation in the protein structure has been inferred long ago from hydrogen-deuterium exchange.[1] Fluorescence depolarization measurements[2] and unusual fluorescence quenching[3] have provided additional indication of internal mobility, and so have even the earliest NMR experiments on proteins.[4] Extensive fluctuations in the protein structure are predicted theoretically by Karplus.[5] Little, however, is known experimentally about the details of such motions.

High-resolution NMR is in principle a very powerful method for the study of this problem, first because it permits simultaneous observation of spectral lines from many different residues in the protein, and second because all of the parameters measurable on each line are to some extent sensitive to motion. Numerous efforts have therefore been made to derive information on internal motions in proteins from NMR and particularly from relaxation measurements.[4b,6-8]

The principal problem encountered in this type of study is that of interpretation. The theoretical models available for the analysis of relaxation data[9,10] are too crude to reflect internal motions in proteins in all of their potential complexity. The significance of the correlation times calculated from the various alternative equations therefore remains unclear. A general theoretical framework for analyzing relaxation data on macromolecules with several degrees of internal motional freedom has only recently been developed,[11] although several more limited generalizations of existing theory had been proposed earlier.[4b,8b,12]

In the present study, we have been concerned with the fundamental question: How much definitive information on internal mobility in proteins can be deduced from NMR relaxation measurements? To this end, we have carried out an extensive study of relaxation on several proteins and a comparison of different methods of theoretical analysis. Model calculations have previously led us to the following conclusions: (1) The rates of individual motions can be identified with reasonable accuracy, treating the analysis of relaxation data as an eigenvalue problem.[11a] (2) On the other hand, precise information on the nature and the amplitude of each motion is not contained in an individual NMR relaxation parameter since alternative specific models often account for the data equally well.[11b,c] The operator formalism which yields the time constants of the motion as eigenvalues but provides only *relative* amplitudes therefore provides as much unique information as can be extracted from the measurements. This is illustrated here by an analysis of carbon relaxation in the relatively rigid protein bovine pancreatic trypsin inhibitor (BPTI). Correlation of ^{13}C and ^{1}H relaxation in this protein is presented elsewhere.[11d,e]

Experimental Section

The bovine pancreatic trypsin inhibitor (BPTI, Trasylol, registered trademark of Farbenfabriken Bayer AG) was obtained from Bayer AG, West Germany. The protein was further purified over Sephadex columns. The ^{13}C NMR studies were carried out on a 9.8 mM sample of BPTI at pD 5 with 10^{-4} M EDTA and 50 mM NaCl buffer. The chemical shifts were referenced to Me_4Si as an external reference.

^{13}C relaxation data at 90.5 MHz were obtained on the modified Bruker HXS-360 spectrometer at Stanford over a sweep width of ±9100 Hz, using 16K channels and a 90° pulse of 22 μs. ^{13}C relaxation data at 45 MHz were obtained on the Bruker HXS-180 wide-bore instrument at the University of California at Berkeley. On this system, spectra covering a sweep width of ±5000 Hz were accumulated into 8K channels and the 90° pulse was 10 μs.

To minimize radio-frequency pulse defects, the T_1 relaxation measurements were made with phase alternation of the 90° pulse by 180° on every repetition of the standard 180°-τ-90° inversion recovery pulse sequence.[13] The recovery time between pulses was 2.7 s. T_2 values were measured by the Hahn spin-echo method (90°-τ-180°-delay) with a recovery time of 1.5 s. NOE experiments on the HXS-360 instrument were carried out by obtaining four scans with the decoupler turned on all the time, and then four scans with the decoupler off during a delay time of 2.2 s and on during acquisition. Recycling through this loop 2800 times yielded two sets of spectra that were fully decoupled, one with and one without NOE.[6b,14]

In attempting the ^{13}C T_1 measurements at 90.5 MHz, we found that a 180° pulse of 44 μs was not adequate to give total inversion over the large sweep widths of 20 kHz for observation of ^{13}C NMR spectra at the high field. Intensity distortions in particular were noted for large spectral offsets from the quadrature phase detection (QPD) frequency set in the center of the ^{13}C spectrum. To bypass this problem of finite pulse excitation power, the relaxation measurements over the aromatic

0002-7863/80/1502-4040$01.00/0

and aliphatic regions were made as two separate experiments. In the first experiment, the spectrum offset was adjusted so that the aliphatic region fell near the QPD frequency; in the second, the aromatic region was offset near the QPD frequency. Reliable inversion was observed in these cases, and semilogarithmic plots of peak intensities gave linear fits. The T_1 data were also fitted to the three-parameter curve-fitting routine of Nicolet software.[15]

The natural abundance ^{13}C NMR relaxation time measurements of this paper on 10 mM protein typically required 5-10 000 accumulations per free induction decay. To ensure protein stability over the long periods of time necessary for the relaxation measurements, the experiments were purposely carried out at the relatively low temperature of 17 °C. For ^{13}C NMR studies of aqueous protein solutions, the heating effects that accompany proton noise decoupling are generally a function of the buffer ions rather than the dipolar protein molecules.[16] For this reason, the amounts of added salt and EDTA were limited to the 10^{-2}-10^{-4} M concentration region.

Heating effects are also more pronounced at higher field strengths.[16] On the 360-MHz spectrometer, the decoupling power was set to not exceed 5 W. In addition, a high cooling rate of nitrogen flow was employed to ensure minimal heating effects. The temperature control on this instrument is estimated at 17 ± 1 °C. On the Bruker 180-MHz spectrometer, two proton decouplers were available and were employed alternately using a scheme similar to that of Levy et al.[17] One decoupler was used at a low power level during the waiting time to ensure the potential NOE, and a second decoupler was then gated up to 2 W during the data acquisition. With a relatively low duty cycle in the experiment, the sample heating was minimal. Digital readout of the probe temperature in the course of the relaxation measurements at 45 MHz was 16.7 ± 0.3 °C.

Theoretical

The fundamental relations which permit the analysis of measured NMR relaxation parameters T_1, T_2, and NOE in terms of molecular motion assume that the motion is stochastic in nature and may be characterized by a correlation function $\overline{F(t)F(t+\tau)}$, which measures the persistence of a given spatial configuration of the moving system, given by a time-dependent function of the coordinates, denoted by $F(t)$. The relaxation rates (T_1^{-1}, etc.) are assumed to be proportional to the Fourier transform of the correlation function, the spectral density function $J(\omega)$

$$J(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} |\overline{F(t)F(t+\tau)}| \, d\tau \qquad (1)$$

A particularly simple form of the relationship of T_1, T_2, and $J(\omega)$ applies to ^{13}C nuclei with a proton attached to them. The relaxation is dipolar, dominated by the proton, and the internuclear distance is fixed by the covalent bond at $R_{CH} = 1.09$ Å.[18] The relationship can then simply be written as

$$1/T_1 = K_1 R_{CH}^{-6}\{J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)\} \qquad (2)$$

$$1/\pi T_2 = K_2 R_{CH}^{-6}\{4J(0) + 3J(\omega_C) + J(\omega_H - \omega_C) + 6J(\omega_H) + 6J(\omega_H + \omega_C)\}$$

$$\mathrm{NOE} = 1 + \frac{6J(\omega_H + \omega_C) - J(\omega_H - \omega_C)}{J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)} \frac{\omega_H}{\omega_C}$$

where K_1 and K_2 are proportionality constants and ω_H and ω_C are the proton and carbon frequencies.

The *relative* simplicity of interpretation using these expressions accounts for the widespread use of ^{13}C relaxation measurements in the study of molecular motion and for our use of them in the present report. The remaining major theoretical problem, however, is to find an appropriate form of $J(\omega)$, taking into account different motions that may occur in a protein. The original formulation of NMR theory derived $J(\omega)$ for the diffusion of a rigid sphere,[9b] and subsequent modifications include anisotropic diffusion of a rigid body,[9c] a spinning top on a rigid sphere,[9d] or a sequence of rotations in an aliphatic side chain.[12] Such conceptions are inadequate for the analysis of data in terms of potentially complex motions, such as may occur in proteins.

On the very general assumption that each of the motions contributing to relaxation in a macromolecule is a Markov process (see Appendix for a full discussion), it is possible to derive a relationship between $J(\omega)$ and λ, the rate parameters, and ϕ, the amplitude parameters of individual motions for an arbitrary number of motions of an arbitrary nature[11a]—i.e.

$$J_F(\omega) = -2 \sum_{n_1,n_2,\cdots,n_M} \frac{\left|\left\langle F, \prod_{k=1}^{M} \phi_{kn_k} \right\rangle\right|^2 \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)}{\omega^2 + \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)^2} \qquad (3)$$

where λ are the eigenvalues and ϕ the eigenfunctions of the transition operator Ω—i.e.

$$\Omega\phi = \lambda\phi \qquad (4)$$

For the diffusion of a rigid sphere, eq 3 reduces to the usual rigid rotor relaxation model[11c] and $1/\lambda$ can be given the simple meaning of a rotational correlation time related to the diffusion coefficient by the Stokes-Einstein relation.[9a]

The formulation given in eq 2-4 makes it possible to devise a procedure for systematically determining the *number* of motions that are necessary to account for a given set of relaxation measurements and their *rates*. This can in fact be done in one of two ways: (1) One can sequentially test models that assume specific motions, thus defining the eigenfunctions ϕ arbitrarily and, a priori, for one, two, three, or more motions. The sets of relaxation parameters calculated from each model can then be compared to the experimental data *and to each other*. A limited comparison of this type has been reported,[11b] leading to the conclusion that, in many cases, different models can account for the same set of data equally well, while in others discrimination between different models is possible, given a sufficiently large set of experimentally measured values. Alternatively, (2) one can sequentially analyze the experimental data in terms of one, two, or more motions, distinguishable by their frequency, and treating the eigenfunctions ϕ as unknowns. This type of analysis is reported here.

Following the procedure outlined in the Appendix, we can derive an $\bar{\alpha}_i$ and a $\bar{\lambda}_i$ such that

$$J_F(\omega) = \sum_{i=1}^{M} \frac{\bar{\alpha}_i \bar{\lambda}_i}{\bar{\lambda}_i + \omega^2} \qquad (5)$$

where $\bar{\alpha}_i$ and $\bar{\lambda}_i$ are the *effective* amplitude and rate for each motion. While $\bar{\lambda}_i = 1/\tau_i$ has a simple physical meaning, $\bar{\alpha}_i$ does not. If all motions were commensurate—e.g., all involving a rotation through an angle θ in a plane—$\bar{\alpha}_i$ would be simply the relative amplitude of each motion, i.e., $\theta_i/\Sigma_i\theta_i$. In the case of complex motions of the vector of fixed length, R, for which the polar angle ϕ and azimuthal angle θ are not known in detail, $\bar{\alpha}_i$ measures the relative contribution each motion makes to the amplitude of the second spherical harmonic used to describe the rotational motion in general terms. The geometrical interpretation of $\bar{\alpha}_i$ will be discussed in detail elsewhere. Suffice it to say here that its magnitude depends on both the amplitude of the individual motion and on the amplitude of all motions contributing to the relaxation of a given ^{13}C nucleus by its neighboring proton. It is a direct measure of the relative contribution each motion makes to the measured relaxation rates.

The number of relaxation measurements required to solve eq 5 is $N = 2M - 1$, where M is the number of independent motions. The analysis of relaxation data on a macromolecule presumed to be flexible requires that (1) the dominant relaxation mechanism be known; (2) a sufficient number of relaxation parameters be determined to permit an analysis in terms other than the simplest models (three relaxation parameters

Table I. BPTI-CH_3 Motional Frequencies (λ) and Amplitudes (α)[a]

peak/shift, ppm assigned residue(s)	field strength, MHz	T_1, s	T_2, s	NOE	σ^2	motion 1		motion 2		motion 3	
						λ_1, Hz	α_1, %	λ_2, Hz	α_2, %	λ_3, Hz	α_3, %
A	45	0.255	0.115	2.75	0.010	6E8	2	1E7	1	2E10	97
7.56											
$Ile^{18}/Ile^{19}\delta$	90	0.370	0.135	2.45	0.031	6E8	2	1E8	3	3E10	94
B	45	0.235	0.055	1.55	0.000	6E8	7	7E7	8	3E11	85
11.27											
$Ile^{18}/Ile^{19}\delta$	90	0.335	0.090	1.50	0.003	6E8	4	2E8	13	1E11	83
C	45	0.320	0.070	1.75	0.001	6E8	2	8E7	7	8E10	91
13.13											
Met^{52}-SCH_3	90	0.350	0.090	1.75	0.000	6E8	8	7E7	7	7E10	85
D	45	0.300	0.105	1.95	0.000	6E8	6	4E7	2	7E10	92
14.31											
$Ile\gamma/Leu^{29}\delta$	90	0.335	(0.090)	2.40	0.000	6E8	5	3E7	3	3E10	92
E	45	0.130	0.105	2.30	0.083	6E8	11	5E9	10	2E10	79
14.77											
Ala^{48}	90	0.280	0.125	2.80	0.008	6E8	2	2E7	1	2E10	97
F	45	0.190	0.085	2.40	0.000	6E8	6	3E7	1	2E10	94
15.51											
$Ile\gamma$	90	0.270	(0.090)	2.50	0.068	6E8	1	1E8	8	2E10	91
G	45	0.150	0.075	3.00	0.026	6E8	1	1E7	<1	1E10	98
16.02											
Ala^{25}	90	0.220	0.070	2.45	0.014	6E8	7	1E7	1	2E10	92
HIJ	45	0.275	0.080	2.60	0.089	6E8	1	4E7	2	2E10	97
16.71											
$Ala^{16,27,58}$	90	0.240	0.100	2.20	0.129	6E8	10	4E7	2	2E10	87
K	45[b]	0.200	0.090	2.80	0.403	6E8	5	2E7	0.5	2E10	94
17.13											
Ala^{40}	90[b]	0.225	0.100	2.65	0.242	6E8	7	1E7	1	2E10	92
LM	45	0.175	0.060	2.45	0.078	6E8	7	2E7	1	2E10	92
18.36											
Thr^{11}(Leu^6/Leu^{29}/Val)	90	0.195	0.080	2.30	0.048	6E8	11	1E7	1	2E10	88
NOP	45	0.165	0.090	2.10	0.035	6E8	10	2E7	1	3E10	89
19.57											
Thr^{32},Thr^{54}(Leu^6/Leu^{29})	90	0.190	0.095	2.25	0.039	6E8	12	1E7	0.5	2E10	88
Q	45	0.250	0.085	2.10	0.001	6E8	6	4E7	2	4E10	92
20.12											
Leu^6	90	0.255	0.100	2.60	0.000	6E8	4	1E7	0.5	2E10	95
RST	45	0.180	0.070	2.10	0.000	6E8	9	3E7	2	3E10	89
22.01											
Val/Leu/Ile	90	0.190	(0.065)	2.20	0.000	6E8	12	1E7	1	2E10	87

[a] Relaxation data at 17 °C, pD 5. Chemical shifts are referenced to external Me_4Si. The experimental T_1, T_2, and NOE determinations in any single run fluctuate by ±10%. Since the calculated error of fits as measured by the square of the variance (σ^2) is considerably smaller than the experimental error, the relaxation values are rounded off to the nearest five or ten in the third figure and may be said to represent both the experimental and calculated values, except where noted. E represents a power of 10. [b] Calculated T_1, T_2, and NOE for peak K (45 MHz, $\sigma^2 = 0.403$) are 0.192 s, 0.090 s, and 2.40; calculated NOE for peak K (90 MHz, $\sigma^2 = 0.242$) is 2.40.

measured at two or three frequencies may be regarded as a minimum); (3) sequential fitting of the data be carried out, assuming first a single motion and, in succession, two, three, or more motions. The last step can be simplified and the number of required NMR measurements reduced if one of the motions can be studied and its parameters determined by another method. This is usually possible for the overall diffusional rotation of a macromolecule using depolarized light scattering techniques.[19] The parameters of this motion can then be introduced as knowns into the combined system of eq 3-5, the only remaining unknowns being the parameters of internal motions. (4) If the number of experimental measurements is not sufficient to obtain exact solutions of the system of equations, the error of fitting needs to be evaluated and compared for different fits.

A computer program was written in Fortran to carry out this type of analysis, based on the algorithm presented in the Appendix, part B. The experimental parameters T_1, T_2, and NOE are entered in the form $1/T_1$, $1/T_2$, and $(NOE - 1)/T_1$, after normalization for the number of attached protons. Equation 5 is solved numerically by assuming a given number of motions and various frequencies λ_k, and searching for the optimal amplitudes $\bar{\alpha}_k$, subject to the linear condition $\Sigma\bar{\alpha}_k = 1$. To reduce the search time for the computer analysis, $\lambda_1 = 6 \times 10^8$ s^{-1}, describing the rotational diffusion of BPTI, was determined by depolarized light scattering and is used as an additional constraint in the analysis. The other motional frequencies (λ_2 and λ_3 in the present analysis) were then sequentially varied from 10^7 to 10^{12} s^{-1} and values of α_k to fit eq 5 searched for. The algorithm was structured to look for convergent fits for any *combination* of T_1, T_2, and NOE which were assigned limits of accuracy of 20, 30, and 20%, respectively. The constraint was also imposed that the calculated error of fit be smaller than experimental error (cf. Appendix). The precision

Table II. BPTI—Typical CH_2 Motional Frequencies (λ) and Amplitudes (α)[a]

peak/shift, ppm assigned residue(s)	field strength, MHz	T_1, s	T_2, s	NOE	σ^2	motion 1 λ_1, Hz	motion 1 α_1, %	motion 2 λ_2, Hz	motion 2 α_2 %	motion 3 λ_3, Hz	motion 3 α_3, %
CH_2 1 33.29 Glu[7,49]λ CH_2	45	0.115	0.075	2.00	0.000	6E8	9	2E8	12	1E10	79
	90	0.160	(0.070)	1.85	0.020	6E8	21	1E8	24	1E10	55
CH_2 2 36.80 A^3 or Asp[50] β CH_2	45	0.110	0.030	1.67	0.000	6E8	25	5E7	16	3E10	58
	90	0.120	0.040	1.55	0.026	6E8	41	1E8	23	1E10	36
CH_2 3 36.50 Asp^3 or Asp[50] β CH_2	45	0.070	0.035	1.50	0.001	6E8	16	2E8	36	2E10	48
	90	0.135	0.055	2.00	0.001	6E8	15	2E8	23	7E9	62
CH_2 4 39.36 Lys ϵ CH_2	45	0.300	0.055	2.10	0.001	6E8	7	3E7	3	3E10	90
	90	0.295	0.065	2.20	0.107	6E8	14	1E7	2	2E10	84

[a] Relaxation data at 17 °C, pD 5. Chemical shifts are referenced to external Me_4Si. The experimental T_1, T_2, and NOE determinations in any single run fluctuate by ±10%. Since the calculated error of fits as measured by the square of the variance (σ^2) is considerably smaller than the experimental error, the relaxation values are rounded off to the nearest five or ten in the third figure and may be said to represent both the experimental and calculated values, except where noted. E represents a power of 10.

Table III. BPTI—Backbone α-CH Motional Frequencies (λ) and Amplitudes (α)[a]

peak/shift, ppm assigned residue(s)	field strength, MHz	T_1, s	T_2, s	NOE	σ^2	motion 1 λ_1, Hz	motion 1 α_1, %	motion 2 λ_2, Hz	motion 2 α_2, %	motion 3 λ_3, Hz	motion 3 α_3, %
α-CH 1[c] 51.14 unknown	45	0.100	0.040	1.25	0.105	6E8	11	2E8	50	1E8	39
	45[b]	0.100	0.040	1.25	1.602	6E8	28	9E8	68	1E10	2
	90[b]	0.200	0.070	1.15	2.783	6E8	35	2E8	52	5E11	13
α-CH 2 51.74 Ala^{58} α-CH	45	0.120	0.060	1.45	0.003	6E8	23	2E8	41	2E10	36
	90[b]	0.220	0.080	1.20	0.449	6E8	23	2E8	62	9E11	15
α-CH 3 52.13 unknown	45[b]	0.100	0.035	1.20	3.054	6E8	13	1E8	80	6E9	6
	90[b]	0.250	0.055	1.15	0.450	6E8	4	2E8	83	9E11	13
α-CH 4[c] 53.55 unknown	45	0.105	0.045	1.30	0.036	6E8	2	2E8	74	3E10	24
	90	0.205	0.55	1.20	0.000	6E8	7	3E8	78	5E7	15

[a] Relaxation data at 17 °C, pD 5. Chemical shifts are referenced to external Me_4Si. The experimental T_1, T_2, and NOE determinations in any single run fluctuate by ±10%. Since the calculated error of fits as measured by the square of the variance (σ^2) is considerably smaller than the experimental error, the relaxation vales are rounded off to the nearest five or ten in the third figure and may be said to represent both the experimental and calculated values, except where noted. E represents a power of 10. [b] Calculated T_1, T_2, and NOE for α-CH 1 (45 MHz, σ^2 = 1.602): 0.115 s, 0.035 s and 1.35; (90 MHz, σ^2 = 2.783) 0.220 s, 0.060 s, and 1.25. For α-CH 2 (90 MHz, σ^2 = 0.449): 0.245 s, 0.075 s, and 1.20. For α-CH 3 (45 MHz, σ^2 = 3.054): 0.120 s, 0.025 s, and 1.30; (90 MHz, σ^2 = 0.450) 0.275 s, 0.050 s, and 1.20. [c] For α-CH 1 slow motions are observed for λ_1 and λ_2 at 45 and 90 MHz. For λ_3 at 45 MHz alternate solutions of a slow and a fast component exist. The 90-MHz data yield a fast component at λ_3. This situation in which alternate fits are possible suggests that that a three-motion analysis may not be adequate to describe the motion of this group. Discrimination might be possible with a four-motion analysis.

of an individual T_1, T_2, and NOE determination in a single run is usually better than the chosen values, i.e., of the order of ±10%. However, our cumulative experience with duplicate runs on the same as well as on different instruments has led us to the conclusion that the *accuracy* of relaxation measurements on macromolecules at the present state of NMR technology is no better than the stated figures. Similarly, the precision of the calculation is considerably greater than the accuracy of the experimental parameters. The values of T_1, T_2, and NOE reported in Tables I–IV therefore represent *both the experimental and the calculated* values of the relaxation parameters. In all cases it has been possible to find a fit in which the calculated values correspond *exactly* to the experimental values to the last experimentally significant figure. The corresponding errors of fit, given in the tables as the sum of the squares of the variances of the individual parameter fits normalized to the experimental error (cf. Appendix, eq 8′), are in most cases <0.01. In Table V we give data for two peaks to show that alternative fits with a correspondingly larger error constitute a cluster of $\bar{\lambda}_i$ within a factor of 2-3 and $\bar{\alpha}_i$ within 20%. Since the accuracy of the experimental parameters, rather than the precision of the calculated values, is limiting the agreement between the two, we consider these factors to be a present-day limit for the determination of molecular structural or dynamic parameters from NMR measurements. Claims of greater accuracy are uniformly based on inadequate data and oversimplified models and are for this reason misleading. If our experimental error limits represent an overestimate, the smallness of the calculated error of fit for the best fit indicates that the range of allowed values of the molecular parameters would be much narrower. It should be noted, however, that even within the stated limits of uncertainty this type of analysis provides much more information on the rates and relative importance of different molecular motions than can be obtained by any other method. As the accuracy of the experimental parameters improves, so will the accuracy of the calculated molecular parameters.

A check of the validity of using the value of 6×10^8 s^{-1} for the overall rotational correlation time of BPTI was provided by the initial calculations when the erroneous value $\tau_c \approx 2 \times 10^{-8}$ s reported by Wüthrich and Baumann[7a] was used. The calculations showed that a motion at that frequency made a

Table IV. BPTI—Aromatic Carbon Motional Frequencies (λ) and Amplitudes (α)[a]

peak/shift, ppm assigned residue(s)	field strength, MHz	T_1, s	T_2, s	NOE	σ^2	motion 1		motion 2		motion 3	
						λ_1, Hz	α_1, %	λ_2, Hz	α_2, %	λ_3, Hz	α_3, %
				A. Tyr^{3,5}							
W 115.99 Tyr^{3,5}	45	0.150	0.025	1.65	0.000	6E8	48	2E7	16	6E10	36
	90	0.305	0.025	1.35	0.000	6E8	44	2E7	15	7E10	41
X 116.87 Tyr^{3,5}	45[b]	0.120	0.025	1.15	0.416	6E8	2	2E8	90	3E11	8
	90[b]	0.320	0.035	1.15	2.363	6E8	12	1E8	81	5E10	7
Y 117.20 Tyr^{3,5}	45	0.095	0.030	1.15							
	90	0.295	0.040	1.15	0.273	6E8	0.3	9E7	42	2E8	57
Z 118.37 Tyr^{3,5}	45	0.115	0.045	1.55	0.000	6E8	55	8E7	22	9E10	23
	90	0.210	0.050	1.20	0.021	6E8	22	2E8	77	9E10	1
				B. Phe and Tyr^{2,6}							
A′ 128.10 Phe/Tyr^{2,6}	45[c]	0.120	0.030	1.20	1.422	6E8	5	1E8	88	7E10	17
	90[c]	0.290	0.040	1.15	2.707	6E8	16	1E8	80	4E10	4
B′ 128.51 Phr/Tyr^{2,6}	45[c]	0.125	0.035	1.20	1.026	6E8	12	1E8	66	9E11	22
	90[c]	0.320	0.040	1.15	2.481	6E8	12	1E8	84	4E10	4
C′ 128.90 Phe/Tyr^{2,6}	45	0.095	0.040	1.15							
	90	0.215	(0.035)	1.30	0.000	6E8	65	2E7	20	4E10	15
D′ 130.00 Phe/Tyr^{2,6}	45	0.090	0.025	1.15							
	90	0.255	0.025	1.15	0.246	6E8	2	2E8	85	1E7	13
E′[d] 130.41 Phe/Tyr^{2,6}	45	0.105	0.030	1.30	0.156	6E8	28	1E8	62	6E11	9
	90[c]	0.240	(0.030)	1.15	0.584	6E8	8	2E8	74	1E7	18
F′ 131.27 Phe/Tyr^{2,6}	45	0.095	0.040	1.15							
	90	0.180	(0.035)	1.20	0.185	6E8	9	3E8	83	1E7	8
G′[d] 131.56 Phe/Tyr^{2,6}	45[c]	0.135	0.030	1.25	0.602	6E8	11	1E8	58	9E11	29
	90[c]	0.215	0.030	1.15	0.688	6E8	11	2E8	79	1E7	10
H′ 133.25 Phe/Tyr^{2,6}	45	0.095	0.035	1.15							
	90	0.150	(0.030)	1.25	0.109	6E8	11	5E8	80	1E7	9

[a] Relaxation data at 17 °C, pD 5. Chemical shifts are referenced to external Me_4Si. The experimental T_1, T_2, and NOE determinations in any single run fluctuate by ±10%. Since the calculated error of fits as measured by the square of the variance (σ^2) is considerably smaller than the experimental error, the relaxation values are rounded off to the nearest five or ten in the third figure and may be said to represent both the experimental and calculated values, except where noted. E represents a power of 10. [b] Calculated T_1, T_2, and NOE for peak X (45 MHz, $\sigma^2 = 0.416$) 0.130 s, 0.025 s, and 1.20; (90 MHz, $\sigma^2 = 2.363$) 0.385 s, 0.025 s, and 1.20. [c] Calculated T_1, T_2 and NOE for peak A′ (45 MHz, $\sigma^2 = 1.422$): 0.125 s, 0.025 s, and 1.25; (90 MHz $\sigma^2 = 2.707$) 0.345 s, 0.035 s, and 1.20. For peak B′ (45 MHz, $\sigma^2 = 1.026$): 0.145 s, 0.030 s, and 1.25; (90 MHz, $\sigma^2 = 2.481$) 0.370 s, 0.025 s, and 1.20. For peak E′ (90 MHz, $\sigma^2 = 0.584$): 0.280 s, 0.020 s, and 1.20. For peak G′ (45 MHz, $\sigma^2 = 0.602$): 0.155 s, 0.035 s, and 1.30; (90 MHz, $\sigma^2 = 0.688$) 0.255 s, 0.025 s, and 1.20. [d] For these resonances good agreement is seen for λ_1 and λ_2 at 45 and 90 MHz. However, the third motion λ_3 differs. Alternate fits exist in which λ_3 is compatible at the two frequencies. Thus, for these resonances, the three-motional fit carried out may be insufficient. A four-motional analysis may be necessary to distinguish the motions of these carbon groups.

negligible (<1%) contribution to the experimentally observed combination of relaxation parameters. A motion in the frequency range 3-6 × 10^8 s^{-1} was found in the same calculations to make a contribution with $\bar{\alpha}_i$ comparable to those reported in Tables I-IV for the respective groups. The program is therefore capable of rejecting values of λ_i that differ substantially from those relevant to the relaxation process.

Results and Discussion

BPTI was chosen for study because of its small size (58 amino acid residues, 6500 molecular weight[20a]) and accurately determined X-ray structure in the crystalline state[20b,c] and the fact that the globular conformation of BPTI is outstandingly stable toward denaturation by chemicals and by heat. For example, Wagner et al.[20d] suggest from NMR evidence that the backbone conformation in BPTI crystals is maintained over the entire range from 4 to 87 °C. A number of proton and carbon resonance lines, mainly in the aliphatic methyl and aromatic regions, have been identified and the chemical-shift changes of the resonances discussed in terms of changes of the protein structure to pH, denaturation, and binding experiments.[21] BPTI is thus a well-studied, relatively rigid protein, suitable as a logical first member for NMR relaxation studies on protein dynamics. One previous paper[7a] has reported a few

Table V. Correlation of Experimental and Calculated Relaxation Parameters at 90 MHz with the Variance σ^2 [a]

peak		NT_1, s	NT_2, s	NOE	σ^2	motion 1		motion 2		motion 3	
						λ_1, Hz	α_1, %	λ_2, Hz	α_2, %	λ_3, Hz	α_3, %
$Met^{52}CH_3$	exptl	1.050	0.174	1.739							
	calcd	1.0919	0.1732	2.0212	1.243	6E8	7.46	1E7	1.05	5E10	91.48
		1.0730	0.1736	1.8836	0.332	6E8	8.68	1E7	1.03	6E10	90.28
		1.0337	0.1747	1.6502	0.132	6E8	10.41	1E7	4.02	9E10	85.57
		1.0732	0.1741	1.6675	0.085	6E8	10.75	1E7	1.00	9E10	88.25
		0.9801	0.1742	1.3998	2.092	6E8	13.70	1E7	0.96	3E11	85.35
		1.0573	0.1738	1.7818	0.030	6E8	9.52	2E7	2.04	7E10	88.43
		1.0548	0.1737	1.7370	0.012	6E8	9.24	4E7	4.11	7E10	86.65
		1.0519	0.1738	1.7500	0.002	6E8	8.69	6E7	6.22	7E10	85.09
		1.0504	0.1740	1.7411	0.000	6E8	8.31	7E7	7.30	7E10	84.39
		1.0450	0.1748	1.7120	0.012	6E8	6.77	1E8	10.63	7E10	82.60
Glu CH_2	exptl	0.322	0.072	1.830							
	calcd	0.2602	0.0682	2.098	5.968	6E8	26.48	1E7	2.24	7E9	71.28
		0.3047	0.0708	2.032	1.103	6E8	25.85	1E7	2.26	1E10	71.90
		0.3214	0.0728	1.644	0.552	6E8	35.03	2E7	4.23	2E10	60.74
		0.2768	0.0622	2.0705	3.728	6E8	24.42	4E7	10.61	7E9	64.97
		0.3207	0.0738	1.6180	0.714	6E8	34.96	4E7	8.30	2E10	56.74
		0.3220	0.0694	1.9074	0.155	6E8	24.09	7E7	16.78	1E10	59.14
		0.3137	0.0674	1.9370	0.349	6E8	21.91	8E7	20.14	9E9	57.95
		0.3186	0.0698	1.8902	0.093	6E8	23.25	8E7	19.19	1E10	57.57
		0.3206	0.0708	1.8578	0.020	6E8	21.11	1E8	23.96	1E10	54.93
		0.3110	0.0652	1.9467	0.517	6E8	18.02	1E8	26.90	8E9	55.08
		0.3240	0.0781	1.7487	0.187	6E8	1.55	2E8	48.87	1E10	49.58

[a] The experimentally determined T_1, T_2, and NOE in this table are given to three significant figures. The computer program searches for convergent fits for all the relaxation parameters and gives calculated values varying in the fourth and fifth digit. The square of the variance (σ^2) is conveniently used as an index of the goodness of fit and is defined as the cumulative computational error of the average deviation of the calculated from the measured parameters (eq 8′ of the Appendix). Significant differences in the calculated and measured values of *all three* parameters occur when $\sigma^2 \gtrsim 1.0$. When $1.0 > \sigma^2 \gtrsim 0.5$, deviations in *two* parameters are notable. When $0.5 \gtrsim \sigma^2 \gtrsim 0.1$, deviations in one parameter are notable but the remaining measured and calculated parameters round off to give identical numbers. When $\sigma^2 \lesssim 0.1$, the calculated and measured numbers are essentially identical after rounding off. Calculated fits with $\sigma^2 < 0.030$ are essentially indistinguishable from the experimentally measured quantities and are taken as "zero error fits."

relaxation times for this protein system and, on the assumption that the α-CH carbons represent exclusively the molecular tumbling, has suggested that the protein has an average tumbling correlation time of $\sim 2 \times 10^{-8}$ s (frequency $\lambda = 5 \times 10^7$ Hz). This value was subsequently found to be incorrect by measurements made both in this laboratory and by Wüthrich and collaborators.

Figure 1A shows the 90-MHz ^{13}C NMR spectrum of 9.8 mM BPTI protein at 17 °C. A large number of carbon lines are resolvable in the methyl, methylene, methine, aromatic, and carbonyl regions. The 20 methyl carbons in BPTI occur in 13 carbon resonances, as shown in Figure 2A. The assignments shown are those suggested by Richarz and Wüthrich.[21a] The methylene and α-CH region is illustrated in Figure 2B. From pH titration studies, it has been possible to identify the methylene carbons of the ionizable side chain groups of Lys, Asp, and Glu and the C-terminal α-CH (Ala 58 α-CH) resonance.[21] The aromatic region is shown in Figure 2C. BPTI possesses four Phe and four Tyr groups. The eight $Tyr^{3,5}$ carbons arise as the four upfield carbon resonances at ~115–117 ppm. The remaining eight carbon resonances labeled peaks a′–h′ arise from $Tyr^{2,6}$ and/or Phe carbons.

Experimental and calculated T_1, T_2, and NOE relaxation parameters for these resonances are presented in Tables I–IV. In general, the T_1 values of these protonated carbon resonances range from about 100 to 300 ms at 45 MHz and increase to 200–400 ms with increase of field strength to 90 MHz. The T_2 values are generally shorter than T_1 at both fields. The NOE values are obtained by comparison of the decoupled spectrum obtained with the decoupler continuously on (Figure 1A) and the decoupled spectrum obtained with the decoupler turned off during the delay time, so that no spin polarization occurs to give rise to the NOE effect (Figure 1B). Subtraction to give the difference spectrum in Figure 1C gives significant peak intensity only in the aliphatic methyl and methylene carbon regions. The nonprotonated carbonyls and the protonated aromatic and backbone α-methine carbons thus exhibit small NOEs, while the NOEs are appreciable in the methyl and methylene regions. Indeed, a close analysis shows that the NOEs ranged from the minimum of 1.15 for most aromatic and backbone α-CH resonances to values over 2.0 for the aliphatic methyls and methylenes. Overall, the relaxation data appear entirely consistent with interpretations in terms of pure dipolar relaxation mechanisms.[6a,22]

The need for a careful analysis which would account for at least three relaxation parameters simultaneously, as discussed in the section on theory, should be clearly evident. The use of a single relaxation parameter, such as a T_1 value, can lead to any arbitrary value of a tumbling correlation time, depending on the specific model of tumbling assumed. The range of motional frequency values obtained in this manner is clearly too large to even approach an understanding of molecular events. Similarly, a calculation could be carried out using any two relaxation parameters at either one or even two frequencies. The range of possibilities is narrowed, but the question of the uniqueness of interpretation remains.[11b,c] To be physically meaningful, a correlation rate $\bar{\lambda}_i$ and amplitude $\bar{\alpha}_i$ must account for *all* relaxation parameters, at all frequencies used for experimental measurement.

Results of the Analysis

General Features. The analysis of the relaxation data on the 40 ^{13}C resonances in the BPTI spectrum by the procedure outlined above leads to several general conclusions:

(1) For none of the resonances can the set of six (or even three) relaxation parameters (e.g., T_1 T_2, and NOE at 45 and 90 MHz) be *simultaneously* accounted for by a single motional term. This observation clearly indicates that there is more than one motion present in the protein system.

(2) Two motions, one of which is the rotational diffusion of

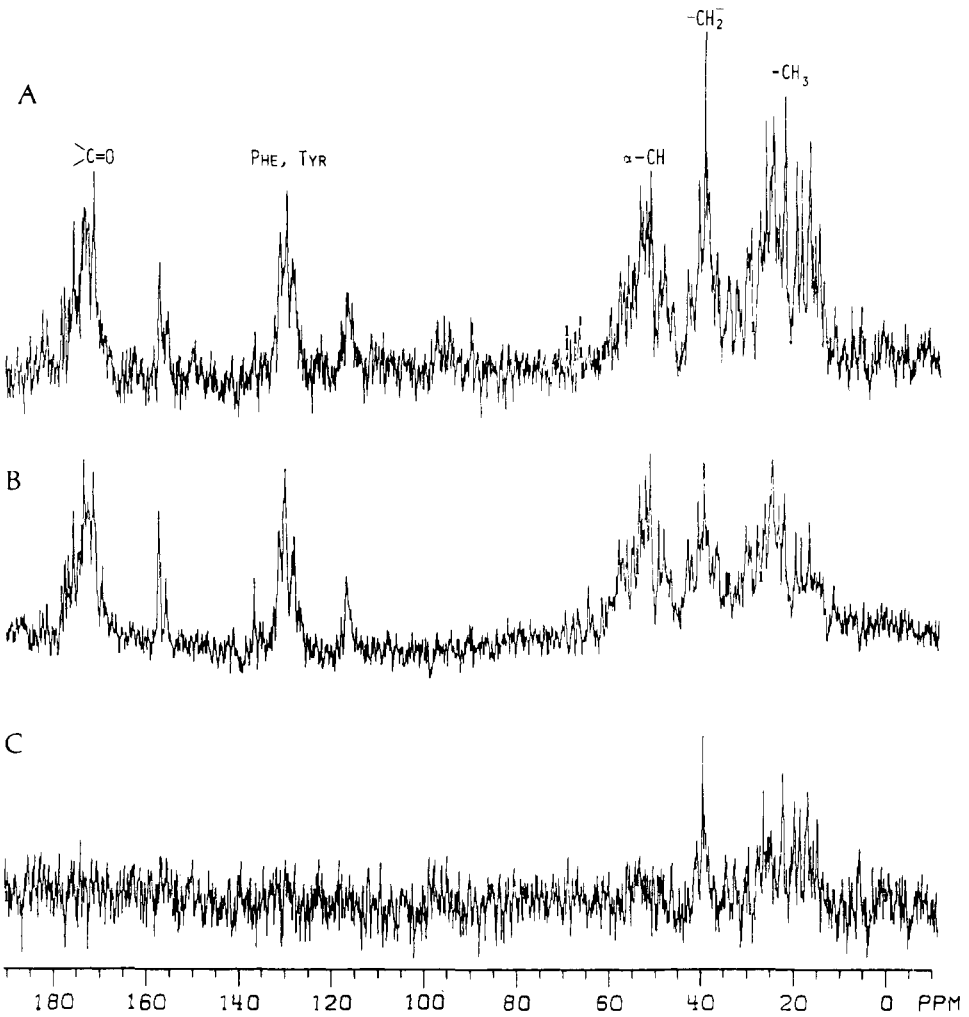

Figure 1. The 90-MHz ^{13}C NMR spectrum (ppm) of 9.8 mM bovine pancreatic trypsin inhibitor with 10^{-4} M EDTA, 50 mM NaCl buffer, and at pD 5 in D_2O. Spectral conditions: 11 200 scans, 16K channels, ±10 000-Hz spectral window, probe temperature 17 °C. (A) Fully decoupled spectrum with NOE. (B) Decoupled spectrum without NOE. (C) Difference spectrum of (A) and (B).

the protein, suffice to account for relaxation in only two of the resonances. Both the 53.5-ppm α-CH line (α-CH no. 4) and the 118.4-ppm $Tyr^{3,5}$ carbon line (Z) gave two motion fits with $\lambda_1 = 6 \times 10^8\ s^{-1}$, $\alpha_1 = 13\%$ and $\lambda_2 = 2 \times 10^8\ s^{-1}$, $\alpha_2 = 87\%$. The nature of the second motion at a frequency of $2 \times 10^8\ s^{-1}$ cannot be specified, but it could be the diffusion of the longer axis of the anisotropic molecule (axial ratio ~3:1) using an approximation of an ellipsoid of revolution. This result is compatible with the notion that these two groups are part of the rigid structure of the protein but does not prove it. In this context it is worth noting that for irregular anisotropic structures such as the approximately pear-shaped BPTI molecule there is no clear definition of the axes whose motion is detected by light scattering. In the analysis of light-scattering data using the Perrin equation for a prolate ellipsoid of revolution[19] it is assumed that the motion of the long axis makes the major contribution. This analysis is not strictly applicable to irregularly shaped objects and the value calculated from the Perrin equation must therefore be regarded as approximate. Without a constraint of the overall tumbling rate at $6 \times 10^8\ s^{-1}$ the analysis of NMR relaxation data yields, as noted above, values of $3\text{-}6 \times 10^8\ s^{-1}$ which can be attributed to the rotational diffusion of the molecule. If it were possible to be certain a priori that parts of the protein structure are rigid, the cited result would indicate that a discrimination between the motions of different axes should be possible on the basis of NMR measurements. This point merits further investigation. For the present it can be said that the principal motions identifiable with the diffusional rotation of BPTI occur in the frequency range $2\text{-}6 \times 10^8\ s^{-1}$ since motions with a frequency of $0.5\text{-}1 \times 10^8$ make a negligible contribution to relaxation. The modes of these motions cannot be specified in detail without recourse to unverifiable assumptions. Internal motions in this frequency range are operationally indistinguishable from the rotation of the molecule as a whole.

(3) A minimum of three motions is *required* to account for the relaxation data on all other resonances. Thus, except for the preceding two, none of the observed resonances can be said to represent groups rigidly held in the protein structure. With three motions and six measured parameters, an exact solution of the six simultaneous relaxation equations is possible. In our fitting procedure, the best fit was obtained separately for the three relaxation measurements at each frequency. A set of $\bar{\alpha}_k$, $\bar{\lambda}_k$, which provides a solution for one set, generally constituted a solution of the other (Tables I-IV).

Specific Features and Correlation with Structure. Figure 3 shows the folded peptide backbone of BPTI determined from X-ray data. As can be seen in this figure, the globular conformation of the polypeptide backbone is characterized by a twisted antiparallel β sheet which extends through the length of the molecule, encompassing residues 16-36. A short α helix is formed by residues 47-56.[20b,c] Much NMR evidence suggests that the overall globular structure observed in the BPTI crystals is closely maintained in solution. Recent molecular dynamics calculations by the Karplus group, however, suggest that the backbone is not static in solution but capable of small rearrangements in the 3-100-ps range.[5]

The analysis of motional frequencies and amplitudes cal-

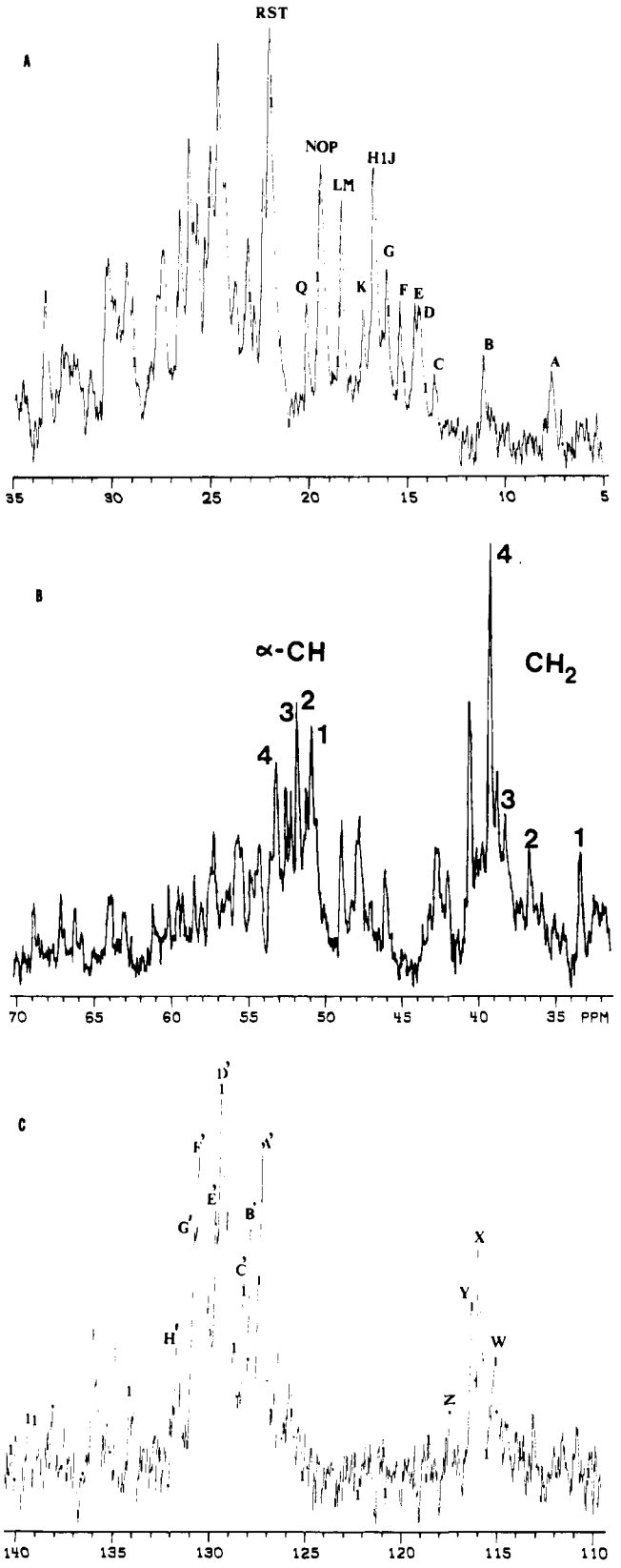

Figure 2. The 90-MHz ^{13}C NMR spectrum (ppm) of 9.8 mM BPTI in D_2O. Conditions as in Figure 1. Peak assignments follow. (A) Methyl region: A, Ile^{18} or $Ile^{18}\delta CH_3$; B, Ile^{18} or $Ile^{18}\delta CH_3$; C, $Met^{52}SCH_3$; D, $Ile\gamma CH_3$ or $Leu^{29}\delta^1CH_3$; E, $Ala^{28}\beta$; F, $Ile\gamma$; G, $Ala^{25}\beta$; HIJ, Ala^{16}, Ala^{27}, Ala^{58}; K, $Ala^{40}\beta$; LM, $Thr^{11}\gamma$ (Leu^6 or Leu^{29}); NOP, $Thr^{54}\gamma$, $Thr^{32}\gamma$; Q, Leu^6; RST, Val or Leu or Ile. (B) Methylene region: CH_2 1, Glu^7 or $Glu^{49}CH_2$; CH_2 2, Asp^3 or Asp^{50} CH_2; CH_2 3, Asp^3 or Asp^{50} CH_2; CH_2 4, Lys ϵ CH_2. Methyne region: α-CH 1, unknown; α-CH 2, Ala^{58} α-CH; α-CH 3, unknown; α-CH 4, unknown. (C) Aromatic region: W, X, Y, Z, $Tyr^{3,5}$ carbons; A′–H′, Phe or $Tyr^{2,6}$ carbons. Assignments are those of Richarz and Wüthrich[21a] and Brown et al.[21b]

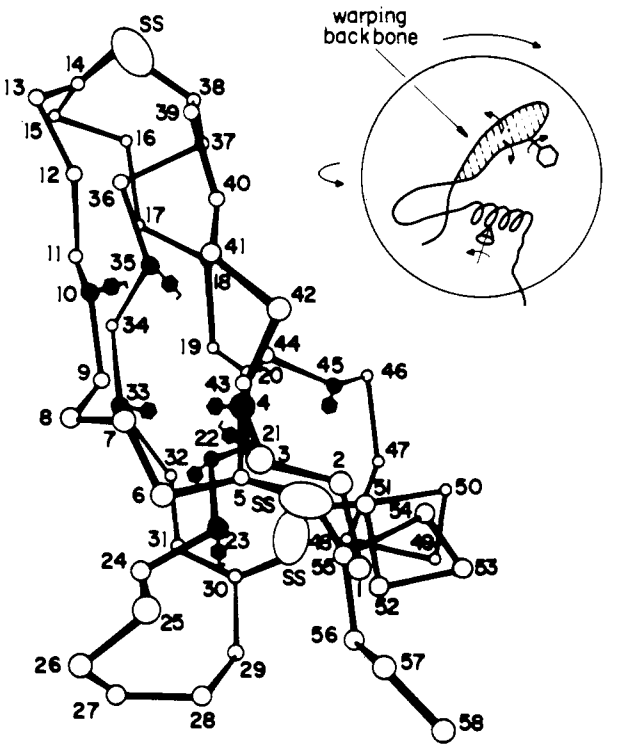

Figure 3. X-ray structure of the peptide backbone (α carbons) and disulfide bonds of BPTI.[20b] The Phe and Tyr aromatic rings are indicated by a hexagon. An extensive β sheet structure permeates the length of the structure at residues 16–36 with a short or helical portion at residues 47–56. The inset shows possible "warping" movements of the protein β sheet structure which might occur at the 10^7-s^{-1} frequency detected for virtually all carbon lines of the NMR analysis.

culated from the ^{13}C NMR relaxation data of this paper (Tables I–IV) reveals a variety of motional phenomena not previously apparent in this protein system:

(1) All groups are found to give at least one fit with the $\lambda_1 = 6 \times 10^8\ s^{-1}$ determined by light scattering for the rotational diffusion of the protein. This finding verifies the requirements imposed by the structure of the protein—i.e., the relaxation of all groups should be sensitive to one common tumbling time. The relative amplitude—i.e., the contribution—of this motion varies. Our data reveal that it is generally larger (30–50%) for the backbone and aromatic carbons and much smaller for the methyl groups (1–10%).

(2) The second and third motions are variable in frequency and amplitude for different groups. However, a low-frequency motion ($\lambda_2 = 1\text{–}2 \times 10^7\ s^{-1}$) makes a small but consistent contribution to the relaxation of virtually all resonances.

Comparison of spectral density functions for different specific motional models[11b,c] indicates that motional frequencies of this magnitude can still be reliably detected by relaxation measurements at 45 and 90 MHz, even though the measurements are most sensitive to motions in the neighborhood of the observing frequency. The existence of a slower motion can also be inferred from the observed T_1/T_2 ratios without detailed analysis. It can readily be shown algebraically that, when several motions contribute to the observed relaxation parameters, this ratio can never be larger than it would be for the slowest contributing motion. For the overall rotational diffusion of $3\text{–}6 \times 10^8\ s^{-1}$ the maximal T_1/T_2 ratio at 45 MHz is 2 in the outer limit and less at 90 MHz. Ratios of 3 or more are frequently seen in Tables I–IV. Therefore at least one motion slower than the rotational diffusion of the molecule makes a contribution to relaxation.

(3) A very high frequency component ($\lambda_3 = 10^{10}\text{–}10^{11}\ s^{-1}$) makes a significant contribution to the relaxation of all the

methyl and methylene resonances and several C_α and aromatic resonances.

(4) One other slow-frequency component (λ_2 or $\lambda_3 = 2 \times 10^8$ s^{-1}) contributes to the relaxation of the remaining C_α and aromatic resonances that do not yield high-frequency motions in the analysis.

(5) The observation of high- and low-frequency components for the relaxation of the C_α resonances clearly indicates that greater caution is necessary in the use of the common assumption that the relaxation of C_α resonances reflects only the motion of the rigid protein backbone.[6d,7a,18] The presence of a high-frequency component is easily understood in the case of the C_α of Ala 58, since this carbon group is near the C terminus of the protein (see Figure 3) and could wobble freely. It speaks against the dynamic stability of a salt bridge between Arg 1 and Ala 58, postulated on the basis of chemical-shift measurements.[21b] No such simple explanation can be advanced for the appearance of the high-frequency component among other C_α's, but it is possible that this component reflects the small rearrangements predicted by the molecular-dynamics calculations.[5]

(6) The nature of the low-frequency component $\lambda = 10^7$ s^{-1} is not defined by this type of analysis. If one assumes that it represents a low-amplitude wobble, the average angle of the wobble can be estimated to be <45°. Other models can be proposed and cannot be distinguished from a wobble on the basis of relaxation data alone. The ubiquity of this component suggests that it may represent a general relatively slow warp of the entire backbone, which would be reflected in the motion of both the backbone and all side-chain carbons (Figure 3). Its origin may well lie in the collisions between protein molecules. This can in principle, but for the present not in practice, be verified by varying the protein concentration.

(7) The rapid component of the motion of the methylene and methyl groups can reasonably be assumed to represent side-chain or methyl-group rotations. The fractional contribution of this motion in the methylene groups of the Glu and Asp side chains is clearly less than that in any methyl group. The fractional contribution of this motion for methyl groups (92–98%) in fact exceeds the theoretical maximum (88.8%) predicted by the appropriate rigid-rotor model of anisotropic motion.[9c] The discrepancy is at the limits of the uncertainty of the present analysis. However, the result and the observation that there exist alternative values of λ_3 in the same frequency range that can account for the relaxation of the methyl groups suggest that methyl-group rotation may not be the only significant motion at this frequency. Discrimination may be possible if the analysis is extended to include four or more degrees of motional freedom. To carry out a reliable analysis of this type will require a set of relaxation parameters at a third frequency.

The high rates of motion for the methyl groups are not unreasonable, since an examination of a molecular model of BPTI (based on Figure 3) reveals that all the aliphatic groups examined are on the surface of the structure.

(8) The second slow component in the range 2×10^8 s^{-1} that dominates the relaxation of a few aromatic and at least one α-CH resonance is interpreted to mean that the relaxation parameters here reflect the asymmetric diffusion of the protein, although contributions of other motions in this frequency range cannot be strictly ruled out (cf. above).

(9) The motions of $2\text{-}6 \times 10^8$ s^{-1} typically account for 50–90% of the relaxation of the aromatic resonances, while they are typically less than 10% for the aliphatic methyls. This is understandable, since the molecular model reveals that the aromatic resonances are generally not on the surface of the protein. The relative amplitudes of the low-frequency motions connected with rotational diffusion are expected to vary greatly, since each aromatic ring clearly has widely different angles of projection on the molecular tumbling axes.

A third motion at low frequency (10^7 s^{-1}) is seen for one Tyr^{3,5} resonance (peak y) and three of the eight Tyr^{2,6}/Phe carbon resonances (peaks d′, e′, and g′). In all other aromatic lines, the third motion appears at high frequency (10^{10} s^{-1}). With lack of exact assignments, it is difficult to correlate the motions of these lines with the microenvironments of the four Tyr—at residues 10, 21, 23, and 35—and four Phe—at 4, 22, 33, and 45 in the molecular model[20] (Figure 3). Nevertheless, it is instructive to compare the results of the relaxation analysis with the extensive line-shape analysis of aromatic rings by ^{1}H NMR.[20d] Based on the assignments of Snyder et al. and Wagner and Wüthrich,[23] Tyr 10 and Tyr 21 are suggested to undergo fast intramolecular ring flipping at a rate of 5×10^4 to 10^8 s^{-1} between 4 and 72 °C. The Tyr 23 ring does not flip until ~15 °C and by 40 °C is implicated in fast intramolecular rotations. On the other hand, the Tyr 35 ring is immobile between 4 and 35 °C and rotates appreciably only above 50 °C.

One Phe ring behaves very similarly to Tyr 10 and Tyr 21. A second Phe ring is immobile until temperatures above 70 °C. A third Phe ring rotates rapidly above 26 °C, and the last Phe ring rotates slowly between 15 and 30 °C. It is thus likely from our data that peak y arises from the 3,5 carbons of Tyr 35, while the corresponding 2,6 carbons may be part of peak d′, e′, or g′.

The high-frequency component of 10^{10} Hz observed for aromatic rings at 17 °C cannot be accounted for by the rates of ring flipping obtained from proton chemical shift data. Thus, the aromatic rings in BPTI exhibit an additional motion. One possibility is "quivering" or "libration" of the aromatic rings with a relatively low amplitude and relatively low probability of reaching the 180° flip required for effective chemical-shift averaging.[11c,14]

Summary and Conclusions

The analysis of a large set of NMR relaxation data in terms of the theory presented here permits a systematic testing and comparison of various models, rather than simple calculation of correlation times from models of molecular motion chosen ad hoc. It is apparent that, when only one or two relaxation parameters are available for a single line, a given model can be chosen and arbitrarily modified to fit the observed data.[11a,c] Thus, conventional analysis of relaxation data can easily lead to incomplete and erroneous results. An analysis of the form reported here—separation of the relaxation observed into relative amplitudes and frequencies of motion *prior* to the testing of specific motional models—appears more meaningful, especially for macromolecules where individual modes of motion may be complex and not readily identifiable. No specific model of molecular motion is assumed a priori. Instead, a range of allowed motional frequencies and their relative contributions are calculated for each group in the macromolecule on the basis of an extensive set of experimental data.

The results reported here for BPTI indicate that, in addition to diffusional tumbling and side-chain rotation, there exist both a very rapid and a relatively slow component in the motion of the protein backbone. Neither of these motions is detected when inadequate data and limited models are used. The detection of the very rapid component is the first experimental verification of molecular-dynamics calculations which predict small rearrangements of the polypeptide backbone in the 10^{10}–10^{12}-s^{-1} range.

The general distinction of slow and fast motions and their relative contributions should prove most useful in locating flexible domains in protein systems and differentiating between freely rotating aliphatic side chains on a protein surface and more restricted groups that may be sterically constrained.

Acknowledgments. This research was supported by NSF Grants PCM78-07930, PCM75-02814, and GP23633, as well

as NIH Grant RR00711. The authors thank Mr. Rudi Nunlist for help with 45-MHz ^{13}C relaxation measurements carried out on the Bruker 180-MHz wide-bore instrument at the University of California at Berkeley and J. Reynolds and R. Bucks for collection of data in the initial stages of the work. The authors also thank Dr. C. C. Wang and Professor R. Pecora for the depolarized light scattering measurements.

Appendix

A. Approximations to Simplify the Calculation of Spectral Density Functions for Multiple Motions. In previous communications,[11a,b] we derived a general expression for the spectral density function of a molecule with multiple levels of internal motional freedom, using the theory of Markov processes (MP). For a series of M independent Markov processes, the number of Lorentzian terms necessary to describe that spectral density function increases exponentially with M. That poses a dilemma: How can we empirically resolve multiple motions without gathering an enormous amount of data? The procedure we present here offers one such solution. Since, at least for microscopically reversible Markov processes, an analogy with quantum mechanics is formally valid, we turn to it for our method. The generator of a MP corresponds to a quantum-mechanical Hamiltonian, and its spectrum, to the quantum-mechanical energy spectrum. In addition, the generator of multiple Markovian motions is analogous to a series of noninteracting Hamiltonians, each with its own spectrum. If the spectra are sufficiently disjoint, we might condense all the energy levels into one average level for each Hamiltonian and sum the projection operators for each energy level of a particular Hamiltonian into a single projection operator for that Hamiltonian. This procedure might not significantly distort the expectation values for certain observations. (Such a technique is often employed in quantum statistics.)

The physical implication of this procedure is that each motion is characterized by a single average correlation rate $\bar{\lambda}_i$. The range $\Delta \log \lambda_i$ of rate constants associated with this motion is small compared to the corresponding difference in $\bar{\lambda}_i$, $\bar{\lambda}_j$; i.e., $\Delta \log \lambda_i \ll \log \bar{\lambda}_i - \log \bar{\lambda}_j$; i.e., the spectrum of motional frequencies is well resolved on a logarithmic frequency scale. Motions for which this condition is not satisfied are not operationally distinguishable in this approximation. To derive the simplified form of eq 3′ suitable for the analysis of a limited set of data, as given in eq 5′, we must define the meaning of $|\langle F, \Pi_{k=1}^{M} \phi_{kn_k}\rangle|^2$ in eq 3′ under the conditions of this approximation. This can be done by the following argument.

In the language of Markov processes, let $P_t(t,y)\,dy$ be the transition probabilities of moving from x to near y in time t. We assume that $x, y \in X$, the state space of the global MP. Suppose that X decomposes into the product of M subspaces—i.e., $X = \Pi_{k=1}^{M} X_k$—and that for $x = (x_1, \ldots, x_M)$, $y = (y_1, \ldots, y_M)$, $P_t(x,y)\,dy = P_{1t}(x_1,y_1)P_{2t}(x_2,y_2)\ldots P_{Mt}\cdot(x_M,y_M)\,dy_1 \ldots dy_M$. That is, the submotions are stochastically independent. Let $p_k(dy_k)$ be the equilibrium density of the kth process and define

$$\langle f,g\rangle = \int_{X_k} f(x)\bar{g}(x)p_k(dx)$$

for complex function f,g on X_k. This defines Hilbert spaces $L(X_k,p_k)$ and a Hilbert space $L^2(X,p) = \otimes_k L^2(X_k,p_k)$ where $p(dy) = p_1(dy_1)\ldots p_M(dy_M)$. P_{kt} defines an operator on $L^2(X_k,p_k)$ by $P_{kt}(f)(x) = \int_{X_k} f(y)P_{kt}(x,y)\,dy$; similarly does P_t.

The generator Ω of the process P_t can be shown to be equal to $\Sigma_k I\otimes \ldots \Omega_k \otimes \ldots I$ where the Ω_k's are the respective subgenerators.

Let ϕ_{kn} be defined so that $\Omega_k\phi_{kn} = \lambda_{kn}\phi_{kn}$—that is, ϕ_{kn} is the nth normalized eigenfunction of the kth process. Let

$$R_{kn} = I\otimes I\otimes \ldots \otimes R_{kn}^{k\text{th place}} \otimes I \ldots \otimes I \qquad (1')$$

where

$$R_{kn}f = \langle f,\phi_{kn}\rangle \phi_{kn}$$

for $f \in L^2(X_k,p_k)$. One can show that

$$\int_{-\infty}^{\infty} E\{F(t)F^*(t+\tau)\}e^{i\omega\tau}\,d\tau \equiv$$

$$J_F(\omega) = -2 \sum_{n_1,n_2\cdots n_M} \frac{\left|\left\langle F, \prod_{k=1}^{M} \phi_{kn_k}\right\rangle\right|^2 \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)}{\omega^2 + \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)^2} \qquad (2')$$

Equations 1′ and 2′ translate to

$$J_F(\omega) = -2 \sum_{n_1,n_2\cdots n_M} \frac{\left|\left\langle F, \bigotimes_{k=1}^{M} R_{kn_k}\right\rangle\right|^2 \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)}{\omega^2 + \left(\sum_{k=1}^{M} \lambda_{kn_k}\right)^2} \qquad (3')$$

As was stated above, the number of terms in eq 2′ and 3′ grows exponentially with M, making it exceedingly difficult to resolve many individual motions in the absence of a correspondingly larger number of measurements. The approximation is therefore introduced that each motion can be characterized by a single average $\bar{\lambda}_k$. As a consequence, motions with nearly equal $\bar{\lambda}_k$ become operationally indistinguishable.

By assumption, $\lambda_{1n_1} \ll \lambda_{2n_2} \ll \ldots \lambda_{Mn_M}$ for all $n_1, \ldots, n_M$—i.e., the motions are well separated in relaxation rates. In addition, let us approximate λ_{kn_k} by $\bar{\lambda}_k$ for all $n_k \neq 0$ and define $\bar{\lambda}_0 = 0$. Also, $\lambda_{k0} = 0$, and $\phi_{k0} = l_k$ since $\Omega_k l_k = 0$ where $l_k = 1$ for all $x \in X_k$. Then we may approximate (3′) by

$$J_F(\omega) \approx -2 \sum_{k=0}^{M} \frac{\left\langle F, \left\{\bigotimes_{i=1}^{k-1} (I) \otimes (I - P_{k0}) \bigotimes_{i=k+1}^{M} P_{i0}\right\} F\right\rangle \bar{\lambda}_k}{\omega^2 + \bar{\lambda}_k^2}$$

$$= -2 \sum_{k=0}^{M} \frac{\langle F,0_kF\rangle \bar{\lambda}_k}{\omega^2 + \bar{\lambda}_k^2} \qquad (4')$$

where $\Sigma_K 0_k = I$. This expression for $J_F(\omega)$ consists of $M+1$ terms, one for each submotion. Let

$$Q_K = I\otimes I\otimes \ldots \otimes R_{(k+1)0} \otimes R_{(k+2)0} \ldots \otimes R_{MO}$$

$$0_0 = Q_0 \text{ and } 0_K = Q_K - Q_{k-1} \qquad (5')$$

Let us turn our attention to the Q_k's. One can see that

$$Q_kF = (I\otimes I\otimes \ldots R_{k+1)0} \otimes \ldots \otimes R_{MO})F$$

$$= E^{\{X_1,\ldots X_k\}}F \equiv E^kF \qquad (6')$$

where $E^{\{X_1,\ldots X_k\}}F$ is the conditional expectation of F given $\{X_1, \ldots X_k\}$. Then $\langle F,Q_kF\rangle = E(FE^kF) = \mathrm{Cov}(F,E^kF) = \{\mathrm{Cor}(F,E^kF)\} \equiv \rho_k^2$, assuming $E(F^2) = 1$, since $E(E^kF\cdot E^kF) = \langle Q_kF,Q_kF\rangle = \langle Q_kF,F\rangle = E(FE^kF)$. [$\mathrm{Cor}(A,B)$ = correlation coefficient between A,B = $\mathrm{Cov}(A,B)/[\mathrm{Cor}(A,A)\cdot \mathrm{Cov}(B,B)]^{1/2}$.]

Intuitively, since E^kF is the best least-squares predictor of F, given knowledge of the composite Markov process on the first k variables $X_1, \ldots, X_k$. $\langle F,Q_kF\rangle$ is just the square of the correlation between the actual random function F and its best partial predictor. By eq 5′ then $\langle F,0_kF\rangle = \rho_k^2 - \rho_{k-1}^2$. $\langle F,0_0F\rangle = \rho_0^2$. Let $\alpha_k = \rho_k^2 - \rho_{k-1}^2$ and $\alpha_0 = \rho_0^2$ be the increment in correlation, squared by adding knowledge of one additional motion. We see that eq 4′ now becomes

$$J_F(\omega) \approx -2 \sum_{k=0}^{M} \frac{\alpha_k\bar{\lambda}_k}{\omega^2 + \bar{\lambda}_k^2} \qquad (7')$$

The α_k's are therefore the amplitudes for each independent motion with average correlation rates λ_k.

B. Least-Squares Algorithm for the Calculation of Relaxation Parameters. Equation 7′ gave us a simplified expansion of $J_F(\omega)$ in terms of a series of Lorentzians. For rotational motion, $1/T_1$, $1/T_2$, and $(\mathrm{NOE}-1)/T_1$ can be written as a linear summation of $J[Y_{2n}(\omega)]$, where Y_{2n} is a second-order spherical harmonic. We can then predict for each series of λ_k's the optimal (in the least-squares sense) set of α_k's, subject to the linear condition that $\Sigma_{k=1}^{M}\alpha_k = 1$. This formulation makes no assumption as to the physical nature of the relaxation mechanism and is equally applicable to the dipolar, chemical shift anisotropy, and other relaxation mechanisms, *provided* that the length of the vector **R** connecting the origin of the relaxing field to the relaxed nucleus may be assumed to be constant. Alternative simplifications are necessary if variations in the length of this vector must be taken into account.

Given a complete set of T_1's, T_2's, and NOE's at N frequencies and the additional assumption of carbon-hydrogen dipolar relaxation, the following algorithm can be used to carry out an iterative fit of the calculated to the observed set of relaxation parameters. Only a slight modification will be necessary to include other relaxation mechanisms.

1. Let $\omega_1{}^{\mathrm{C}} \ldots \omega_N{}^{\mathrm{C}} \ldots$ be the N carbon frequencies; $\omega_1{}^{\mathrm{H}} \ldots \omega_N{}^{\mathrm{H}}$ be the N hydrogen frequencies; $\omega_i{}^{\mathrm{H}}/\omega_i{}^{\mathrm{C}} = X$. Let $\mathrm{HMC}_i = \omega_i{}^{\mathrm{H}} - \omega_i{}^{\mathrm{C}}$; $\mathrm{HPC}_i = \omega_i{}^{\mathrm{H}} + \omega_i{}^{\mathrm{C}}$.

2. Let T_{1i}, T_{2i}, NOE_i be the measured relaxation data for a particular peak at the ith carbon frequency.

3. Let $Z_{1i} = 1/T_{1i}$, $Z_{2i} = 1/T_{2i}$, $Z_{3i} = (\mathrm{NOE}_i - 1)/T_{1i}$.

4. For fixed λ_1, λ_2, ω define

$$Y(\lambda_1, \lambda_2, \omega) = \frac{2\lambda_1}{\lambda_1{}^2 + \omega^2} - \frac{2\lambda_2}{\lambda_2{}^2 + \omega^2}$$

$$X(\lambda_2, \omega) = \frac{2\lambda_1}{\lambda_2{}^2 + \omega^2}$$

5. Fix sequence of λ_m's, $M = 1, \ldots, M$, so that $\lambda_1 < \lambda_j \ldots < \lambda_M$.

6. To take into account that $\Sigma_{k=1}^{M}\alpha_k = 1$, define the unconstrained contribution to the spectral density function F_{kij} such that if $k = 1$, then $F_{kij} = R_1\{2Y(\lambda_j,\lambda_M,\mathrm{HMC}_i) + 6Y(\lambda_j,\lambda_M,\omega_i{}^{\mathrm{C}}) + 12Y(\lambda_j,\lambda_M,\mathrm{HPC}_i)\}$; if $k = 2$, then $F_{kij} = R_2\{4Y(\lambda_j,\lambda_M,0) + 3Y(\lambda_j,\lambda_M,\omega_i{}^{\mathrm{C}}) + Y(\lambda_j,\lambda_M,\mathrm{HMC}_i) + 6Y(\lambda_j,\lambda_M,\omega_i{}^{\mathrm{H}}) + 6Y(\lambda_j,\lambda_M,\mathrm{HPC}_i)\}$; if $k = 3$, then $F_{kij} = R_3\{6Y(\lambda_j,\lambda_M,\mathrm{HPC}_i) - Y(\lambda_j,\lambda_M,\mathrm{HMC}_i)\}$; where R_1, R_2, and R_3 are the constants relating the spectral density functions to the appropriate relaxation parameter, given the specific relaxation mechanism.

For dipolar relaxation $R_1 = 1.075 \times 10^9$, $R_2 = 5.375 \times 10^8$, and $R_3 = 3.7004 \times 10^{-9}$.

Similarly define the constrained contribution G_{ki} such that

$$G_{1i} = R_1\{2X(\lambda_M,\mathrm{HMC}_i) + 6X(\lambda_M,\omega_i{}^{\mathrm{C}}) + 12X(\lambda_M,\mathrm{HPC}_i)\}$$

$$G_{2i} = R_2\{4X(\lambda,0) + 3X(\lambda_M,\omega_i{}^{\mathrm{C}}) + X(\lambda_M,\mathrm{HMC}_i) + 6X(\lambda_M,\omega_i{}^{\mathrm{H}}) + 6\mathrm{X}(\lambda_{\mathrm{M}},\mathrm{HPC}_i)\}$$

$$G_{3i} = R_3\{6X(\lambda_M,\mathrm{HPC}_i) - X(\lambda_M,\mathrm{HMC}_i)\}$$

7. Let r_k = fractional error in Z_{ki} with $k = 1, 2, 3$.

8. Let the cumulative computational error measuring the average deviation of the calculated from the measured parameters be

$$E(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{M-1}) = \sum_{\substack{k=1,2,3 \\ i=1,\ldots,N}} \frac{\left(Z_{ki} - \sum_{j=1}^{M-1} F_{kij}\alpha_j - G_{ki}\right)^2}{r_k{}^2 Z_{ki}{}^2} \quad (8')$$

9. Differentiating with respect to α_l

$$\frac{E}{\alpha_l} = -2\sum_{k,i}\left\{\frac{\left(Z_{ki} - G_{ki} - \sum_{j=1}^{M-1} F_{kij}\alpha_j\right)F_{kil}}{r_k{}^2 Z_{ki}{}^2}\right\} \quad (9')$$

or

$$\sum_{j=1}^{M-1}\left\{\sum_{k,i}\frac{F_{kij}F_{kil}}{r_k{}^2 Z_{ki}{}^2}\right\}\alpha_j = \sum_{k,i}\frac{F_{kil}(Z_{ki} - G_{ki})}{r_k{}^2 Z_{ki}{}^2} \quad (10')$$

Let

$$A_{lj} = \sum_{k,i}\frac{F_{kij}F_{kil}}{r_k{}^2 Z_{ki}{}^2}$$

and

$$D_l = \sum\frac{F_{kil}(Z_{ki} - G_{ki})}{r_k{}^2 Z_{ki}{}^2}$$

Equation 10′ is now, expressed in matrix form

10. $$\sum_{j=1}^{M-1} A_{lj}\alpha_j = D_l \quad \text{for } l = 1, \ldots, M-1 \quad (11')$$

Let $\tilde{A} = (A_{lj})$, $\tilde{D} = (D_l)$, and $\tilde{\alpha} = (\alpha_j)$. Then eq 11′ reduces to the matrix equation

11. $$\tilde{A}\tilde{\alpha} = \tilde{D} \quad (12')$$

12. Solve (12′) for $\tilde{\alpha}$ resulting with $\tilde{\alpha} = \tilde{A}^{-1}\tilde{D}$. Let $\alpha_M = 1 - \Sigma_{j=1}^{M-1}\alpha_j$.

13. If it is not the case that $\alpha_j \geq 0\ J = 1, \ldots, M$, then return to step 5 (also increment the $\lambda_1, \ldots, \lambda$'s); otherwise

14. Let $L_{ki} = \Sigma_{j=1}^{M-1} F_{kij}\alpha_j - G_{ki}$.

15. Let $T_{1i}{}^{\mathrm{est}} = 1/L_{1i}$, $T_{2i}{}^{\mathrm{est}} = 1/L_{2i}$, and $\mathrm{NOE}_i{}^{\mathrm{est}} = (L_{3i}/L_{1i}) + 1$.

16. Let $\Delta_{1i} = (T_{1i} - T_{1i}{}^{\mathrm{est}})/L_{1i}$, $\Delta_{2i} = (T_{2i} - T_{2i}{}^{\mathrm{est}})/L_{2i}$, and $\Delta_{3i} = (\mathrm{NOE}_i - \mathrm{NOE}_i{}^{\mathrm{est}})/\mathrm{NOE}_i$, and print

17. $E(\alpha_1, \ldots, \alpha_{M-1})$, $\alpha_1, \ldots, \alpha_M$, $T_{1i}{}^{\mathrm{est}}$, $T_{2i}{}^{\mathrm{est}}$, $\mathrm{NOE}_i{}^{\mathrm{est}}$, Δ_{1i}, Δ_{2i}, Δ_{3i}, for $i = 1, \ldots, N$. Note that $E(\alpha_1 \ldots \alpha_{n-1}) = \Delta_{1i}{}^2/r_1{}^2 + \Delta_{2i}{}^2/r_2{}^2 + \Delta_{3i}{}^2/r_3{}^2$.

18. Go to step 5 and repeat the calculation incrementing the $\lambda_1 \ldots \lambda_M$'s.

References and Notes

(1) (a) Linderstrøm-Lang, K. U.; Schellman, J. A. In "The Enzymes", Boyer, P. D., Lardy, H., Myrback, K., Eds.; Academic Press: New York, 1959; Vol. I, pp 443–510. (b) Nakanishi, M.; Tsubol, M.; Ikegamal, A. *J. Mol. Biol.* **1972**, *70*, 351–361.

(2) Yguerabide, J.; Epstein, H. F.; Stryer, L. *J. Mol. Biol.* **1970**, *51*, 573–590.

(3) (a) Lackowicz, J. R.; Weber, G. *Biochemistry* **1973**, *12*, 4161–4170. (b) Eftink, M. R.; Ghiron, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3290–3294.

(4) (a) Saunders, M.; Wishnia, A. *Ann. N.Y. Acad. Sci.* **1957**, *70*, 870–874. (b) Jardetzky, O. *Adv. Chem. Phys.* **1964**, 499–531.

(5) (a) Gelin, B. R.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2002–2006. (b) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature (London)* **1977**, *267*, 585–590. (c) McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* **1979**, *18*, 927–942. (d) Karplus, M.; McCammon, J. A. *Nature (London)* **1979**, *277*, 578.

(6) (a) Doddrell, D.; Glushko, V.; Allerhand, A. *J. Chem. Phys.* **1972**, *56*, 3683–3689. (b) Opella, S. J.; Nelson, D. J.; Jardetzky, O. *J. Am. Chem. Soc.* **1974**, *96*, 7157–7159. (c) Campbell, I. D.; Dobson, C. M.; Williams, R. J. P. *Proc. R. Soc. London, Ser. B* **1975**, *189*, 503—509. (d) Visscher, R. B.; Gurd, F. R. N. *J. Biol. Chem.* **1975**, *250*, 2238–2242. (e) Hull, E.; Sykes, B. D. *J. Mol. Biol.* **1975**, *98*, 121–153.

(7) (a) Wüthrich, K.; Baumann, R. *Org. Magn. Reson.* **1976**, *8*, 532–535. (b) Opella, S. J.; Nelson, D. J.; Jardetzky, O. "Magnetic Resonance in Colloid and Interface Science", ACS Symposium Series; American Chemical Society: Washington, D.C., 1976; pp 397–417. (c) Deslauriers, R.; Ralston, E.; Somorjai, R. L. *J. Mol. Biol.* **1977**, *113*, 697–710. (d) Torchia, D. A.; Vanderhart, D. L. *J. Mol. Biol.* **1976**, *104*, 315–321.

(8) (a) Bull, T.; Nome, J. E.; Reimarsson, P.; Lindman, B. *J. Am. Chem. Soc.* **1978**, *100*, 4643–4647. (b) Wittebort, R. J.; Szabo, A. *J. Chem. Phys.* **1978**, *69*, 1722–1736. (c) Gilman, J. G. *Biochemistry* **1979**, *18*, 2273–2279.

(9) (a) Abragam, A. "The Principles of Nuclear Magnetism"; Clarendon Press: Oxford, 1961. (b) Bloembergen, N.; Purcell, E. M.; Pound, R. V. *Phys. Rev.* **1948**, *73*, 670–712. (c) Woessner, D. E. *J. Chem. Phys.* **1962**, *36*, 1–4. (d) *Ibid.* **1962**, *37*, 647–654. (e) Woessner, D. E.; Snowden, B. S.; Jr.; Meyer,

G. H. *Ibid.* **1969**, *50*, 719–721.
(10) For a review of relaxation theory, see: Noack, F. In "NMR: Basic Principles and Progress", Diehl, P., Ed.; Springer-Verlag: West Berlin, 1971; Vol. 3, pp 83–144.
(11) (a) King, R.; Jardetzky, O. *Chem. Phys. Lett.* **1978**, *55*, 15–18. (b) King, R.; Maas, R.; Gassner, M.; Nanda, R. K.; Conover, W. W.; Jardetzky, O. *Biophys. J.* **1978**, *6*, 103–117. (c) Jardetzky, O. In "NMR and Biochemistry: A Symposium in Honor of Mildred Cohn", Opella, S., Lu, P., Eds.; Marcel Dekker: New York, 1979; pp 141–167. (d) Jardetzky, O.; Ribeiro, A.; King, R. *Biochem. Biophys. Res. Commun.*, **1980**, *92*, 883–888. (e) Ribeiro, A.; Wade-Jardetzky, N. G.; King, R.; Jardetzky, O. *Appl. Spectrosc.*, in press.
(12) (a) Levine, Y. K.; Partington, P.; Roberts, G. C. K. *Mol. Phys.* **1973**, *25*, 497–514. (b) Levine, Y. K.; Birdsall, N. J. M.; Lee, A. G.; Metcalfe, J. C.; Partington, P.; Roberts, G. C. K. *J. Chem. Phys.* **1974**, *60*, 2890–2899.
(13) (a) Cutnell, J. D.; Bleich, H. E.; Glasel, J. A. *J. Magn. Reson.* **1976**, *21*, 43–46. (b) Vold, R. L.; Waugh, J. S.; Klein, M. P.; Phelps, D. E. *J. Chem. Phys.* **1968**, *48*, 3831–3832.
(14) Nelson, D. J.; Opella, S. J.; Jardetzky, O. *Biochemistry* **1976**, *15*, 5552–5560.
(15) Kowaleski, J.; Levy, G. C.; Johnson, L. F.; Palmer, L. *J. Magn. Reson.* **1977**, *26*, 533–536.
(16) Led, J. J.; Petersen, S. B. *J. Magn. Reson.* **1978**, *32*, 1–17.
(17) Levy, G. C.; Peat, I. R.; Rosanke, R.; Parks, S. *J. Magn. Reson.* **1975**, *18*, 205–208.
(18) R = 1.09 Å is commonly used as the internuclear distance between a carbon and its directly bonded proton. E.g., see: Allerhand, A.; Doddrell, D.; Glushko, V.; Cochran, D. W.; Werhert, E.; Lawson, P. J.; Gurd, F. R. N. *J. Am. Chem. Soc.* **1971**, *93*, 544–546. A recent note by K. Dill and A. Allerhand (*J. Am. Chem. Soc.* **1979**, *101*, 4376) points out that a 2–3% error in this value can cause an error of as much as a factor of 2 in the value of the correlation time calculated from a rigid rotor model. It should be noted that the *experimental* errors in determining T_1, T_2, and NOE on existing spectrometers are sufficiently large to cause uncertainties in the calculated correlation times of a similar order of magnitude. It is therefore doubtful that more precise values of τ_c or R_{CH} can at present be derived from relaxation measurements, as proposed by Dill and Allerhand.
(19) Berne, B.; Pecora, R. In "Dynamic Light Scattering"; Wiley: New York, 1976.
(20) (a) Tschesche, H. *Angew. Chem., Int. Ed. Engl.* **1974**, *13*, 10–28. (b) Deisenhofer, J.; Steigmann, W. *Acta Crystallogr., Sect. B* **1975**, *31*, 238–250. (c) Huber, R.; Kukla, D.; Ruhlmann, A.; Steigemann, W. *Cold Spring Harbor Symp. Quant. Biol.* **1971**, *36*, 141–150. (d) Wagner, G.; De Marco, A.; Wüthrich, K. *Biophys. Struct. Mech.* **1976**, *2*, 139–158.
(21) (a) Richarz, R.; Wüthrich, K. *Biochemistry* **1978**, *17*, 2263–2269. (b) Brown, L. R.; De Marco, A.; Richarz, R.; Wagner, G.; Wüthrich, K. *Eur. J. Biochem.* **1978**, *88*, 87–95.
(22) Kuhlmann, K. F.; Grant, D. M.; Harris, D. M. *J. Chem. Phys.* **1970**, *52*, 3429–3448. At higher magnetic field strengths, chemical-shift anisotropy (CSA) becomes the dominant relaxation mechanism for *nonprotonated unsaturated* carbons. For the hydrogen-bearing unsaturated carbons of aromatic side chains, the CSA mechanism could contribute slightly (possibly as high as 10%) to the relaxation. The measured relaxation parameters have uncertainties at least of this magnitude. The contribution of the CSA mechanism to the relaxation of protonated, saturated carbons (such as α carbons and CH_2 and CH_3 groups of a protein) is totally negligible for all magnetic field strengths available today for high-resolution NMR.
(23) (a) Snyder, G. H.; Rowen III, R.; Karplus, S.; Sykes, B. D. *Biochemistry* **1975**, *14*, 3765–3777. (b) Wagner, G.; Wüthrich, K. *J. Magn. Reson.* **1975**, *30*, 435–445.

Ab Initio, Quantum Chemical Analysis of Noncovalent Interactions between Peptides as Modeled by Dimers and a Trimer of Formamide

Ernest L. Mehler

Contribution from the Department of Structural Biology, Biocenter, The University of Basle, CH-4056 Basle, Switzerland. *Received July 23, 1979*

Abstract: The self-consistent, nonorthogonal group function approximation has been applied to dimers and a trimer of formamide in various geometries constructed to simulate intrachain, noncovalent interactions between peptides. The interactions simulated are hydrogen-bonded and nonbonded pairs in the α helix and the 3_{10} helix and the doubly hydrogen bonded trimer in the α helix. The interactions are decomposed into Coulomb-exchange, polarization, and charge-transfer contributions, a detailed analysis of the dimers and trimer is given, and it is shown that the main source of deviations from pair additivity is the polarization effect. On the basis of this analysis estimates of the interaction energy and dipole moment are obtained for hydrogen-bonded complexes of any length. Finally, it is estimated that the positive cooperativity effect of multiple hydrogen bonds in the infinite chain increases from 10% in the purely pair-additive interaction to 23% when the deviations from pair additivity are included.

I. Introduction

The fundamental role of noncovalent interactions in stabilizing polypeptide structure has long been recognized. Nevertheless, there have been relatively few investigations, beyond peptide pairs at the ab initio quantum chemical level, aimed at analyzing the nature of these interactions and assessing the importance of their contributions to the interaction energy. Such studies are of considerable importance since they can form the basis for formulating new approaches applicable to larger polypeptide chains, or indicate where the limited success[1] of current empirical approaches can be improved.

Ab initio or near-ab initio studies of peptide interactions in single strands have been reported by Shipman and Christoffersen[2] using the fragment molecular orbital method,[3] and Kleier and Lipscomb[4] using the partial retention of diatomic differential overlap (PRDDO) approximation.[5] The results are conflicting: Shipman and Christoffersen find the α-helical structure less stable than the fully extended (FE) structure, which is also found by Kleier and Lipscomb. The latter also find that the α helix is less stable than the 3_{10} helix, although it is by far the most commonly observed helical structure in globular proteins. Finally, Scheiner and Kern[6] have used the PRDDO method to compute peptide pair interactions and have subsequently calibrated an empirical potential function based on these computations. They find the α helix to be slightly more stable than the FE structure and the 3_{10} helix.

These results indicate that a more fundamental analysis is required to clarify the various interactions. The hydrogen-bonded (H-bonded) formamide pair interaction has been extensively studied. Most of these studies have been done at arbitrary or optimal geometries,[7] but a few papers have restricted the dimeric geometry to simulate noncovalent interactions in various types of protein secondary structure.[8] Beyond the dimer very few ab initio studies have been reported, although these are useful for studying cooperative effects and deviations from pair additivity in multiply H-bonded chains. Cooperativity

0002-7863/80/1502-4051$01.00/0