ACCOUNTS OF CHEMICAL RESEARCH

VOLUME 14 NUMBER 10 OCTOBER, 1981

NMR Studies of Macromolecular Dynamics

OLEG JARDETZKY

Stanford Magnetic Resonance Laboratory, Stanford University, Stanford, California 94305

Received March 17, 1981

The importance of internal mobility to the function of biological macromolecules is well recognized. It is apparent from X-ray structures of macromolecules that transient local packing defects must occur to admit small ligands—as oxygen to hemoglobin[1] or intercalating dyes into DNA.[2] "Induced fit"—the selection of specific local conformations on both protein side chains an ligands—is often necessary to produce optimum binding.[3-5] Rearrangement of subunits in response to ligand binding is known to be involved in the control functions of allosteric enzymes and a variety of other regulatory proteins.[6] Large-scale movements are essential for the mechanical functions of antibodies, muscle proteins, and the proteins of the mitotic apparatus.[7,8] There is also mounting evidence for the existence of extended flexible regions in some proteins, which become more rigid as a result of specific interactions with other macromolecules. In some cases, such as those of tobacco mosaic virus (TMV), a flexible peptide segment is necessary to permit the packaging of viral RNA inside the protein subunit.[9] In others, such as the muscle protein myosin,[10] the role of flexible regions is not yet understood.

Internal motions in macromolelcules can be detected by a variety of spectroscopic, diffraction, and other techniques, notably hydrogen exchange, fluorescence depolarization and quenching, and NMR; their existence has been known for nearly 30 years.[11-14] However, the information obtained by most methods used in the past has been too fragmentary to have permitted the construction of a coherent picture. The most detailed information on the amplitudes of very fast atomic motions in proteins comes from X-ray diffraction,[15] but the method provides no information on their time dependence or on slower motions.

Two recent developments offer hope for a more complete understanding of macromolecular dynamics: (1) With the advent of high-frequency high-resolution NMR spectrometers (360–600 MHz for ^1H), relatively accurate measurements on macromolecules have become possible. This, in turn, has led to advances in the theoretical interpretation of NMR data in terms of molecular events. (2) Computer simulations of molecular dynamics in proteins of known crystal structure[16] have provided a conceptual basis for predicting the types of motions that might be encountered.

An understanding of macromolecular dynamics will be achieved when it becomes possible to describe on an atomic scale the changes of macromolecular structures as functions of time in terms of solutions to an appropriate set of equations of motion and to correctly predict on this basis all experimental parameters in which motion is reflected. Much remains to be done to reach this goal. At present, theoretical calculations are limited to very fast motions ($<10^{-10}$ s), while many experimentally detectable motions are on longer time scales (10^{-9}–10^{-3} s). Nevertheless, the foundations of an understanding have been laid.

In this account we discuss three types of NMR observations which can provide information on internal motions in macromolecules: exchange of labile NH or OH hydrogens with the solvent, averaging of chemical

Oleg Jardetzky is Professor of Pharmacology at Stanford University and Director of the Stanford Magnetic Resonance Laboratory. After completing undergraduate work in chemistry and biology at Macalester College in St. Paul, MN (B.A. 1950, hon. D.Sc. 1974), he studied at the University of Minnesota, where he received both the M.D. and the Ph.D. degrees, the latter with W. N. Lipscomb and M. B. Visscher. He was a National Research Council fellow at Caltech with Linus Pauling and on the faculty of the Harvard Medical School from 1957 to 1967, with a year at Cambridge University, before becoming Director of the Department of Biophysics and Pharmacology at Merck Sharp & Dohme Research Laboratories and then Executive Director for Basic Medical Sciences at the Merck Institute for Therapeutic Research.

(1) Perutz, M. F.; Mathews, F. S. *J. Mol. Biol.* **1966**, *21*, 199.
(2) Bresloff, J. L.; Crothers, D. M. *J. Mol. Biol.* **1975**, *95*, 103.
(3) Koshland, D. E. *Cold Spring Harbor Symp. Quant. Biol.* **1963**, *12*, 88.
(4) Reeke, G. N.; Hartouck, J. A.; Ludwig, M. L.; Quiocho, F. A.; Steitz, T. A.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 2220.
(5) Meadows, D. H.; Roberts, G. C. K.; Jardetzky, O. *J. Mol. Biol.* **1969**, *45*, 491.
(6) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88.
(7) Yguerabide, J.; Epstein, H. F.; Stryer, L. *J. Mol. Biol.* **1970**, *51*, 573.
(8) Mendelson, R. A.; Morales, M. F.; Botts, J. *Biochemistry* **1973**, *12*, 2250.
(9) Jardetzky, O.; Akasaka, K.; Vogel, D.; Morris, S.; Holmes, K. C. *Nature (London)* **1978**, *273*, 564.
(10) Highsmith, S.; Akasaka, K.; Konrad, M.; Goody, R.; Holmes, K.; Wade-Jardetzky, N.; Jardetzky, O. *Biochemistry* **1979**, *18*, 4238.
(11) Linderstrom-Lang, K. *Chem. Soc. Spec. Publ.* **1955**, *2*, 1.
(12) Kirkwood, J. G. *discuss. Faraday Soc.* **1955**, *20*, 78.
(13) Weber, G. *Adv. Protein Chem.* **1975**, *29*, 1.
(14) Jardetzky, O.; Jardetzky, C. D. *Methods Biochem. Anal.* **1962**, *9*, 235.
(15) Frauenfelder, H.; Petsko, G. A.; Tsernoglou, D. *Nature (London)* **1979**, *280*, 558.
(16) McCammon, J. A.; Karplus, M. *Annu. Rev. Phys. Chem.*, **1980**, *31*, 29.

0001-4842/81/0114-0291$01.25/0

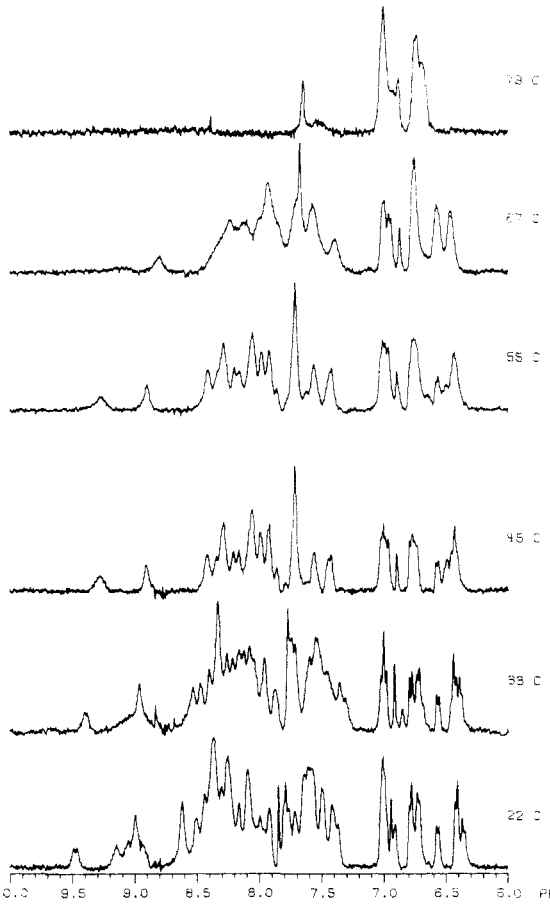

Figure 1. ^1H NMR spectra at 360 MHz of 0.2 mM HP in H_2O, high phosphate buffer, pH 7.5, as a function of temperature. All spectra taken with the Redfield technique, 2.82-s delay between pulses, 2000 scans. The carrier frequency was in the 8.7-ppm region.

shifts (and coupling constants) by group rotation, and the measurement of relaxation times. Our emphasis is on relaxation studies, which in principle can provide the most detailed information.

Hydrogen Exchange

The monitoring of hydrogen exchange by NMR has a significant advantage over exchange studies by other techniques. Provided individual spectral lines can be resolved and assigned, the sites at which exchange is occurring can be identified and the rates of exchange at different sites can be compared. This was first shown by Glickson, McDonald, and Phillips[17] for tryptophan in lysozyme, and extensive mapping of hydrogen exchange in the bovine pancreatic trypsin inhibitor,[18] oligonucleotide duplexes,[19] and tRNA[20] has been carried out since.

The type of information that is obtained can be illustrated by our recent study of the structure of the N terminal ("headpiece") fragment of the *lac* repressor.[21] Figure 1 shows the 360-MHz ^1H NMR spectra of the exchangeable protons of the headpiece in H_2O. In ^2H_2O only the very slowly exchanging protons (lifetimes >100 s) can be observed, while in H_2O those with lifetimes in the range of 0.01–100 s can be seen as well. In the *lac*-repressor headpiece there are 23 ± 3 protons that exchange very slowly and an equal number of protons exchanging more rapidly. The former can be shown by NOE (nuclear Overhauser effect) experiments to represent H-bonded protons in the polypeptide backbone, indicating that approximately one-half of the amino acid residues are part of an extensive secondary structure. Analysis of other parts of the NMR spectrum suggests that HP contains both a C-terminal and an N-terminal helix.[22] The existence of NOE between C-terminal and N-terminal residues shows that the peptide chain is folded back on itself.[23] Since the crystal structure of the headpiece is not known, the main known structural features of the headpiece are derived from NMR data.

The exchange rates can be measured either by saturation transfer or for more slowly exchanging protons by following the decrease in line intensity as a function of time following addition of ^2H_2O. If it can be shown that the intrinsic exchange rate of a given proton is rapid, the measured rate represents the rate of structure fluctuation at the site of the proton. The assumption of a rapid intrinsic exchange rate probably holds for the amide protons of the polypeptide backbone but may not hold for nucleic acid protons.[24] Caution in interpreting the measured rate is necessary in all cases, since the effects of local electric fields on the intrinsic rate remain largely unknown. In the case of HP the secondary structure persists without extensive fluctuation over many minutes. The principal limitation of hydrogen exchange as an indicator of internal motion is that it provides no direct information on the amplitude—and hence on the nature—of the fluctuation.

Chemical Shift Averaging

Averaging of chemical shifts is also an excellent qualitative indicator of internal motion. In the free aromatic amino acids phenylalanine and tyrosine the 2,6 and 3,5 ring protons appear equivalent as a result of rapid ring rotation. When such a ring is rigidly held within the protein structure, the equivalence is abolished and separate lines are seen for the 2 and 6 or 3 and 5 protons, respectively. In many protein spectra, however, the aromatic region is simple, indicating motional averaging.

Conformational averaging of tyrosine NMR spectra was first reported by Markley and Jardetzky[25] for staphylococcal nuclease and has been extensively investigated in several proteins, especially BPTI,[26] lysozyme,[27] and parvalbumin.[28] Strictly speaking, the nature of the averaging motion cannot be deduced from the spectrum. However, for aromatic rings it is natural to assume that the dominant motion is rotation about the axis of symmetry along the C_β–C_γ bond. The rates

(17) Glickson, J. D.; McDonald, C. C.; Phillips, W. D. *Biochem. Biophys. Res. Commun.* **1969**, *35*, 492.
(18) Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1979**, *130*, 31.
(19) Hilbers, C. W.; Patel, D. J. *Biochemistry* **1975**, *14*, 2656.
(20) Johnston, P. D.; Redfield, A. G. *Nucl. Acids Res.* **1977**, *4*, 3599.
(21) Wemmer, D.; Shvo, H.; Ribeiro, A.; Bray, R. P.; Jardetzky, O. *Biochemistry* **1981**, *20*, 3351.
(22) Ribeiro, A. A.; Wemmer, D.; Bray, R. P.; Wade-Jardetzky, N. G.; Jardetzky, O. *Biochemistry* **1981**, *20*, 823.
(23) Ribeiro, A. A.; Wemmer, D.; Bray, R. P.; Jardetzky, O. *Biochem. Biophys. Res. Commun.* **1981**, *99*, 668.
(24) McConnell, B. *Biochemistry* **1978**, *17*, 3168.
(25) Campbell, I. D.; Dobson, C. M.; Ratcliffe, R. G. *J. Magn. Reson.* **1977**, *27*, 455.
(26) Cited in: Jardetzky, O. "Ciba Foundation Symposium on Molecular Properties of Drug Receptors"; Churchill: London, 1970; p 113.
(27) Karplus, S.; Snyder, G. H.; Sykes, B. D. *Biochemistry* **1973**, *12*, 1323.
(28) Cave, A.; Parello, J. "Biophysics of Membranes and Intracellular Communications"; Balian, A.; Chabre, M., Eds.; North-Holland: Amsterdam, in press.

can only be estimated when motional averaging is slow and generally fall in the range of 10^2–10^5 s^{-1}. In the case of Tyr-35 of BPTI a molecular dynamic calculation[29] based on the assumption that the rate-limiting step is the relaxation of the protein structure surrounding the ring has led to a reasonable agreement between the calculated and observed rates. As in the case of hydrogen exchange, observation of motional averaging provides no direct information on the amplitudes of the motions involved.

Relaxation Studies

NMR relaxation parameters, by virtue of their sensitivity to both the frequencies and the amplitudes of motions, are potentially the richest source of experimental information on macromolecular dynamics, largely for motions in the frequency range of 10^6–10^{12} s^{-1}. This range may not include some of the functionally more interesting motions; however, biological macromolecules are highly structured systems, in which forces can be coupled. It is therefore likely that the interesting slower fluctuations will ultimately be understood as subharmonics of the more rapid motions or solitons. An understanding of the dynamics of a macromolecule over the entire frequency range will be required for the understanding of any specific motion.

The interpretation of the measured relaxation parameters presents formidable problems and can be carried out at several different levels of sophistication, with some useful information accessible at each level.

Detection of Segmental Motion in Large Proteins

Large differences in the observed relaxation rates in macromolecules have long been known to indicate greater mobility in some parts of the molecule relative to others.[14] In small proteins or nucleic acid fragments (M_w 6000–20 000), the overall rotational diffusion of the molecule makes a significant contribution to the observed relaxation rate, and qualitative arguments relating differences in the relaxation rates to different modes of internal motion are difficult to sustain. However, in large proteins or nucleic acids (M_w >100 000), simple models—e.g., diffusion of a rigid sphere or of an ellipsoid of revolution, or a spinning top on a rigid sphere—predict very broad NMR spectral lines (>100 Hz) and short T_2. The appearance of narrow lines in the spectrum can therefore be taken as an indication of more extensive internal mobility. A good example of this phenomenon is seen in the 360-MHz ^1H NMR spectra of the *lac* repressor (Figure 2).

Comparison of the spectra of the native repressor with that of the cleavage products reveals that the majority of the narrow lines in the repressor spectrum originate from the headpiece. The headpiece is, however, clearly not a random coil. The existence of structure is most readily apparent in the aromatic and the high-field methyl region of the spectrum. The four tyrosines of HP are not equivalent at room temperature and only become equivalent on denaturation at 70 °C. The same behavior is seen for the high-field methyls. The conclusion can therefore be reached from this simple experiment that the headpiece is a separate structured domain in the repressor molecule, attached to the core by a short flexible stalk of residues 51–59.[30]

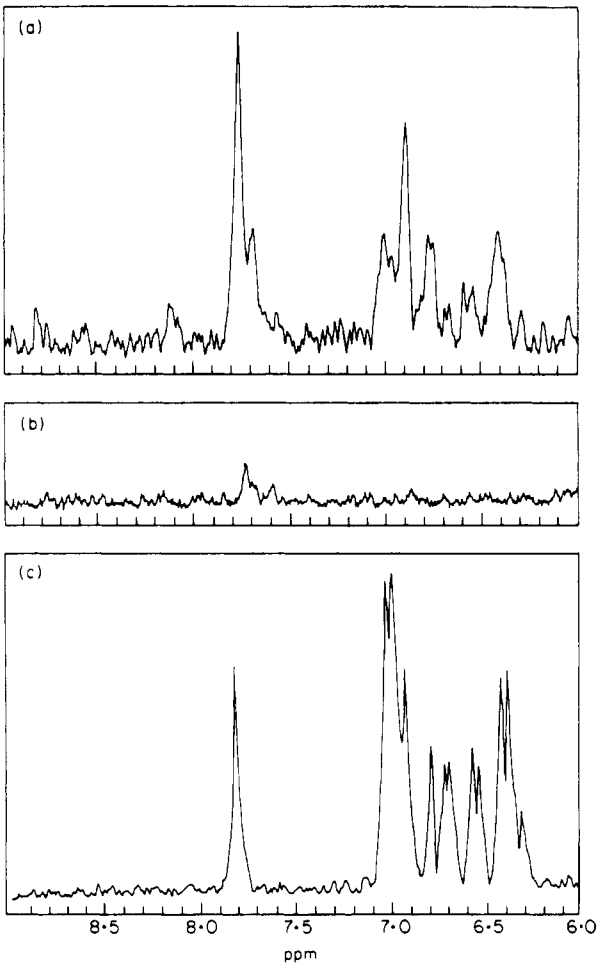

Figure 2. Resolution-enhanced ^1H NMR spectra at 360 MHz of the 6–9 ppm region of (a) intact *lac* repressor protein; (b) T core; (c) headpiece.

Other examples of segmental motion have been described for the coat protein of tobacco mosaic virus[9] and the muscle protein myosin and its proteolytic fragments: heavy meromyosin (HMM) with subfragments S1 and S2 and LMM, the light meromyosin.[10,31,32]

Qualitative observations of this type provide clear evidence for the existence of motions rapid compared to the rate of rotational diffusion of the macromolecule as a whole, and the relatively mobile regions within the molecule can be identified.[9,10,30] However, no conclusions concerning the nature of the motion are possible. For characterization of the motion a mathematical analysis which correctly predicts the observed relaxation rates is necessary.

Analysis of NMR Relaxation Data Using Motional Models

Quantitative analysis of NMR relaxation measurements requires first an accurate knowledge of the dominant relaxation mechanism. The measured relaxation rates ($1/T_1$, $1/T_2$, NOE) can be related to the rates of molecular motion by a time-independent proportionality factor which contains the relevant molecular constants and differs for different mechanisms—

(29) Gelin, B. R.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2002.
(30) Wade-Jardetzky, N. G.; Bray, R. P.; Conover, W. W.; Jardetzky, O.; Geisler, N.; Weber, K. **1979**, *128*, 259.
(31) Highsmith, S.; Jardetzky, O. *FEBS Lett.* **1980**, *121*, 55.
(32) Highsmith, S.; Jardetzky, O. *Biochemistry* **1981**, *20*, 780.

e.g., for dipolar, chemical shift anisotropy or quadrupolar interactions.[33] Since the molecular motions are random and only their average effects are observed, their effectiveness is measured by a statistical function, known as the spectral density function $J(\omega)$.[33]

The exact form of $J(\omega)$ strongly depends on the motional model used. Generally $J(\omega)$ will be a function of the rates of motion λ (or the correlation times $\tau = 1/\lambda$) for different motions and of appropriate amplitude factors ϕ. For most models, $J(\omega)$ will contain a weighted sum of Lorentzian factors $\lambda/(\omega^2 + \lambda^2)$ (cf. below).

It was originally pointed out by Allerhand et al.[34] that the simplest form of analysis of relaxation measurements on macromolecules is possible for proton-bearing carbons, since the relaxation mechanism of the ^{13}C nucleus is in this case dipolar and dominated by the neighboring proton at 1.09 Å; other contributions can usually be neglected. More recently it has been shown that the relaxation of phosphates in the ^{31}P spectra of nucleic acids is also dominated by a dipolar mechanism and permits a similar simple analysis.[35] Analysis of relaxation in ^1H spectra of macromolecules is more complicated because of extensive cross-relaxation.[33] Relatively little has as yet been done to accurately identify the various contributing factors. Most of the reported work from which quantitative conclusions have been drawn is therefore based on ^{13}C and ^{31}P measurements.

For the dipolar mechanism, a simple form of the equations for relaxation rates, neglecting cross-relaxation and other terms, can be used,

$$\frac{1}{T_1} = K_1 R^{-6}\{J(\omega_H - \omega_C) + 3J(\omega_H + \omega_C)\}$$

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

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

where K_1 and K_2 are products of atomic constants, R is the distance between nuclear dipoles, and $J(\omega)$ is the spectral density function to be calculated for a specific motional model; ω_H and ω_C are the resonance frequencies for hydrogen and carbon, respectively.

Our recent study of internal motions in double-stranded DNA[34,36,37] can serve as an example of the possibilities and limitations of model building. The qualitative observation that lines 2 orders of magnitude narrower than those expected on the assumption that DNA is a rigid rod are seen in ^1H, ^{13}C, and ^{31}P spectra of 150–600 base pair long DNA fragments shows unequivocally that relatively rapid internal motions prevail in the molecule (Figure 3). The known cylindrical geometry of DNA fragments and the independent knowledge of their hydrodynamic behavior make the model in this case relatively simple and well defined.

For long rods or ellipsoids which experience a single internal motion with a correlation time τ_{int} which is much shorter than τ_L or τ_S, the correlation times of the long and short axes, respectively, the relaxation rates T_1^{-1} and T_2^{-1} and NOE can be expressed as a linear sum of internal and overall motions:

$$\frac{1}{T_2(\mathrm{obsd})} = \frac{\alpha}{T_2(\tau_L)} + \frac{1-(\alpha+\beta)}{T_2(\tau_S)}$$

$$\frac{1}{T_1(\mathrm{obsd})} = \frac{1-(\alpha+\beta)}{T_1(\tau_S)} + \frac{\beta}{T_1(\tau_{int})}$$

$$\frac{\mathrm{NOE}-1}{\gamma_I/\gamma_S} = \frac{1-(\alpha+\beta)\phi(\tau_S) + \beta\phi(\tau_{int})}{1-(\alpha+\beta)\chi(\tau_S) + \beta\chi(\tau_{int})}$$

$$\chi(\tau) = \frac{\tau}{1+(\omega_H-\omega_C)^2\tau^2} + \frac{3\tau}{1+\omega_C^2\tau^2} + \frac{6\tau}{1+(\omega_H+\omega_C)^2\tau^2}$$

$$\phi(\tau) = \frac{6\tau}{1+(\omega_H+\omega_C)^2\tau^2} - \frac{\tau}{1+(\omega_H-\omega_C)^2\tau^2} \quad (2)$$

where the individual values $T_1(\tau)$, $T_2(\tau)$, $\phi(\tau)$, and $\chi(\tau)$ refer to expressions calculated for a single correlation time τ; α and β are coefficients which describe the relative contribution of τ_L and τ_{int} to NMR relaxation and are determined exclusively by molecular geometry.[37] Since the sum of these coefficients is normalized to 1, the coefficient corresponding to τ_S is expressed as $1-(\alpha+\beta)$.

Equation 2 does not per se imply a specific motional model, only that the internal motion is independent of and faster than the overall motion of the helix and that it may be described by a single correlation time. A particular model is necessary only to interpret the coefficients α and β in terms of angular fluctuations. This can be meaningful in the case of DNA because the constraints on such fluctuations are well defined. The dominant contribution to the line width in DNA arises from the slow motion of the long helix axis. The motion of the short helix axis is much faster and makes a much smaller contribution to the line widths. Therefore, *as a first approximation*, only internal motions relative to the *long* helix axis need to be considered. The internal motion can be modeled as a two-state fluctuation of relaxation vectors **r**(H–H, C–H, P–H) between two angles θ_1 and θ_2 defined by **r** and the long helix axis. Using this model it is possible to calculate[37]

$$\alpha \simeq \tfrac{1}{16}[1 + \cos^2\bar{\theta}\cos\Delta]$$

$$\beta = \tfrac{3}{4}\sin^2\Delta \quad (3)$$

Here $\bar{\theta} = (\theta_1 + \theta_2)/2$ is an average orientation of **r** with respect to the long axis and $\Delta = \theta_1 - \theta_2$ is the amplitude of the two-state fluctuation. τ_L and τ_S can be determined independently on monodisperse rodlike DNA fragments of defined length.

For a true internal motion, α, β, and τ_{int} should not vary with changes in the overall correlation times. Therefore the relaxation equations may be solved uniquely if T_1 and T_2 or the NOE and T_2 are measured for DNA fragments with different known values τ_L and τ_S. A map of internal motions at several positions in the DNA helix obtained from a combination of ^1H, ^{13}C, and ^{31}P data is shown in Figure 4. Motion is depicted

(33) Jardetzky, O.; Roberts, G. C. K. "NMR in Molecular Biology"; Academic Press: New York, 1981.
(34) Allerhand, A.; Doddrell, D.; Komoroski, R. *J. Chem. Phys.* **1971**, *55*, 189.
(35) Hogan, M.; Jardetzky, O. *Proc. Natl. Acad. Sci. U.S.A* **1979**, *76*, 6341.
(36) Hogan, M.; Jardetzky, O. *Biochemistry* **1980**, *19*, 2079.
(37) Hogan, M.; Jardetzky, O. *Biochemistry* **1980**, *19*, 3460.

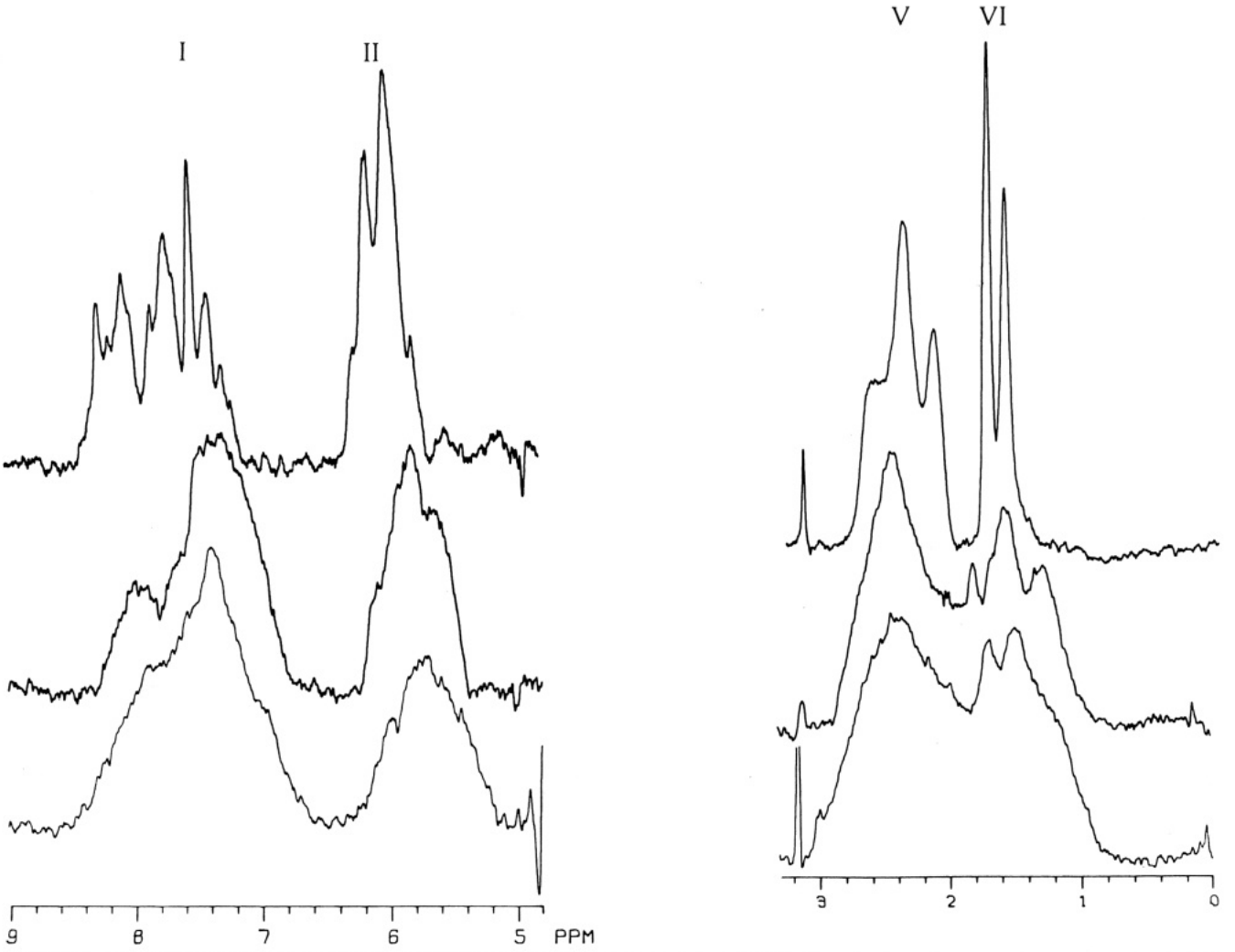

Figure 3. Effect of temperature on the ^1H NMR spectrum of 260 base pair long DNA. Fourier transform spectra were accumulated on a Bruker HXS-360 spectrometer with a delay between acquisitions equal to 5 times the longest measured T_1 relaxation time. The DNA concentration is 4.8 mM in base pairs. In all cases the spectra represent the average of 1000 scans. (a) (75 °C, denatured DNA) 20 mM NaCl, 6 mM Na_2HPO_4, 1 mM Na_2EDTA, pH 7.2, and D_2O. (b) (75 °C, native DNA) 200 mM NaCl, 6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM Na_2EDTA, pH 7.2, and D_2O. (c) (25 °C, native DNA) 200 mM NaCl, 6 mM Na_2HPO_4, 3 mM NaH_2PO_4, 1 mM Na_2EDTA, pH 7.2, and D_2O.

as an arc directed toward the long helix axis, and the amplitudes indicated in the figure are approximately equal to the vertical projection of those motions which occur in the helix.

Figure 4 summarizes several important conclusions. (1) Large amplitude motions occur at the base plane, the deoxyribose sugar, and the sugar phosphate backbone. The estimated τ_{int} for these motions is ($\sim$1–2) $\times 10^{-9}$ s. The similarity of τ_{int} cannot prove the idea that these motions are coupled to one another but is consistent with it. (2) ^1H and ^{13}C relaxation calculations of motion at position 2′ are in good agreement, both in terms of calculated amplitude and correlation time. (3) The internal motions are not only rapid but have a small calculated activation energy $\Delta E = 0.6$ kcal/mol. On the time average nearly every nucleotide in the helix experiences the motions, implying that the extremes of conformational change are nearly isoenergetic. (4) For consistency with the known hydrodynamic properties of DNA, the motion cannot lead to a local bending of the helix greater than 4° or 5°. (5) The large magnitude of ΔH for unstacking a base pair ($\geq$20 kcal) also makes it unlikely that the internal motions involve a major changes in base stacking.

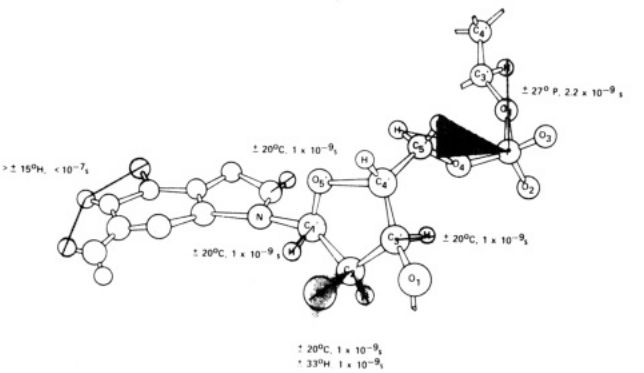

Figure 4. A map of DNA internal motions.

The fundamental problem with this model, as with all models, attractive as it may be, is that it is not necessarily unique. The two-state jump can be understood as an equilibrium between the canonical B DNA structure and a propeller-twisted structure, as proposed by Levitt,[38] or between two structures whose average corresponds to B DNA. Multistate models and

(38) Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 640.

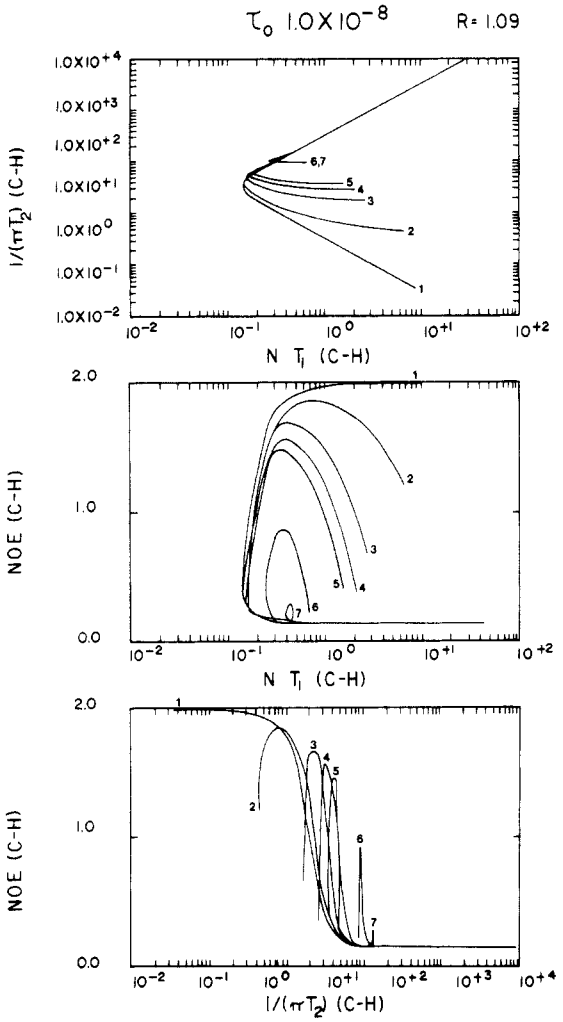

Figure 5. Relaxation phase diagram for ^1H–^{13}C relaxation in a spherical macromolecule with $\tau_0 = 1.0 \times 10^{-8}$ s. (1) Isotropic rotation; (2) 90° wobble on a cone superimposed on a three-state jump; (3) three-state jump; (4) 71° wobble on a cone; (5) 90°; (6) 45°; (7) 15° 2-state wobble.

oscillations about a single energy minimum can also account for the data.

Many different types of motional models have been proposed for the analysis of NMR relaxation data and are discussed in a recent review by London.[39] An important question, however, remains: what physical meaning can be attributed to the rate and amplitude parameters calculated from measured relaxation rates, using a particular model?

Our first approach to answering this question was to determine how different motional models are reflected in the calculated relaxation rates.[40] To compare the predictions of different models, we found it useful to plot T_1 vs. T_2, T_1 vs. NOE, and T_2 vs. NOE, as shown in Figure 5. We call such plots relaxation phase diagrams. A line on a relaxation phase diagram represents the locus of all possible combinations of two relaxation parameters that can be accounted for by a given model at a given observing frequency. If a pair of observed values of, say, T_1 and T_2 does not fall on this line, the model is clearly inappropriate for the analysis of the data. A very important, but often neglected, point immediately emerges from the study of such diagrams—i.e., that a minimum of *two measured* relaxation parameters is necessary to test any model. A correlation time τ can *always* be calculated from a single relaxation measurement (e.g., T_1) whatever the model used. To be physically meaningful, on the other hand, a correlation rate λ and amplitude ϕ must account for *all* relaxation parameters measured at *all* frequencies. Thus, much more extensive studies of relaxation behavior in macromolecules than have so far been reported must be carried out before specific conclusions about internal motions can be drawn. Figure 5 shows a comparison of several specific motional models, assuming nearest-neighbor ^{13}C–^1H relaxation at 90 MHz and isotropic overall rotation of the macromolecule.

Several general conclusions discussed in detail elsewhere[41] can be drawn from this type of study. The most important of these is that the correlation times or rate parameters are more readily assigned a unique value than are the amplitude parameters. Many cases can be found in which several models, differing largely in the formulation of amplitude factors, account equally well for the experimental data. The dual danger of model building is therefore that motions which actually contribute to relaxation will be suppressed in the analysis of data by an unfortunate choice of the model, or, alternatively, that nonexistent motions will give a satisfactory prediction of the data and will therefore be assumed to exist. If more than one model accounts for a given set of measurements, the true information content of the set is clearly limited. It is important to know these limits. In order to define them we have developed a rigorous general formalism which permits the identification of the relative contributions to relaxation made by different motions occurring in a macromolecule without resort to any specific model.[42] Our method bears a formal resemblance to the harmonic analysis of vibrations. Given a large enough data set, it permits the discovery of all motional frequencies contributing to relaxation and provides a useful guideline for the formulation of appropriate equations of motion on which the ultimate interpretation of relaxation data will eventually rest.

Frequency Analysis of Relaxation Data

The sole assumption necessary for our formulation is that the resultant of all motions contributing to relaxation is a Markov process. (The alternative assumption that each contributing motion is a Markov process leads to an equivalent result.) This assumption should hold for all molecular systems which have no memory. On this assumption it is possible to show, by use of the general theory of Markov processes, that $J(\omega)$ can be expressed as a function of the eigenvalues (rate parameters) λ and eigenfunctions (amplitude parameters) ϕ of a transition operator Ω describing the transitions occurring in relaxation:[41,42]

$$\Omega\phi_n = \lambda_n\phi_n \qquad (4)$$

The observed relaxation parameters $\rho_1 = 1/T_1$, $\rho_2 = 1/T_2$, and $\rho_3 = (\text{NOE} - 1)/T_1$ can then be expressed as linear sums of corresponding relaxation rates for simple

(39) London, R. E. In "Magnetic Resonance in Biology"; Cohen, J. S., Ed.; Wiley-Interscience: New York, in press.
(40) King, R.; Maas, R.; Gassner, M.; Nanda, R. K.; Conover, W. W.; Jardetzky, O. *Biophys. J.* **1978**, *24*, 103.
(41) Jardetzky, O. In "NMR and Biochemistry"; Opella, S. J., Lu, P., Eds.; Marcel Dekker, New York, 1979; p 141.
(42) King, R.; Jardetzky, O. *Chem. Phys. Lett.* **1978**, *55*, 15.

Table I
BPTI Typical Motional Frequencies (λ) and Amplitudes (α) from ^{13}C Relaxation Data

chemical shift,[a] (assigned residue(s))	field strength, MHz	T_1, s	T_2, s	NOE	motion 1		motion 2		motion 3	
					λ_0, Hz	α_0, %	λ_1, Hz	α_1, %	λ_2, Hz	α_2, %
7.55	45	0.256	0.100	2.77	6×10^8	2	1×10^7	1	2×10^{10}	97
(Ile^{18} or Ile^{19} CH_3)	90	0.370	0.135	2.46	6×10^8	3	1×10^7	1	3×10^{10}	96
13.129	45	0.322	0.068	1.73	6×10^8	5	1×10^7	1	7×10^{10}	94
(Met^{52} SCH_3)	90	0.350	0.057	1.74	6×10^8	10	1×10^7	1	8×10^{10}	89
16.708	45	0.276	0.085	2.61	6×10^8	1	1×10^7	1	2×10^{10}	98
(Ala^{16,27,58} CH_3)	90	0.242	0.083	2.19	6×10^8	5	1×10^7	<1	2×10^{10}	94
33.286	45	0.111	0.026	1.67	6×10^8	25	3×10^7	10	2×10^{10}	65
(Asp β-CH_2)	90	0.122	0.033	1.55	6×10^8	48	3×10^7	7	1×10^{10}	45
39.355	45	0.303	0.082	2.13	6×10^8	7	1×10^7	1	3×10^{10}	92
(Lys ϵ-CH_2)	90	0.264	0.050	2.20	6×10^8	14	1×10^7	3	2×10^{10}	83
51.738	45	0.117	0.062	1.45	6×10^8	33	2×10^8	35	9×10^{11}	32
(Ala^{58} α-CH)	90	0.220	0.062	1.18	6×10^8	23	2×10^8	61	9×10^{11}	15
53.546	45	0.106	0.044	1.27	6×10^8	1	2×10^8	74	3×10^{10}	24
(unassigned α-CH)	90	0.207	0.036	1.20	6×10^8	63	1×10^7	6	2×10^{10}	31
116.866	45	0.121	0.026	1.16	6×10^8	4	9×10^7	87	9×10^{10}	9
(Tyr^{3,5} C)	90	0.322	0.035	1.15	6×10^8	13	1×10^8	87	3×10^{10}	<1

[a] Relaxation data at 17 °C, pD 5, chemical shifts are referenced to external Me_4Si.

isotropic motion at each of the frequencies occurring in the system, i.e.

$$\rho_1 = \sum_i^N \alpha_i \rho_{1i} \lambda_i$$

$$\rho_2 = \sum_i^N \alpha_i \rho_{2i} \lambda_i$$

$$\rho_3 = \sum_i^N \alpha_i \rho_{3i} \lambda_i \qquad (5)$$

with the constraints

$$\sum_i^N \alpha_i = 1, \ \alpha_i \geq 0 \qquad (6)$$

where $\alpha \propto \langle \phi^2 \rangle$ and $\rho_i(\lambda_i) = K \alpha_i \lambda_i / (\omega_0^2 + \lambda_i^2)$, with K a constant determined by the relaxation mechanism and ω_0 the observing frequency. It can further be shown using Caratheodory's theorem[43,44] that the extent to which the analysis can be rigorously carried out is data limited. If there are N experimental values, it is always possible to find an N-term Lorentzian expansion to fit the data. Given a sufficiently large set of data, it is therefore possible to identify all motional frequencies occurring in the macromolecular system and their *relative* contributions to relaxation. Different motions with identical rates $\lambda_i = \lambda_j$ remain indistinguishable from each other.

This method of analysis can be illustrated by our recent study of ^{13}C relaxation data on BPTI, M_w 6500,[45,46] by use of a computer data-fitting program based on eq 5. The analysis proceeds sequentially, assuming one ($J = 1$), two ($J = 2$), or more ($J = n$) motional frequencies. As an object for the initial comprehensive study of relaxation behavior, BPTI has several advantages: (1) it is small, so that many individual lines can be resolved and in part assigned to specific amino acid residues in the sequence; (2) its crystal structure is known; (3) it is thought to be *relatively* rigid, considering the extensive β-pleated sheet backbone, so that one might expect the number of terms in eq 5 to be relatively small; (4) theoretical calculations of its molecular dynamics have been made, and the results of the calculations can be compared to the rates and amplitudes derived from the experiment.

An additional constraint on the analysis using eq 5 can be introduced if the overall rotational diffusion rate of the macromolecule, λ_0, is known. For BPTI, $\lambda_0 = 6 \times 10^8$ s^{-1} has been determined by light scattering and is assumed in the analysis. A representative set of ^{13}C relaxation data for assigned peaks[47] in the ^{13}C spectrum of BPTI is given in Table I.

The requirements of the analysis imposed by the structure of the protein are (1) that a term containing the same λ_0 be common to all groups and (2) that, for a side chain containing several carbons, the motion of each succeeding carbon includes the motion of the preceding carbon. Thus, if the motion of the C_α in alanine can be described by two terms, λ_0, λ_1, the motion of the C_β methyl will contain three terms, λ_0, λ_1, λ_2. The relative contribution of these terms will be expressed by the relative magnitudes of α_0, α_1, and α_2. Thus, if the motion of C_α makes a negligible contribution to the relaxation of C_β, α_1 will be negligibly small.

Table I also summarizes representative results of the analysis. It is noteworthy that none of the data for side-chain carbons can be fitted on the assumption $J = 1$, i.e., with λ_0 as the sole rate-determining relaxation. The relaxation of some groups can be accounted for in the second step of the analysis, with $J = 2$. For these groups, a second frequency in the range of $\lambda_1 = (2\text{–}6) \times 10^8$ s^{-1} appears to be common. We interpret this to mean that these groups are relatively rigidly held and the relaxation parameters reflect the asymmetric diffusion of the protein: BPTI is a pear-shaped molecule which be approximated by an ellipsoid of revolution with an axial ratio 1:3. For alanine methyls, a third rate, $\lambda_2 = 10^{10}$–10^{11}, is necessary, which in most cases dominates the relaxation process. Particularly noteworthy is the case of the C-terminal Ala-58, where both the C_α and C_β relaxation can be accounted for by $J = 2$, $\alpha_0 = 0.05$, $\alpha_1 = 0.95$, and $\lambda_1 = 10^{11}$, indicating that the residue

(43) In Eggleston, H. G. "Convexity"; Cambridge University Press: London, 1963; p 35.
(44) King, R.; Jardetzky, O., to be published.
(45) Jardetzky, O.; Ribeiro, A. A.; King, R. *Biochem. Biophys. Res. Commun.* **1980**, *92*, 883.
(46) Ribeiro, A. A.; King, R.; Restivo, C.; Jardetzky, O. *J. Am. Chem. Soc.* **1980**, *102*, 4040.
(47) Richarz, R.; Wüthrich, K. *Biochemistry* **1978**, *17*, 2263.

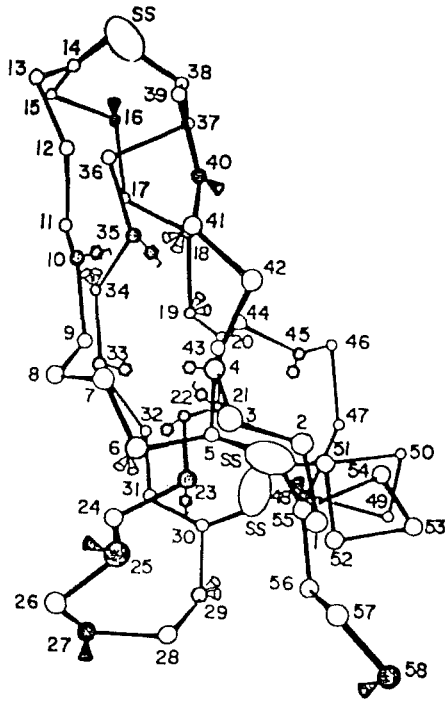

Figure 6. Location of assigned groups in the BPTI structure.

is swinging freely with respect to the remainder of the structure. This is already evident from the long T_2 values in resolution enhanced 2DFT spectra.[48] For most groups, a small but significant slow component ($\alpha_3 = 0.1–0.2$; $\lambda_3 = (2–5) \times 10^7\ s^{-1}$) also appears. Our recent calculations[44] indicate that this frequency may reflect the passage of a soliton through the peptide backbone.

Correlation of these findings with molecular structure can be made insofar as the assignments of individual lines are known. As shown in Figure 6, the rapid rotation of alanine methyls for Ala-16, -27, -40, and -58 protruding from the surface is predictable from the model. So is the relatively slow rotation of the severely constrained methyls of Val-34 or Ile-18, -19. The relaxation of the methyls can be accounted for, almost in its entirety, in terms of the anisotropic diffusion of the relatively rigid structure and the rotation of the methyl groups. Ala-58 is an exception, as already noted. The slow component in the $J(\omega)$ for CH, CH_2, and CH_3 groups, on the other hand, indicates that, in addition, there exists a slow warp of the backbone structure possibly arising from collisions between protein molecules—but more likely from a soliton. To affect relaxation, the motion must be of relatively high amplitude. The analysis of relaxation data on this protein thus confirms the notion that it is relatively rigid on the relaxation time scale, subject to (a) a slow deformation of the backbone and (b) side-side rotations at frequencies varying from 10^7 to $10^{11}\ s^{-1}$, as determined by steric constraints in the local environments.

How does this picture compare to the predictions of the molecular dynamics calculations? Since the detailed calculations predict deformations on a picosecond time scale, a comparison is possible only for the high-frequency motions observed by NMR. It is likely that the high-frequency component seen in the motions of side chain and some α-carbons correspond to those found in the molecular dynamics simulations. The motions of many other carbons do not show this frequency. It should be borne in mind that the NMR experiment is sensitive to high amplitude motions, and if such motions did exist in the frequency range $10^{10}–10^{12}\ s^{-1}$, they would dominate the observed values of T_1, T_2, and NOE. The fact that they do not appear in the analysis of relaxation data can be taken as an indication that, insofar as the motions at these frequencies predicted by molecular dynamics simulations are ubiquitous, they cannot have a very high angular amplitude. That is, for these carbons they do not involve displacements equivalent to rotations of the C–H internuclear vector over wide angles ($\theta > 30–45°$). Similarly, if a concerted high amplitude "breathing" motion or a continuous "rapid" unfolding or refolding of the protein, as proposed by Anfinsen,[49] exist, they must be slow on the relaxation time scale ($\lambda < 10^7\ s^{-1}$).

Conclusion

Analysis of this type, while limited to the determination of motional frequencies, has the important advantage of permitting the discovery of existing frequencies, which could be suppressed by an unfortunate choice of a specific simplified motional model. The ultimate bridging of the gap between theory and experiment will come when it becomes possible to construct correlation functions—and hence spectral density functions—from equations of motion, rather than from simplified models. An attempt in this direction has recently been made by Levy and Karplus[50] who reported reasonable agreement between their calculated and our[46] measured parameters for rapidly moving side chains. While theoretical methods suitable for extending such calculations to the entire experimentally accessible frequency range are being developed, much can be learned about the occurrence and distribution of different motional frequencies in different parts of biological macromolecules by using the analysis of experimental data outlined here.

All members of the SMRL staff in the period 1975–1980 and many collaborators have contributed to the program outlined in this review. The specific examples are taken from the work of M. Hogan, R. King, A. Ribeiro, N. Wade-Jardetzky, and D. Wemmer. The work has been supported by grants from the National Institutes of Health (RR00711 and GM18098) and the National Science Foundation (GP23633 and PCM78-07930).

(48) Jardetzky, O.; Conover, W. W.; Sullivan, G. R.; Basus, V. J. *Proc. Congr. AMPERE, Tallinn, USSR, 20th*, **1979**, 19.

(49) Anfinsen, C. *Science (Washington, D.C.)* **1973**, *181*, 223.

(50) Levy, R. M.; Karplus, M. *Biophys. J.* **1981**, *33*, 136a.