J. Mol. Biol. (1992) **224**, 1087–1101

Internal Motional Averaging and Three-dimensional Structure Determination by Nuclear Magnetic Resonance

Carol Beth Post

Department of Medicinal Chemistry
Purdue University
West Lafayette, IN 47907, U.S.A.

(Received 9 July 1991; accepted 13 December 1991)

Dynamic averaging effects from internal motions on interproton distances estimated from nuclear Overhauser effects (NOE) are determined by using a molecular dynamics simulation of lysozyme. Generalized order parameters measuring angular averaging and radial averaging parameters are calculated. The product of these two parameters describes the full averaging effects on cross-relaxation. Analysis of 2778 non-methyl NOE interactions from the protein interior and surface indicates that distances estimated by assuming a rigid molecule have less than 10% error for 89% of the NOE interactions. However, analysis of 1854 methyl interactions found that only 68% of the distances estimated from cross-relaxation rates would have less than 10% error. Qualitative evaluation of distances according to strong, medium and weak NOE intensities, when used to define only the upper bound for interproton separation, would misassign less than 1% of the distance constraints because of motional averaging. Internal motions do not obscure the identification of secondary structure, although some instances of significant averaging effects were found for interactions in α-helical regions. Interresidue NOEs for amino acids more than three residues apart in the primary sequence are more extensively averaged than intraresidue or short-range interresidue NOEs. Intraresidue interactions exhibit a greater degree of angular averaging than those involving interresidue proton pairs. An internal motion does not equally affect all NOE interactions for a particular proton. Thus, incorporation of averaging parameters in nuclear magnetic resonance structure determination procedures must be made on a proton-pair-wise basis. On the basis of the motional averaging results, particular fixed-distance proton pairs in proteins are suggested for use as distance references. A small percentage of NOE pairs localized to three regions of the protein exhibit extreme averaging effects from internal motions. The regions and types of motions involved are described.

Keywords: n.m.r. structure; methyl rotation; molecular dynamics simulation; lysozyme dynamics; NOE order parameters

1. Introduction

The nuclear Overhauser effect (NOE†) arises from ^1H–^1H cross-relaxation, and is the basis for determination of three-dimensional structure by n.m.r. Interproton distances evaluated from NOEs generally are estimated by assuming that the protein behaves as a rigid molecule, disregarding internal motion contributions. The simplification of a single overall correlation time is made, and the NOE intensity is related to a fixed interproton distance. Initial applications of n.m.r. for structure determination have successfully demonstrated the ability of the technique. More recent efforts are being made to obtain more accurate distance estimates by accounting for multiple-spin relaxation pathways (Olejniczak *et al.*, 1986; Olejniczac, 1989; Borgias & James, 1988; Post *et al.*, 1990*b*; Thomas *et al.*, 1991) and by improving the quantification of cross-peak intensities (Denk *et al.*, 1986; Olejniczak *et al.*, 1989; Holak *et al.*, 1987). With these efforts to increase the accuracy of the n.m.r. structural solution, it is important to consider fully the rigid-molecule assumption in the interpretation of the NOE intensity. That n.m.r. relaxation parameters are sensitive to fast timescale motions of proteins has been demonstrated (Gurd & Rothgeb, 1979;

† Abbreviations used: NOE, nuclear Overhauser effect; n.m.r., nuclear magnetic resonance; r.m.s., root-mean-square; NOESY, 2-dimensional NOE spectroscopy.

0022-2836/92/081087-15 $03.00/0

Olejniczak *et al.*, 1981, 1984*b*; Torchia, 1984; London, 1989), including recent measurements of main-chain dynamics by two-dimensional, heteronuclear ^{15}N–^{1}H experiments (Kay *et al.*, 1989; Clore *et al.*, 1990). In spite of the tightly packed structure of proteins, atomic fluctuations can be large enough to contribute to n.m.r. relaxation. Practical means for accounting for internal motions in the interpretation of distances from NOE intensities have been discussed (Braun *et al.*, 1981; Baleja *et al.*, 1990; Koning *et al.*, 1991), however, the contribution from internal motions to cross-relaxation rates and NOE intensities and, consequently distances estimated from them, has not been investigated for a protein molecule as a whole.

The present study examines picosecond fluctuations in the protein lysozyme and their contribution to the ^{1}H–^{1}H cross-relaxation rate by calculation of time-correlation functions from a 102 picosecond molecular dynamics simulation of chicken lysozyme (Post *et al.*, 1986). The motions on the timescale of the trajectory include vibrational and librational motions and some dihedral angle transitions. Longer timescale events corresponding to hundreds of picoseconds up to the overall correlation time are also relevant to n.m.r. relaxation but are not sampled in a 102 picosecond trajectory. Nonetheless, examining the influence of picosecond motional averaging on distances estimated from NOE intensities provides some indication of the validity of the rigid molecule assumption.

Simulated dynamics of proteins have been used previously to investigate internal motional averaging of n.m.r. relaxation, including ^{13}C longitudinal relaxation (Levy *et al.*, 1981*a*,*b*, 1982) and ^{1}H cross-relaxation (Olejniczak *et al.*, 1984*b*; LeMaster *et al.*, 1988). The earlier studies focused on the motional properties of a small number of interactions. The investigation reported here extends other studies utilizing protein dynamics simulations by considering motional averaging effects for all proton pairs separated by less than 4·5 Å in a protein (1 Å = 0·1 nm). While specific features of motional averaging are well described by these earlier investigations, the large set of interactions (4632 pairs) examined in the present study provides a better basis for assessing the effects of motional averaging on distance estimates and for analyzing the motional characteristics of different types of interactions, such as those used to locate secondary structure and of methyl groups.

As suggested by previous analyses of fewer NOE interactions (Olejniczak *et al.*, 1984*b*; LeMaster *et al.*, 1988), picosecond motions have negligible effects on the evaluation of the large majority of distances. In contrast to earlier conclusions (Olejniczak *et al.*, 1984*b*; LeMaster *et al.*, 1988), analysis of the more complete set of interactions indicates that the full averaging effect is not diminished by the counterbalance of angular and radial components of averaging; the radial component dominates the averaging from internal motions, although the exact degree of averaging depends on the model for the rigid protein, as described in Results. We also note that even though averaging by picosecond motions of most NOE interactions is insignificant, a small percentage of NOE interactions (approx. 11% for non-methyl NOE pairs and 32% for pairs with 1 or 2 methyl protons) are averaged to an extent that distances estimated assuming a rigid molecule have a greater than 10% error. The error can lead to either underestimated or overestimated distances.

The theory for the calculation of relaxation rates is briefly outlined in the following section, followed by a description of the molecular dynamics trajectory and the proton pairs considered in this study. Motional averaging effects are described in Results by comparison of the cross-relaxation rates calculated from the molecular dynamics trajectory with the rates for a static structure. In the final section the consequences of internal motions on structure determination are stated and the motions leading to significant averaging of certain NOE interactions are discussed.

2. Theory and Methods

(a) *n.m.r. cross-relaxation and motional averaging from internal dynamics*

The NOE (Noggle & Schirmer, 1971) is the result of cross-relaxation between 2 protons; fluctuations in the local magnetic field at a proton due to the modulation of the proton dipolar interaction by the molecular motion lead to magnetic relaxation. Cross-relaxation efficiency depends on the frequency ω of motions and is a function of the spectral densities $J(\omega)$. For the case of a macromolecule undergoing isotropic, overall rotational motion and with internal motions which are independent of overall tumbling, $J(\omega)$ is of the form (Lipari & Szabo, 1982):

$$J(\omega) = \int_0^\infty \langle A(0)A(t)\rangle \exp\left(-\frac{t}{\tau_0}\right)\cos(\omega t)dt. \qquad (1)$$

The single exponential describes the decay in correlation due to overall tumbling with a correlation time τ_0 (= $1/6D$, where D is the rotational diffusion constant). The internal motion correlation function for a proton pair is:

$$\langle A(0)A(t)\rangle = \sum_{m=-2}^{2} \left\langle \frac{Y_m^2(\Omega, 0)\, Y_m^{2*}(\Omega', t)}{r^3(0)r^3(t)} \right\rangle \qquad (2)$$

$Y_m^2(\Omega, t)$ are second-order spherical harmonics and $\langle\ \rangle$ represents a time correlation function. Ω are the polar angles (θ, ϕ) for the orientation within a fixed molecular frame of the interproton vector with length r.

At long times, when the correlation in the internal motion decays to a plateau level on a timescale t_p, the internal correlation function is equal to the equilibrium orientational distribution (Lipari & Szabo, 1982; Olejniczak *et al.*, 1984*b*):

$$\left.\begin{aligned}\langle A(0)A(t)\rangle &= \sum_{m=-2}^{2} \left|\left\langle \frac{Y_m^2(\Omega, 0)}{r^3(0)} \right\rangle\right|^2 \\ &\equiv S^2 \left\langle \frac{1}{r^3} \right\rangle^2\end{aligned}\right\} t > t_p. \qquad (3)$$

Eqn (3) derives from the long-term behavior of correlation functions when the internal motion correlation function

decays rapidly to a plateau level, and does not depend on a motional model to describe the dynamics of the internuclear vector. The time-dependent contribution to the internal correlation function $\langle A(0)A(t)\rangle$ for $t < t_p$ is negligible when the correlation time for the internal motions is small relative to τ_0 and S^2 does not approach 0·0 (Lipari & Szabo, 1982). This assumption is valid here (see below) and the contribution has been neglected in this analysis. For internal motions slower than those addressed here, the time-dependence of $\langle A(0)A(t)\rangle$ is important†. The quantity S^2 is the square of the generalized order parameter for re-orientation of an internuclear vector (Lipari & Szabo, 1982). It is a measure of the spatial restriction on the internal motion due to the surrounding protein lattice. Values of S^2 range from 0 to 1·0, where 0 indicates less restricted motion and 1·0 corresponds to a rigid molecule. The decay in correlation of the interproton vector by internal motions comprises both angular and radial components. As defined by eqn (3), the angular component and the radial component of averaging are assumed to be uncoupled.

In this study, the internal motion time-correlation function is calculated from a molecular dynamics trajectory. Since the timescale of the internal motions sampled in the 102 ps trajectory is roughly 2 orders of magnitude smaller than the overall correlation time ($t_p < 0{\cdot}05$ ns and $\tau_0 \approx 10$ ns) and $S^2 > 0{\cdot}01$, the time dependence of the internal motion correlation function (eqn (2)) is unimportant (Olejniczak *et al.*, 1984*b*). The fraction of $\langle A(0)A(t)\rangle$ contributed over the period $t < t_p$ for the values of S^2 found in this study are negligible compared to the term shown in eqn (3). The internal motions are manifest in the plateau level to which the internal correlation function decays at t_p. As such, the internal motional averaging uniformly alters the spectral density according to the equilibrium orientational distribution eqn (3) (Lipari & Szabo, 1982). The spectral density after evaluating the integral takes the form:

$$J(\omega) = \frac{S^2}{4\pi}\left\langle\frac{1}{r^3}\right\rangle^2\left(\frac{\tau_0}{1+\omega^2\tau_0^2}\right). \tag{4}$$

An observed cross-relaxation rate σ_{ij}^{obs}, including internal motion contributions, can be expressed as a change relative to that which would be obtained for a rigid molecule σ_{ij}^{rig} using eqn (4):

$$\begin{aligned}\sigma_{ij}^{obs} &= \frac{2\pi}{5}\gamma^4\hbar^2[6J_{ij}(2\omega) - J_{ij}(0)] \\ &= S^2\frac{r_{rig}^6}{\langle r^{-3}\rangle^{-2}}\sigma_{ij}^{rig},\end{aligned} \tag{5}$$

where r_{rig}^6 is the interproton distance corresponding to a rigid structure. S^2 describes the angular component of the motional averaging, while the radial component is accounted for in the ratio $r_{rig}^6/\langle r^{-3}\rangle^{-2} \equiv R$. The complete effect from internal dynamic averaging on an observed cross-relaxation rate is therefore given by Q:

$$Q = S^2\frac{r_{rig}^6}{\langle r^{-3}\rangle^{-2}} = S^2R. \tag{6}$$

Q is the factor by which σ^{rig} is scaled due to internal motions. At short mixing times in a NOE experiment, where the NOE intensity is linearly related to σ (Dobson *et al.*, 1982), the effect of dynamic averaging is direct scaling of the intensity by Q.

† The spectral density, eqn (35) of Lipari & Szabo (1982), has a term which is a function of the effective correlation time of the internal motion τ_c.

(b) *Methyl interactions*

Methyl rotation occurs on a timescale of 0·01 to 0·2 ns (Olejniczak *et al.*, 1984*b*; Torchia, 1984), much faster than the time (~10 ns) for overall rotation of proteins the size of lysozyme. Given this fast rotation and the magnetic equivalence of methyl proton resonances, examination of the dynamical effects must take into account the methyl group averaging. Three approaches to methyl group averaging are considered here.

Following Olejniczak *et al.* (1984*a*,*b*), we use the 3-site jump rotational model (Tropp, 1980) where the internal correlation function is averaged over methyl proton sites for both the simulation configurations and the rigid model. That is, the rigid model has averaging of the 3 methyl protons in an otherwise static structure. The measured σ_{ij} value is related to this 3-site(Ω, r) model by the full averaging parameter:

$$Q_{mf} = \frac{\sum_{m=-2}^{2}\left|\left\langle\frac{Y_m^2(\Omega_i,0)}{r_i^3(0)}\right\rangle^M\right|^2}{\sum_{m=-2}^{2}\left|\left(\frac{Y_m^2(\Omega_i)}{r_{rig,i}^3}\right)^M\right|^2}, \tag{7}$$

where Ω_i and r_i are the polar angles and length of an internuclear vector of a methyl proton, and the superscript M denotes the average over the protons of a methyl group:

$$\begin{aligned}(f_i)^M &\equiv \frac{1}{M}\sum_{i=1}^{M}(f_i) \\ \langle f_i\rangle^M &\equiv \frac{1}{M}\sum_{i=1}^{M}\langle f_i\rangle,\end{aligned} \tag{8}$$

$\langle\ \rangle^M$ represents a time-average function and $(\ )^M$ represents a rigid-model function. M equals 3 or 9 for a methyl–non-methyl or a methyl–methyl interaction, respectively. The full methyl averaging parameter Q_{mf} is therefore the ratio of the internal correlation function from the dynamics simulation to that from the rigid model, both averaged over the methyl proton positions. Angular (Y_m^2) and radial ($r_{rig,i}^{-3}$) components are averaged in the rigid model.

Another model which might be used in the application of NOE distance restraints for methyl interactions is based on averaging r^{-3} over the 3 methyl sites, without averaging the angular terms in the correlation function (Keepers & James, 1984; Nilges *et al.*, 1991). The motional averaging parameter for the 3-site(r) model is given by:

$$Q_{mr} = \sum_{m=-2}^{2}\left|\left\langle\frac{Y_m^2(\Omega_i,0)}{r_i^3(0)}\right\rangle^M\right|^2 r_{m'}^6, \tag{9}$$

$$r_{m'} = [(r_{rig,i}^{-3})^M]^{-1/3}. \tag{10}$$

Although averaging the radial component of the correlation function over 3 sites while ignoring the angular component is not rigorously consistent, constraints based on the average of r^{-3} are readily implemented in practice and therefore were considered in this analysis.

A final model for NOE distance constraints of methyl groups is to use the geometrical center of the 3 methyl proton positions (Wüthrich *et al.*, 1983) as the reference for the constraint distance, r_{gc}. The measured σ_{ij} value is then related to the rigid-molecule value based on r_{gc} by

the motional averaging parameter:

$$Q_{gc} = \sum_{m=-2}^{2} \left| \left\langle \frac{Y_m^2(\Omega_i, 0)}{r_i^3(0)} \right\rangle^M \right|^2 r_{gc}^6. \quad (11)$$

$R(\equiv r_{rig}^6/\langle r^{-3}\rangle^{-2})$ for the different methyl models is calculated with $r_{rig} = r_{m'}$ for 3-site(Ω, r) or 3-site(r), and with $r_{rig} = r_{gc}$ for the geometrical center. Values for S^2 are from eqn (6) with Q from eqns (7), (9) or (11) and the corresponding values of R.

(c) *Molecular dynamics trajectory*

The effects of picosecond internal motions on cross-relaxation in proteins was determined by using a 102 ps molecular dynamics trajectory of chicken lysozyme (129 residues; 14,400 Da). The dynamics of 1531 atoms, including 53 ST2 water molecules of structural importance, was simulated with the CHARMM program (Brooks *et al.*, 1983). Details of the trajectory calculation have been reported (Post *et al.*, 1986). The trajectory is stable as indicated by maintaining constant temperature and energy without coupling to a heat bath, and a constant r.m.s. deviation (1·5 Å for the main chain) from the crystallographic structure. Polar protons were explicitly included in the calculated trajectory, while non-polar protons were accounted for by the extended-atom technique. Aliphatic hydrogen atoms therefore were built onto the simulation structures by using standard geometry (Olejniczak *et al.*, 1984*b*) and a bond length of 1·08 Å. The analysis of dynamic averaging used 1038 co-ordinate sets separated by 0·098 ps. Overall rotation and translation were removed from the configurations prior to the analysis.

The model chosen to represent the rigid molecule for values of r_{rig} is the optimized molecular dynamics average structure. The trajectory average structure is preferred over the crystallographic average structure for the rigid-molecule model in order to achieve a consistent comparison between static and averaged quantities given that the trajectory average deviates slightly from the crystallographic average (Post *et al.*, 1986). The averaged trajectory co-ordinates were optimized by energy minimization to remove strained internal co-ordinates before building the proton co-ordinates. An additional motivation for using the optimized average structure is that n.m.r.-determined structures most often conform to idealized geometry with low energy, thus an appropriate comparison of time-averaged distances should involve distances in a static structure with good geometry. Small differences in the extent of radial averaging were found for r_{rig} taken from the unoptimized average structure instead of the optimized average structure, but the general characteristics described here were unchanged. The differences are described below. We also note that a distance r_m calculated from the optimized average structure does not equal the time-averaged distance $\langle r\rangle$. The difference between r_m and $\langle r\rangle$ is usually less than 0·25 Å, although differences greater than 0·75 Å occur.

The radial parameter for non-methyl proton pairs is therefore:

$$R = \frac{r_m^6}{\langle r^{-3}\rangle^{-2}}, \quad (12)$$

where the time-average of r^{-3} is obtained from the trajectory and r_m^6 is calculated from the 102 ps average simulation structure after energy minimization. For methyl proton pairs, R is based on $r_{m'}$ and r_{gc} (described above) calculated from the optimized average structure. S^2 describes the angular effects and is calculated according to eqn (3). The combined parameter Q (eqn (6)) gives the complete averaging effect. Angular averaging can only reduce σ_{ij} (as noted above, $0 < S^2 < 1$), while radial averaging can enhance or decrease the NOE intensity depending on the details of the motion. Q, the product of S^2 and R, can therefore be greater than or less than 1·0.

Motions on timescales longer than ~20 ps, such as dynamics involving large loops or whole domains, are not well represented in a 102 ps simulation because of the limited sampling by the trajectory. Nevertheless, it is expected that the results are not significantly influenced by any rare dynamics events. The dominant motions modeled in the trajectory are high-frequency fluctuations† (Swaminathan *et al.*, 1982; Post *et al.*, 1989). In addition, the time-development of the atomic fluctuations averaged over all atoms is nearly converged over the simulation period, and the fluctuation time-development and magnitudes calculated from the 2 halves of the trajectory are similar (Post *et al.*, 1989). Examination of a subset of the time-correlation functions of eqn (2) also finds that the large majority of the functions decay to a plateau within 5 to 10 ps, although a small number of internal correlation functions do not reach a plateau level within 50 ps. As such, the motional effects on σ_{ij} reported here reflect well picosecond timescale motions such as vibrations, dihedral librations, side-chain dihedral transitions and small loop movements. Other, longer timescale motions could lead to greater changes in experimental NOE intensities than those estimated by this study.

(d) *Proton pairs*

Overhauser interactions for lysozyme amide protons (126 protons) and non-polar protons (696 protons of which 183 are methyl protons) were evaluated in this study. Lysozyme consists of 15% β-sheet (3-stranded, antiparallel), 34% α-helix (4 helices), and 9% 3_{10} helix (2 helices). The analysis included 2778 non-methyl proton pairs obtained with the criterion of $r_m < 4{\cdot}5$ Å, a cut-off close to the longest distance observed in an NOE experiment. A distance cutoff of 4·0 Å based on $r_{m'}$ (eqn (10)) was used to select methyl NOE pairs. The number of individual methyl proton pairs found for lysozyme was 1854. However, each methyl–methyl or methyl–non-methyl interaction was counted only once in the analysis, reducing the number of interactions to 492.

Two classes of non-methyl interactions were defined according to the surface exposure of the 2 protons. Interior protons are those with less than 0·01 Å^2 surface area accessible to a 1·4 Å probe (Lee & Richards, 1971) in the optimized average structure of lysozyme. Taking the remaining protons as surface protons gives 436 interior protons and 203 surface protons. If 2 protons of a pair are from the interior set than the pair is designated an interior pair, while a pair comprising both surface protons or one surface and one interior proton is designated a surface pair. Each class is further grouped to reflect the categories commonly used in n.m.r. structural studies as described below.

(1) Long-range interresidue pairs comprise protons from residues more than 3 residues apart in primary sequence.

(2) Short-range interresidue pairs comprise protons from residues 1, 2 or 3 residues apart in primary sequence, excluding pairs used to locate secondary structure.

† Motions with timescales less than 20 ps contribute about 80% of the r.m.s. fluctuations (Post *et al.*, 1989).

Table 1

Average values of the motional averaging parameters for groups of non-methyl NOE pairs in the interior and on the surface of lysozyme

	Pairs	$\langle Q \rangle$ [a]	$\langle S^2 \rangle$ [b]	$\langle R \rangle$ [c]
A. *Interior*				
Total	1771	0·94±0·36	0·90±0·09	1·05±0·44
Interresidue, long-range	538	0·94±0·51	0·91±0·08	1·04±0·62
Interresidue, short-range	219	0·97±0·26	0·93±0·06	1·05±0·30
Secondary structure markers	497	0·99±0·31	0·93±0·05	1·07±0·36
Intraresidue	459	0·89±0·23	0·86±0·12	1·05±0·33
Geminal	58	0·82±0·16	0·82±0·16	—
B. *Surface*				
Total	1007	0·87±0·29	0·83±0·13	1·05±0·32
Interresidue, long-range	113	0·86±0·49	0·86±0·09	0·99±0·53
Interresidue, short-range	100	1·01±0·34	0·91±0·06	1·11±0·37
Secondary structure markers	184	0·96±0·25	0·89±0·07	1·08±0·32
Intraresidue	516	0·84±0·23	0·80±0·14	1·05±0·26
Geminal	94	0·74±0·17	0·74±0·17	—

[a] Eqn (6).
[b] Eqn (3).
[c] Eqn (12), reference distance from the time-averaged structure after energy minimization.

(3) Secondary-structure markers are the pairs (Wüthrich *et al.*, 1984) αN, NN, $\alpha\beta$ for residues i, $i+3$; αN, NN for residues i, $i+2$; αN, NN, βN for residues i, $i+1$; and intraresidue αN.

(4) Intraresidue pairs comprise protons from the same residue excluding secondary structure markers and geminal pairs.

(5) Geminal pairs.

The number of interproton vectors in each group is given in Table 1 for interior and surface pairs.

3. Results

(a) *Motional averaging of interior pairs*

We compare time-averaged relaxation with relaxation in a rigid molecule, and discuss the dynamics of the interproton vector in terms of angular, S^2, and radial, R, components, and their full combination, Q. Only the full combined effect is observed experimentally. Features of Q, S^2 and R for the interior proton pairs are discussed first for the set as a whole, followed by features specific to the different groups. Each of the averaging parameters relates cross-relaxation in a flexible molecule to that in a rigid-molecule and is a scale factor of σ_{ij}^{rig} (eqn (5)) such that values $<1{\cdot}0$ or $>1{\cdot}0$ indicate diminished or enhanced cross-relaxation, respectively. Average values are listed in Table 1 and distributions of $\log(Q)$, $\log(S^2)$ and $\log(R)$ are plotted in Figure 1. (Log values are plotted in the Figures to display values corresponding to diminished and enhanced cross-relaxation rates with equal weighting.)

On the average, the full effect of internal motions is not large; $\langle Q \rangle$ for all interior pairs equals 0·94, $\langle S^2 \rangle$ equals 0·90, and $\langle R \rangle$ equals 1·05. The distribution of Q over the interior pairs is nearly symmetrical about the average ($\log\langle Q \rangle = -0{\cdot}02$) (Fig. 1). Cross-relaxation rates are enhanced by dynamic averaging almost as frequently as they are diminished because of the radial component of internal motional averaging; the distribution of Q

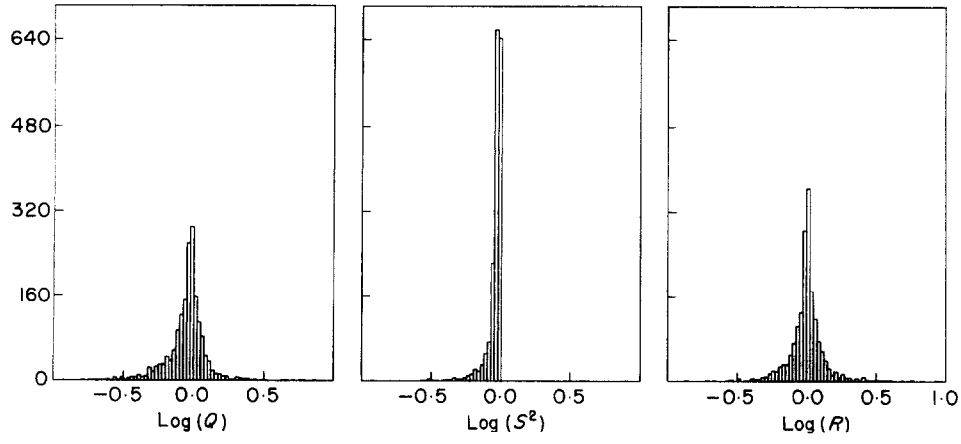

Figure 1. Histograms for 1771 interior proton pairs of $\log(Q)$, $\log(S^2)$ and $\log(R)$, the full, angular and radial averaging parameters, respectively.

Table 2
Number of extreme values for the motional averaging parameters Q, S^2 *and* R

		Extreme values[a]		
	Pairs	Q(%)[b]	S^2[c]	R[d]
A. *Interior*				
Total	1771	193 (11)	37	142
Interresidue, long-range	538	110 (20)	6	92
Interresidue, short-range	219	15 (7)	1	15
Secondary structure markers	497	20 (4)	0	14
Intraresidue	459	42 (9)	24	21
Geminal	58	6 (10)	6	
B. *Surface*				
Total	1007	132 (13)	86	60
Interresidue, long-range	113	29 (26)	3	23
Interresidue, short-range	100	10 (10)	0	10
Secondary structure markers	184	7 (4)	2	7
Intraresidue	516	65 (13)	60	20
Geminal	94	21 (22)	21	

[a] Extreme values are <0.6 or >1.7 and correspond to an error of approximately 10% or greater in the distance estimated from σ_{ij}^{obs}.
[b] Eqn (6).
[c] Eqn (3).
[d] Eqn (12), reference distance from the time-averaged structure after energy minimization.

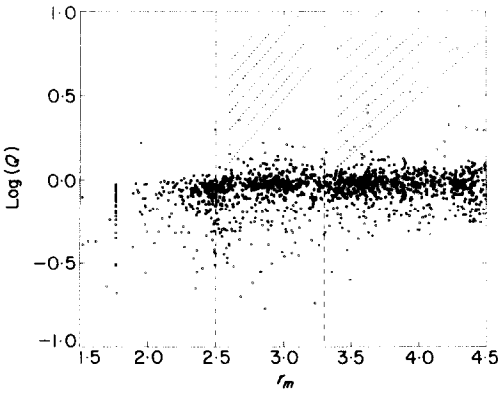

Figure 2. Values of log(Q) for 1771 interior proton pairs plotted as a function of the interproton distance from the energy-minimized, simulation average structure, r_m. Broken lines at 2·5 Å and 3·3 Å separate short, medium and long-distance categories which correspond to strong, medium and weak NOE intensities for mixing times where the intensity is linearly related to the cross-relaxation rate σ_{ij}. The shaded areas indicate regions where dynamic averaging enhances cross-relaxation to an extent that the NOE interaction would be incorrectly specified as strong or medium and assigned a distance constraint which is too short.

strongly reflects the shape of the distribution of R. The symmetrical distribution of log(R) indicates that there is no systematic decrease in the effective distance as a result of radial averaging; a strong tendency for $\langle r^{-3}\rangle^{-1/3}$ to be smaller than r_m would be manifested as a higher frequency of positive values in Figure 1. For cases where $R < 1.0$, r_m is less than $\langle r\rangle$. Although the nature of the radial averaging varies for different proton pairs, one type of distance distribution which leads to $r_m < \langle r\rangle$ is a bimodal distribution such that distances for both modes are greater than the distance corresponding to the configuration intermediate to the two modes, i.e. the average structure. In contrast to R, S^2 is restricted to values between 0 and 1·0, and is asymmetrically distributed near 1·0. Further, S^2 is narrowly distributed while the distribution for R is broader due to the strong distance dependence of σ_{ij}. (When r_{rig} is from the unoptimized average structure, the width of the distribution of R is somewhat narrower, and the value of $\langle R\rangle$ is slightly reduced to 1.04 ± 0.30.)

Because S^2 is less than or equal to 1·0 while R can be greater than 1·0, the possibility exists that effects from angular averaging may be offset by those from radial averaging in the product Q (Olejniczak *et al.*, 1984*b*; LeMaster *et al.*, 1988). However, on the whole, compensation does not occur. The distribution of Q is not narrower than that for R, as would be expected if angular-averaging effects tended to offset radial averaging effects. Further, although on average values of Q are intermediate between S^2 and R (i.e. $\langle S^2\rangle < \langle Q\rangle < \langle R\rangle$), the full averaging parameter $\langle Q\rangle$ is not nearer the rigid-molecule value 1·0 than is $\langle R\rangle$. Indeed, for some interactions radial and angularing averaging effects are reinforcing rather than offsetting in that both S^2 and R are less than 1·0, with the result that Q is less than S^2 or R (see also section (c), below).

Although $Q \approx 1.0$ for the large majority of Overhauser interactions, there is a small number of interactions with values far from 1·0, and for which motional averaging effects are significant. To examine the consequences of motional averaging on the interpretation of distances from NOESY measurements, we have considered the number of pairs for which the scale factors Q, S^2 or R deviate from 1·0 to the extent that the error in the calculated value for r_{ij} would be greater than 10%. The r^{-6} dependence of σ_{ij} makes the calculated distance insensitive to small deviations from 1·0. Extreme values are therefore defined as Q, S^2 or R less than 0·6 or greater than 1·7 ($|\log| > 0.22$) and correspond to errors of approximately 10% or more in a distance estimated from σ_{ij}. The number of interactions with extreme values for Q, S^2 or R are listed in Table 2. Of all interior pairs only 11% (193/1771) undergo motional averaging to a degree which would produce a significant error in the calculated distance.

The relationship between Q and the interproton distance r_m is shown in Figure 2 to examine the nature of dynamic averaging as a function of distance. Internal motions resulted in either enhanced or diminished cross-relaxation between proton pairs separated by medium-size distances (2.5 Å $< r_m < 3.3$ Å) and long distances (3.3 Å $< r_m < 4.5$ Å). In contrast, cross-relaxation between protons at short distances ($r_m < 2.5$ Å) is almost exclusively diminished by averaging; log(Q) is less than 0·0. In addition, in regard to the number of

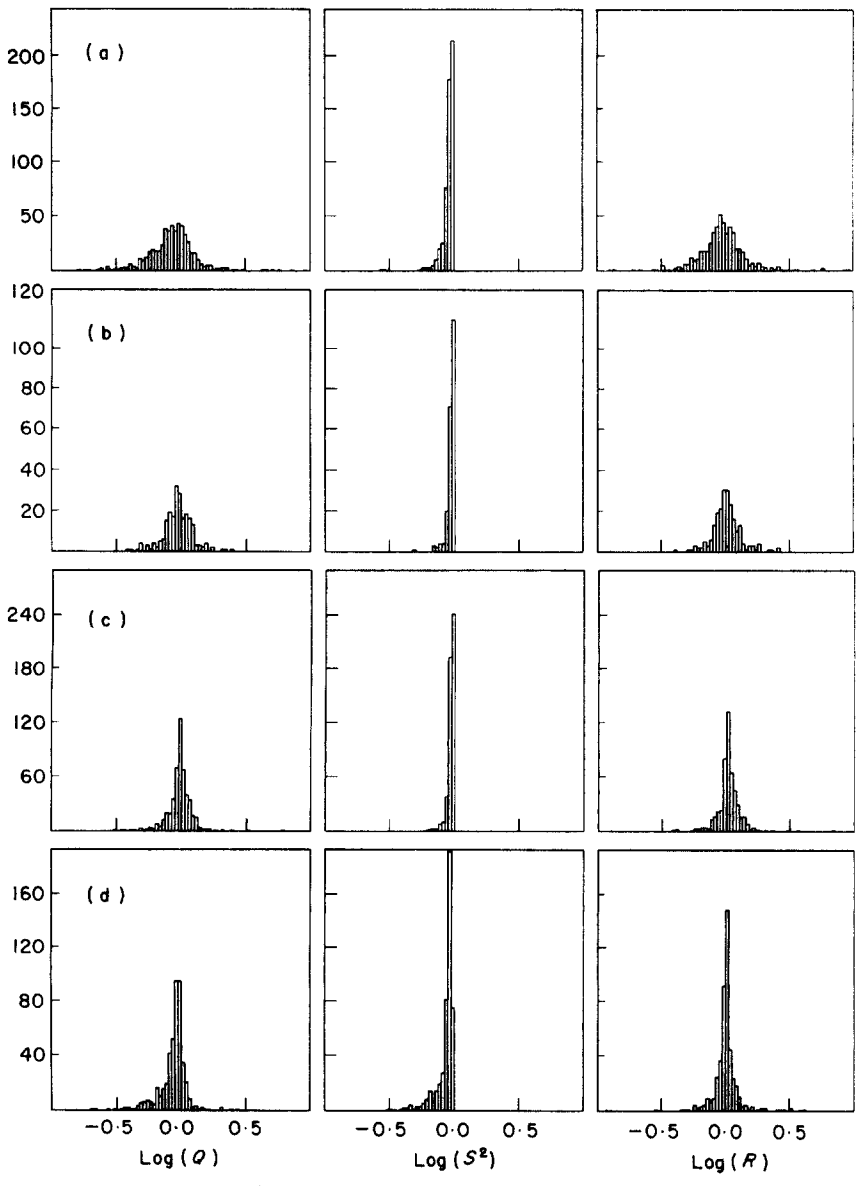

Figure 3. Histograms of log(Q), log(S^2) and log(R) for 1771 interior proton pairs. Values are grouped into (a) long-range interresidue pairs, (b) short-range interresidue pairs, (c) secondary structure markers, and (d) intraresidue pairs.

extreme averaging values shown in Figure 2, there are more interactions with decreased σ_{ij} ($Q < 0{\cdot}6$) than with increased σ_{ij} ($Q > 1{\cdot}7$).

Some structure determination procedures evaluate distances by categorizing NOE cross-peaks according to strong, medium and weak NOE intensities. Dynamic averaging effects on this type of distance evaluation were considered since a change in σ_{ij} and NOE intensity could lead to peaks being incorrectly categorized. In many of the current protocols, the NOE intensities define only the upper bound of the distance constraint while the lower bound is set by steric contact using the van der Waals term in the empirical force field (Clore & Gronenborn, 1989). As such, motionally reduced NOEs, which would give too long an apparent distance, do not lead to distance constraints with detrimental effects. However, errors from interactions which were enhanced by internal motions, and therefore incorrectly placed in the short or medium-range groups, would have negative consequences. An estimate of this type of averaging, assuming the intensity is proportional to σ_{ij} (i.e. the limit of short mixing times), is indicated in Figure 2 by the shaded areas. The result is that few interactions (<1%) appear in the areas, therefore fast internal motions are not a significant source of error for distances evaluated by this approach.

(i) *Interresidue pairs*

Distributions of log(Q), log(S^2) and log(R) for the groups of interior pairs are shown in Figure 3. $\langle Q \rangle$ equals 0·94 for long-range and 0·97 for short-range interresidue pairs. Motional averaging affects long-range interresidue interactions to a greater extent than all other types of interactions; there is a broad distribution of R (Fig. 3(a)) giving rise to extreme values of Q for 20% of the pairs (110/538), more

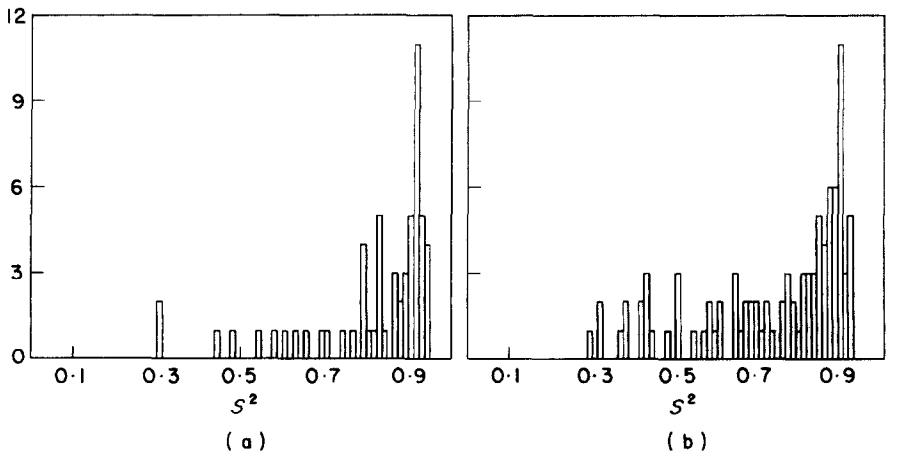

Figure 4. Histograms for the angular averaging parameter S^2 for geminal proton pairs for (a) interior pairs and (b) surface pairs.

than twice the percentage of other groups (Table 2). However, the degree of radial averaging depends on energy minimization (see below).

(ii) *Secondary-structure markers*

Because of the importance of identifying helical and sheet structural elements, the effects of picosecond motions on the cross-relaxation of secondary structure markers were examined (Fig. 3(c)). Dynamic averaging is minimal for the proton pairs used to define secondary structural elements; $\langle Q \rangle = 0{\cdot}97$ and only 4% of the pairs (20/497) have extreme values. Of the 20 extreme Q values, none is from β-sheet residues and 12 are from α-helix residues. Given the short distances involved, none of the 20 NOE intensities would be reduced to the extent that an NOE would not be observed. Interestingly, six of the 12 values from α-helices involve residues at the beginning of three of the four helices. Three of the 12 are $NN(i, i+1)$ interactions, a strong, short-distance NOE which is a hallmark for helices (Wüthrich *et al.*, 1984; Clore & Gronenborn, 1987). Two $NN(i, i+1)$ interactions are enhanced by internal motions two- to threefold, rather than reduced, so that the expected connectivity would be observable. Cross-relaxation of the third $NN(i, i+1)$ interaction is reduced by about half and would remain visible given the short distance ($\sim$2·0 Å) involved.

(iii) *Intraresidue pairs*

An asymmetry in the distribution Q for intraresidue interactions (Fig. 3(d)) contrasts with the nearly symmetrical distributions of the interresidue interactions. The asymmetry is a tendency for decreased cross-relaxation rates due to more extensive angular averaging of intraresidue interactions. $\langle S^2 \rangle$ equals 0·86, a larger deviation from the rigid-molecule value of 1·0 than found for the interresidue groups, and there are a greater number of extreme S^2 values. This angular component is reflected in a smaller $\langle Q \rangle$ (Table 1) and the slight bias in the distribution of Q toward values less than 1·0 (Fig. 3(d)). The smaller S^2 values are not offset by larger values of R. Indeed, for many interactions both S^2 and R are less than 1·0 so that the angular component reinforces the radial component, and the combined averaging result is $Q < S^2$ or R.

(iv) *Geminal pairs*

Only re-orientation of the interproton vector affects fixed-distance geminal pairs. As found for the other intraresidue interactions, the angular averaging can be significant, with S^2 values equal to 0·5 or less in some cases (see Fig. 4). Approximately 10% of the geminal interactions (6/58), a high percentage of extreme S^2 values in comparison to the other groups, are diminished by angular averaging to the extent of giving a greater than 10% error in distance. Non-negligible effects on cross-relaxation rates from motions of geminal protons have also been observed experimentally (Olejniczak *et al.*, 1981).

(b) *Comparison of interior and surface pairs*

Two classes of NOE interactions based on the surface exposure of the proton pair were defined to distinguish the dynamical properties of buried protons from those of protons only partially surrounded by protein atoms. The qualitative features of such a comparison are informative, even though the trajectory includes only structural water molecules and not bulk solvent. From $^3J_{\alpha\beta}$ coupling constants of lysozyme it appears that conformational flexibility of the side-chains of surface residues is more pronounced than that of interior side-chains (Smith *et al.*, 1991).

Interactions of surface pairs were averaged somewhat more than those from the protein interior; $\langle Q \rangle$ equals 0·87 for the surface pairs compared to 0·94 for the interior pairs (Table 1). The smaller $\langle Q \rangle$ value results primarily from angular averaging of intraresidue and geminal pairs consistent with the greater flexibility in dihedral angles of some surface side-chains observed with $^3J_{\alpha\beta}$ (Smith *et al.*, 1991).

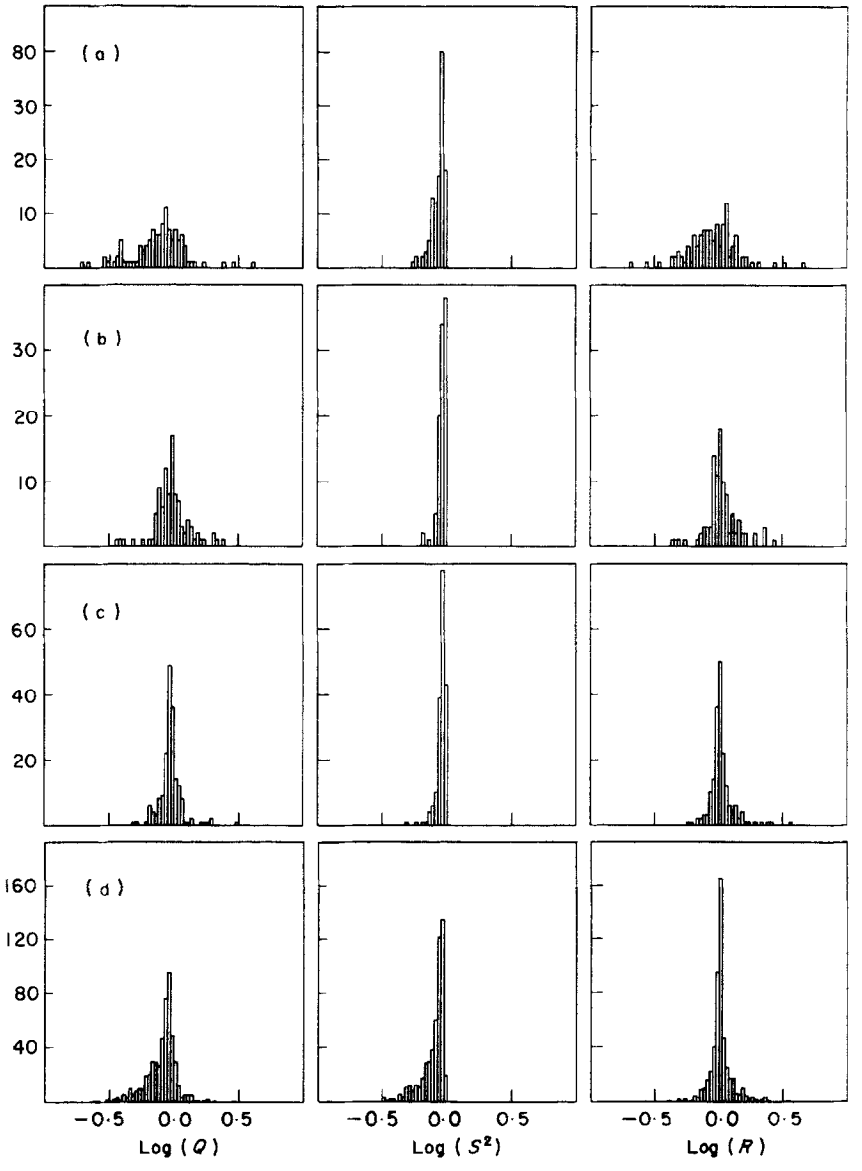

Figure 5. Histograms of $\log(Q)$, $\log(S^2)$ and $\log(R)$, for 1007 surface proton pairs. Values are grouped into (a) long-range interresidue pairs, (b) short-range interresidue pairs, (c) secondary structure markers, and (d) intraresidue pairs.

Although the surface interactions have smaller Q values on average, the general features found for proton pairs from the interior of lysozyme also apply to pairs on the surface of the molecule. Similar behavior was found in that: (1) the large majority of interactions are not significantly effected by the internal dynamics ($0{\cdot}8 < \langle Q \rangle < 1{\cdot}0$ for the individual groups except the geminal pairs) and less than 13% of the pairs (132/1007) have extreme Q values (Tables 1 and 2); (2) the distribution for R is broader than that for S^2 showing radial averaging is more pronounced than angular averaging (Fig. 5); (3) the group of interresidue interactions has the highest frequency of extreme values of Q; (4) motional averaging of secondary structure markers is negligible; and (5) more extensive angular averaging is evident for intraresidue pairs, which leads to a slight asymmetry in the distribution of Q values.

(c) *Motional averaging of methyl pairs*

In the case of methyl proton NOEs, we consider the picosecond dynamically averaged cross-relaxation rate relative to that based on different static models used to account for methyl rotation. Because R expresses the internally averaged distance relative to a static distance in the rigid-molecule model the value of Q depends on the treatment of methyl proton distances in the rigid-molecule model.

Three values for Q (eqns (7), (9) and (11)) were calculated to fit with different methyl group averages: 3-site(Ω, r) radial and angular average, 3-site(r) radial average, and geometrical center. In all three cases, the distribution of Q shows a strong tendency toward less efficient cross-relaxation due to dynamic averaging, consistent with the relatively small NOE intensities for methyl protons frequently

Table 3
Motional averaging parameters for 492 methyl–methyl and methyl–non-methyl pairs

	$\langle Q \rangle$	$\langle S^2 \rangle$	$\langle R \rangle$
3-site(Ω, r)	0·72 ± 0·24 [a]	0·81 ± 0·16	0·91 ± 0·29 [d]
3-site(r)	0·57 ± 0·21 [b]	0·64 ± 0·15	0·91 ± 0·29 [d]
Geom center	0·66 ± 0·31 [c]	0·64 ± 0·15	1·03 ± 0·39 [e]
	No. extreme values [f]		
	Q(%)	S^2	R
3-site(Ω, r)	132 (32) [a]	44	62
3-site(r)	241 (58) [b]	145	62
Geom center	180 (43) [c]	145	53

[a] Eqn (7).
[b] Eqn (9).
[c] Eqn (11).
[d] $r_{rig} = r_{m'}$, calculated from the time-averaged structure after energy minimization.
[e] $r_{rig} = r_{gc}$, calculated from the time-averaged structure after energy minimization.
[f] Extreme values are <0.6 to >1.7 and correspond to an error of approximately 10% or greater in the distance σ_{ij}^{obs}.

observed experimentally. Depending on the type of methyl averaging, $\langle Q \rangle$ ranges from 0·57 to 0·72 (Table 3), significantly smaller than the values of 0·87 and 0·94 for the non-methyl surface and interior pairs (Table 1). The small Q values result from $R < 1.0$, which reinforces the decrease in relaxation due to angular averaging. Furthermore, the number of methyl pairs with extreme averaging ranges from 32% to 58% (Table 3), a much larger percentage than that for the non-methyl pairs: 11% and 13% (Table 2).

The results in Table 3 show that the diminished cross-relaxation observed for methyl NOE interactions originates from non-rotational picosecond fluctuations as well as methyl group rotation; even relative to a rigid structure with methyl rotation, the values of Q calculated from the simulation are significantly different from 1·0. A second feature evident from the values in Table 3 is that averaging the angular terms of the correlation function Y_m^2 over the methyl group in the rigid model gives better agreement with σ_{ij}^{obs} measured from initial rates or a relaxation matrix analysis. $\langle Q \rangle$ for the 3-site(Ω, r) averaging is nearer to 1·0 than for the other two averaging models primarily because S^2 is closer to 1·0.

(d) *Apparent restraint distances*

The effect of motional averaging and different types of methyl averaging is readily seen by comparing an apparent restraint distance r_{rstr} calculated from the measured σ_{ij}^{obs} with the actual distance r_χ obtained from the static structure (Fig. 6). Assuming the large molecule limit in which $J_{ij}(0)$ dominates cross-relaxation and an accurately determined σ_{ij}^{obs}, the apparent distance is given by:

$$r_{rstr} = Q^{-1/6} r_\chi, \qquad (13)$$

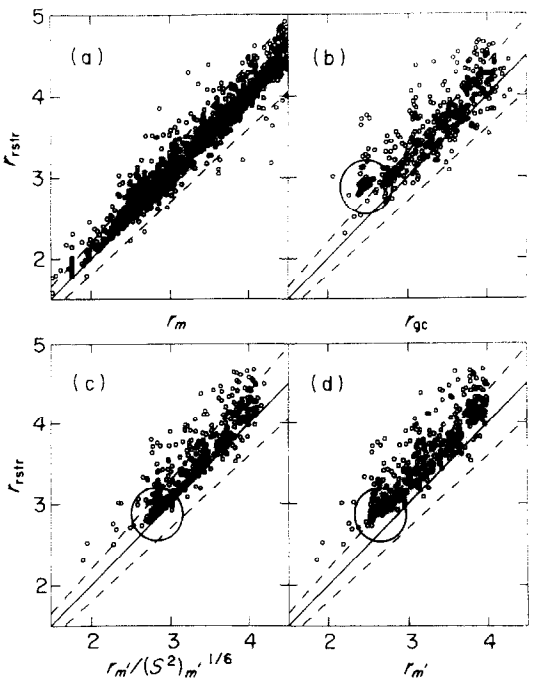

Figure 6. The effective restraint distances in the large molecule limit (eqn (13)) from NOE interactions of interior, non-methyl proton pairs (a) and methyl–methyl or methyl–non-methyl pairs ((b), (c) and (d)). The broken lines delimit the region of less than 10% deviation between r_{rstr} and the actual distance in the rigid molecule plotted on the abcissa. (b) Geometrical center: Q_{gc} from eqn (11) and $r_\chi = r_{gc}$; (c) 3-site(Ω, r): Q_{mf} from eqn (7) and $r_\chi = (S^2)_{m'}^{-1/6} r_{m'}$; (d) 3-site($r$): Q_{mr} from eqn (9) and $r_\chi = r_{m'}$.

where r_χ is from the optimized, average simulation structure. In the case of a methyl group, r_χ equals r_{gc} for the geometrical center model, $r_{m'}$ for 3-site(r) averaging, and $(S^2)_{m'}^{-1/6} r_{m'}$ for the 3-site(Ω, r) averaging (i.e. the inverse 1/6th power of the denominator in eqn (7)). Q for the respective averaging procedures is from equations (7), (9) or (11). Hence, r_{rstr} is the same value for the three methyl averages, as expected, while the value r_χ varies. The apparent r_{rstr} value for the large majority of methyl interactions is larger than the corresponding distance in the static structure regardless of the type of methyl averaging; the points fall above the continuous line in Figure 6(b), (c) and (d) for the most part. This bias is in contrast to the non-methyl interactions (Fig. 6(a)) for which the points cluster around the continuous line and dynamic averaging effects are nearly equally likely to increase or decrease the NOE interaction strength.

Comparison of the three methyl models finds that the greatest discrepancy occurs for a set of NOE interactions with r_{rstr} near 3·0 Å (the circled areas in Fig. 6). The deviation from the rigid structure distance is reduced in the progression geometrical center → 3-site(r) → 3-site(Ω, r), that is the progression to the more rigorous averaging model. This set comprises primarily pairs involving a non-methyl proton and a methyl group separated by a single torsion angle, i.e. H^α–H^β in Ala, H^β–H^γ in Val, Thr

or Ile, and H^{γ}–H^{δ} in Ile or Leu. For distance restraints referenced to the geometrical center of the methyl protons, r_{rstr} estimated from the measured cross-relaxation rate would give a value which is ~0·5 Å greater than the actual distance from the geometrical center of the three methyl protons (Fig. 6(b)). The error would be ~0·25 Å (Fig. 6(d)) for 3-site(r) averaging in the rigid structure model, while the apparent r_{rstr} deviates insignificantly from the actual distance in the rigid structure determined with angular and radial averaging over the methyl group, the 3-site(Ω, r) model (Fig. 6(c)). Thus, motional averaging of NOE interactions involving a proton and a methyl group across a single torsion angle can be adequately accounted for with 3-site methyl rotation, and other fast timescale fluctuations do not significantly alter cross-relaxation.

(e) *Model of the rigid-molecule*

The assessment of radial averaging from internal motions is made in this study by comparing r_{rig}^{6} to $\langle r^{-3}\rangle^{-2}$, where r_{rig} is from the optimized average simulation structure, r_m (eqn (12)). This choice of structure for the rigid molecule gives a consistent comparison between static and averaged distances from the trajectory and corresponds to an n.m.r. structural solution in which standard geometry and favorable van der Waals and electrostatic contacts are maintained. As stated above, some of the details of motional averaging effects are altered by energy minimization. Distributions of R like those in Figures 3 and 5 were also calculated for r_{rig} obtained from the average simulation structure without energy minimization. R distributions based on the unoptimized average structure are narrower with standard deviations smaller by a factor of 2 for most groups, and less symmetrical about the average value. In the case of the intraresidue interactions, the asymmetry in R is toward values less than 1·0, reinforcing the more extensive angular averaging noted for this group and leading to a more pronounced asymmetry in the distribution of the full averaging parameter Q than observed with the optimized average structure. In the case of the long-range interresidue interactions, the distribution is skewed toward values greater than 1·0 manifesting the apparent decrease in distance expected as a result of dynamic averaging. That is, $\langle R\rangle$ equals $1{\cdot}14 \pm 0{\cdot}45$, compared to $1{\cdot}04 \pm 0{\cdot}62$ listed in Table 1. The narrower R distributions of the unoptimized average structure result in fewer extreme Q values; 6% of the Q values from the interior pairs are less than 0·6 or greater than 1·7, compared to 11% with the optimized structure. (The percentage for the surface interactions is near 13% for both rigid-molecule models.) Energy minimization has the greatest effect on the number of extreme Q values for the long-range interresidue interactions, increasing the percentage from 5% to 20% and 26% for the interior and surface pairs, respectively.

In further consideration of the influence on the model chosen to be the rigid structure, a comparison was made of the distances before (r_u) and after (r_m) energy minimization of the average simulation structure. The ratio $(r_u/r_m)^6$ was evaluated for all proton pairs. Approximately 10% of the ratios were less than 0·6 or greater than 1·7, similar to the overall percentage of extreme Q values.

On the basis of the above discussion, the general features of motional averaging effects described in this work pertain to the rigid protein modeled from the average structure either with or without energy minimization, although minimization does alter the details and aspects related to specific interactions. We note that the effects from motional averaging on n.m.r. structure determination presented in Results, sections (a) and (b), would be reduced if r_{rig} were obtained from the unoptimized average simulation structure. Thus, the disparity between the inter-proton distance of a rigid-molecule and the effective distance from an ensemble of dynamic structures is in part a result of imposing good geometry and favorable non-bonded contacts by energy minimization.

(f) *Protein regions associated with the extreme averaging effects*

It is of interest to examine the nature of the motions of interior proton pairs with extreme values for Q and where they are located in the lysozyme structure. Half of the intraresidue interactions involve side-chain protons of arginine or lysine. Of the remaining extreme Q values from the interior many result from fluctuations in three regions of lysozyme. It should be emphasized that not all interactions in these regions are extensively averaged. Only some of the NOE interactions for a given residue are altered as a result of an internal motion depending on the orientation of the inter-proton vector relative to the direction of motion. Q is therefore not similar in value for spatially close pairs, or for all NOE interactions of a particular proton.

The three regions which exhibit extensive averaging are an eight-residue loop connecting the α-helices αC and αD, the hydrophobic "box" (Poulsen *et al.*, 1980) of non-polar residues at the interface of three α-helices, and the residues near Ser91, the side-chain which contacts the four internal water molecules of lysozyme first described by crystallographic studies (Imoto *et al.*, 1972; Blake *et al.*, 1983). In all three regions the motions are dihedral transitions between rotameric states in either the main-chain (CD-loop) or side-chain (Met105 in the hydrophobic box or Ser91). The CD-loop motion is a localized change in the main-chain conformation which affects only the residues 100 to 107 and has been described in detail (Post *et al.*, 1989). Met105 undergoes transitions in χ_2, and Ser91 undergoes transitions in χ_1 and χ_2. In the case of Ser91, the motion allows alternating hydrogen bond partners of the hydroxyl group with the four internal water molecules. Similarly transient

behavior occurs for other hydrogen bonds in lysozyme (Post *et al.*, 1990*a*).

4. Conclusions

(a) *Motional averaging in three-dimensional structure determination*

The effect of internal motions on ^1H–^1H cross-relaxation in proteins was investigated to elucidate possible consequences of the rigid-molecule assumption on the determination of three-dimensional structure by n.m.r. A description of the nature of picosecond motional averaging effects on all NOE interactions in lysozyme is obtained by using a molecular dynamics trajectory. Longer timescale motions such as large loop displacements or domain motions are not accounted for in this analysis. As such, the full averaging parameter Q calculated from the simulation may not agree quantitatively with experimental values, and longer simulations with all-atom techniques might be useful to improve this approach to study n.m.r. relaxation phenomenon.

NOE cross-relaxation can be diminished or enhanced because of radial averaging, unlike ^{13}C longitudinal relaxation where only angular averaging, and thus a decrease in rate, occurs. We find that the average effect of internal motions on non-methyl interactions is to decrease the cross-relaxation rate by 9% ($\langle Q \rangle = 0{\cdot}91$ for 2778 NOE pairs in lysozyme), with values of Q nearly symmetrically distributed above and below the average. For interpreting the large majority of NOE intensities, the contribution to relaxation from picosecond motions is negligible with respect to estimating interproton distances so that a rigid-molecule model is appropriate. With a qualitative evaluation of distances based on strong, medium and weak intensities, and when n.m.r. distances specify only upper bound restraints (i.e. when only the enhancement of intensities would have negative consequences), picosecond motions result in incorrectly categorizing less than 1% of the interactions (Fig. 2). Even in the case of more quantitative evaluations of distance from σ_{ij}^{obs}, there is less than a 10% error in r calculated assuming a rigid-molecule for 89% of the non-methyl NOE interactions from the interior of the protein. This size of error is insignificant since constraints to the n.m.r. distances are loosely enforced in most procedures for structure determination (Clore & Gronenborn, 1987; Kaptein *et al.*, 1988; Wüthrich, 1989) and is small relative to that which results from experimental error in NOESY intensities or when multiple spin-pair effects are disregarded in large molecules (Olejniczak *et al.*, 1986; Borgias & James, 1988; Post *et al.*, 1990*b*).

Nonetheless, picosecond motions have non-negligible effects on some interactions, even in the protein interior (e.g. Met105 and Ser91), which could lead to difficulties in a conformational search. Averaging in some instances alters cross-relaxation rates by more than fourfold, which would give >20% error in r. For the most extensively averaged interactions, the cross-relaxation rates are usually diminished rather than enhanced (Fig. 6), resulting in smaller NOE intensities than expected for a rigid molecule. For example, if initial rates are measured, the larger decreases found for σ_{ij} would give r values of 2·5 instead of 2·0 Å, or 3·8 instead of 3·0 Å, or 4·3 instead of 3·4 Å. A restraint to an incorrect distance would produce an error in the structure related to the particular proton pair, and could also negatively influence neighboring interactions leading to poor convergence in satisfying the simultaneous set of distance restraints.

Cross-relaxation involving methyl protons is motionally averaged to a greater extent than that involving non-methyl protons; $\langle Q \rangle$ equals 0·72 for 492 interactions even when methyl rotation is fully accounted for in the rigid-structure model. In nearly all cases, the cross-relaxation rate is reduced. Distances estimated from σ_{ij}^{obs} would give a greater than 10% error for 32% of the methyl–methyl and methyl–non-methyl separations.

An internal motion does not equally affect all NOE interactions for a residue or even for a particular proton. If an NOE between a particular proton of residue i and a proton of residue j is enhanced relative to the rigid-molecule value, then it does not follow that all NOEs of that proton of residue i are enhanced, nor that all NOEs between residue i and residue j are enhanced. For example, $Q = 1{\cdot}74$ for the NOE between Leu56 $H^{\beta 1}$ and Trp108 $H^{\eta 2}$, yet the NOE between Leu56 $H^{\beta 1}$ and Trp108 $H^{\zeta 2}$ is essentially unaveraged so that $Q = 0{\cdot}88$. It is the dynamic averaging of the interproton vector which is relevant to cross-relaxation, and averaging depends on the orientation of the particular vector relative to the direction of motion. This situation contrasts with one where a more general dependence on protein mobility applies, such as that involving crystallographic temperature factors and atomic fluctuations. In the case of temperature factors, nearly all atoms in a flexible region of a protein have elevated temperature factors. On the other hand, NOE cross-relaxation is effected only if there is an interproton vector component perpendicular to the direction of motion. To illustrate, rotation about χ_2 in a phenylalanine residue influences cross-relaxation of H^{ε}–H^{ζ}, but not H^{δ}–H^{ε} where the interproton vector is parallel to the direction of motion. Hence incorporation of dynamic information by using individual averaging factors in n.m.r. structure determination procedures must be made on a pair-wise basis, and cannot be specified by regions of the protein, or a particular residue, or even a specific proton.

Consideration of the effects from motional averaging of σ_{ij} would be important to a spectral comparison between experimental NOESY intensities and intensities calculated from a rigid-molecule model. Such a comparison is useful for assessing the quality of an n.m.r. structural solution in a fashion analogous to the R-factor in crystallographic methods. The distribution of Q (Fig. 1) would

increase an *R*-factor determined from the difference between observed NOESY intensities and simulated intensities calculated with a relaxation rate matrix and a static molecule. For short mixing times, the disparity between measured and calculated intensities would relate linearly to Q, and at longer mixing times in a more complicated fashion as indirect relaxation paths become important. Assuming NOE intensities vary linearly with σ_{ij} and Q, the rigid-molecule model of lysozyme would yield for the non-methyl interior pairs an *R*-factor equal to 0·22 due to the picosecond motions described here. The *R*-factor summed over methyl group intensities would equal 0·30.

(b) *Considerations for structure determination*

Analysis by groups of non-methyl proton pairs find that long-range interresidue interactions, both in the interior and on the surface, were most effected by motional averaging by having the highest percentage of extreme values for the full averaging parameter Q (extreme averaging parameter is defined in Table 2). Angular averaging is more pronounced for the intraresidue vectors than for interresidue vectors, leading more often to a reduction in cross-relaxation efficiency for this group of NOEs. Proton pairs in arginine and lysine side-chains exhibited the largest degree of averaging, thus weighting these interactions less heavily in constrained molecular dynamics or distance geometry calculations might be appropriate.

The full averaging parameter Q is model-dependent since the radial averaging parameter R expresses the internally averaged distance relative to the distance in the structural model of the rigid molecule. For NOE interactions involving one or two methyl groups the distribution of Q varies somewhat depending on the procedure taken to account for conformational averaging due to methyl rotation in the rigid-structure model. Dynamically averaged values of σ_{ij} correspond most closely to cross-relaxation rates estimated by including angular and radial averaging over the methyl group in the rigid structure, i.e. 3-site(Ω, r) averaging. Hence, the most accurate procedure to implement NOE-distance restraints involving methyl groups is to average both angular and radial terms ($(Y_m^2/r^{-3})^2$) when calculating the instantaneous distance during the conformational search procedure.

One type of NOE pair that was found to have a systematic error due to dynamic averaging is the case of methyl–non-methyl proton pairs separated by a single torsion angle. Representing the methyl group by the average of r^{-3}, or by the geometrical center of the three methyl proton positions, leads to an estimated distance 0·25 to 0·5 Å too long in almost all cases. However, this systematic error is removed if the instantaneous distance during the conformational search is calculated by 3-site(Ω, r) averaging.

Motional averaging is more pronounced for methyl than non-methyl interactions even if methyl rotation is completely accounted for by the rigid structure model; there are picosecond motions in addition to methyl rotation which give substantial averaging. The nature of the averaging is to reduce cross-relaxation, leading to apparent distances which are too long. The result suggests that it would be appropriate to weight methyl distance restraints less heavily in the conformational search procedure.

The relatively large angular averaging found for geminal proton interactions is important in regard to the choice of a reference distance, used to gauge unknown distances. Since as many as 10% of the geminal proton interactions from the interior and more than 20% from the surface showed significant angular averaging, an unfortunate choice for a reference distance could lead to a systematic error in the n.m.r.-determined distances. A more reliable type of fixed-distance pair would be Tyr or Phe H^{δ}–H^{ε}. $\langle Q \rangle$ for these aromatic fixed-distance pairs is $0{\cdot}91 \pm 0{\cdot}04$, equal to the average over all non-methyl NOE pairs, and there were no extreme Q values.

The nature of the secondary structure markers makes it unlikely that dynamic averaging by internal motions would obscure the identification of a helical or sheet structure. First, the markers involve main-chain–main-chain or main-chain–H^{β} proton pairs, which are not generally mobile; main-chain atoms have smaller atomic fluctuations than side-chain atoms and regions of lowest mobility in a protein correspond to helical and β-sheet elements (Post *et al.*, 1989). Of the total 193 extreme Q values for interior proton pairs, only 26 are main-chain–main-chain interactions, and 25 are main-chain–H^{β} interactions. Second, the markers are short-distance proton pairs so that a measurable NOE would remain even if dynamic averaging reduced the intensity by as much as a factor of 3. Finally, a secondary structure is identified by a pattern of NOEs involving several sequential residues, thus mobility of one or a few NOE pairs should not prevent identification of the secondary structure. Given that there was significant averaging of interactions between protons at the N terminus of three of the four helices in lysozyme, the precise beginning of an α-helix may be difficult to determine by n.m.r.

(c) *Significantly averaged regions in lysozyme*

Most of the NOE interactions with significant averaging in lysozyme involve Ser91 or Met105, or are located in the CD-loop. Ser91 is surrounded by hydrogen bonds to four internal water molecules. It is worthy of note that an increase in mobility of residues associated with internal water molecules has also been reported for interleukin-1β (Clore *et al.*, 1990). Both the CD-loop and Met105 play a role in binding the hexasaccharide substrate of lysozyme. The CD-loop makes direct contact with the substrate and the motions of this loop are altered in the presence of substrate (Post *et al.*, 1986, 1989). Met105 indirectly interacts with the substrate through Trp108. The mobility of Met105 is observed

experimentally (Olejniczak *et al.*, 1981, 1984*a*) and the coupling between 105 and 108 has been studied theoretically (Olejniczak *et al.*, 1984*b*). Further studies are planned to include motional averaging effects in an analysis of experimental two-dimensional NOESY intensities.

The author thanks Martin Karplus and Christopher M. Dobson for many thoughtful and stimulating discussions throughout the course of this work, and for helpful comments on the manuscript. This work was supported by a grant from the Lucille P. Markey Charitable Trust for the development of structural studies at Purdue.

References

Baleja, J. D., Moult, J. & Sykes, B. D. (1990). Distance measurement and structure refinement with NOE data. *J. Magn. Reson.* **87**, 375–384.

Blake, C. C. F., Pulford, W. C. A. & Artymiuk, P. J. (1983). X-ray studies of water in crystals of lysozyme. *J. Mol. Biol.* **167**, 693–723.

Borgias, B. A. & James, T. L. (1988). COMATOSE, a method for constrained refinement of macromolecular structure based on two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.* **79**, 493–512.

Braun, W., Bosch, C., Brown, L. R., Go, N. & Wüthrich, K. (1981). Combined use of proton–proton Overhauser enhancements and a distance geometry algorithm for determination of polypeptide conformations. *Biochim. Biophys. Acta*, **667**, 377–396.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.

Clore, G. M. & Gronenborn, A. M. (1987). Determination of three-dimensional structures of proteins in solution by nuclear magnetic resonance spectroscopy. *Protein Eng.* **1**, 275–288.

Clore, G. M. & Gronenborn, A. M. (1989). Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *CRC Crit. Rev. Biochem. Mol. Biol.* **24**, 479–564.

Clore, G. M., Driscoll, P. C., Wingfield, P. T. & Gronenborn, A. M. (1990). Analysis of the backbone dynamics of interleukin-1β using two-dimensional inverse detected heteronuclear ^{15}N–^{1}H NMR spectroscopy. *Biochemistry*, **27**, 7387–7401.

Denk, W., Baumann, R. & Wagner, G. (1986) Quantitative evaluation of cross-peak intensities by projection of two-dimensional NOE spectra on a linear space spanned by a set of reference resonance lines. *J. Magn. Reson.* **67**, 386–390.

Dobson, C. M., Olejniczak, E. T., Poulsen, F. M. & Ratcliffe, R. G. (1982). Time development of proton nuclear Overhauser effects in proteins. *J. Magn. Reson.* **48**, 97–110.

Gurd, F. R. N. & Rothgeb, T. M. (1979). Motions in proteins. *Advan. Protein Chem.* **33**, 73–165.

Holak, T. A., Scarsdale, J. N. & Prestegard, J. H. (1987). A simple method for quantitative evaluation of cross-peak intensities in two-dimensional NOE spectra. *J. Magn. Reson.* **74**, 546–549.

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972). Vertebrate lysozymes. In *The Enzymes* (Boyer, P. D., ed.), vol. 7, pp. 665–868. Academic Press, New York.

Kaptein, R., Boelens, R., Scheek, R. M. & van Gunsteren, W. F. (1988). Protein structures from NMR. *Biochemistry*, **27**, 5389–5395.

Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972–8979.

Keepers, J. W. & James, T. L. (1984). A theoretical study of distance determinations NMR. Two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.*, **57**, 404–426.

Koning, T. M. G., Boelens, R., van der Marel, G. A., van Boom, J. H. & Kaptein, R. (1991). Structure determination of a DNA octamer in solution by NMR spectroscopy. Effect of fast local motions. *Biochemistry*, **30**, 3787–3797.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

LeMaster, D. M., Kay, L. E., Brünger, A. T. & Prestegard, J. H. (1988). Protein dynamics and distance determination by NOE measurements. *FEBS Letters*, **236**, 71–76.

Levy, R. M., Karplus, M. & McCammon, J. A. (1981*a*). Increase of ^{13}C NMR relaxation times in proteins due to picosecond motional averaging. *J. Amer. Chem. Soc.* **103**, 994–996.

Levy, R. M., Karplus, M. & Wolynes, P. G. (1981*b*). NMR relaxation parameters in molecules with internal motion: exact langevin trajectory results compared with simplified relaxation models. *J. Amer. Chem. Soc.* **103**, 5998–6011.

Levy, R. M., Dobson, C. M. & Karplus, M. (1982). Dipolar NMR relaxation of nonprotonated aromatic carbons in proteins. Structural and dynamical effects. *Biophys. J.* **39**, 107–113.

Lipari, G. & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Amer. Chem. Soc.* **104**, 4546–4559.

London, R. E. (1989). Interpreting protein dynamics with nuclear magnetic resonance relaxation measurements. In *Methods of Enzymology* (Oppenheimer, N. J. & James, T. L., eds), pp. 358–375. Academic Press, San Diego.

Nilges, M., Jabazettl, J., Brunger, A. T. & Holak, T. A. (1991). Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J. Mol. Biol.* **219**, 499–510.

Noggle, J. H. & Schirmer, R. E. (1971). *The Nuclear Overhauser Effect*, Academic Press, New York.

Olejniczak, E. T. (1989). Including methyl rotation in simulations of spin-lattice relaxation experiments. *J. Magn. Reson.* **81**, 392–394.

Olejniczak, E. T., Poulsen, F. M. & Dobson, C. M. (1981). Proton nuclear Overhauser effects and protein dynamics. *J. Amer. Chem. Soc.* **103**, 6574–6580.

Olejniczak, E. T., Poulsen, F. M. & Dobson, C. M. (1984*a*). Distance dependence of proton nuclear Overhauser effects in proteins. *J. Magn. Reson.* **59**, 518–523.

Olejniczak, E. T., Dobson, C. M., Karplus, M. & Levy, R. M. (1984*b*). Motional averaging of proton nuclear Overhauser effects in proteins. Predictions from a molecular dynamics simulation of lysozyme. *J. Amer. Chem. Soc.* **106**, 1923–1930.

Olejniczak, E. T., Gampe, R. & Fesik, S. (1986).

Accounting for spin diffusion in the analysis of 2D NOE data. *J. Magn. Reson.* **67**, 28–41.

Olejniczak, E. T., Gampe, R. T. & Fesik, S. W. (1989). Utility of fitting two-dimensional NOE spectra. *J. Magn. Reson.* **81**, 178–185.

Post, C. B., Brooks, B. R., Karplus, M., Dobson, C. M., Artymiuk, P. J., Cheetham, J. C. & Phillips, D. C. (1986). Molecular dynamics simulations of native and substrate-bound lysozyme. *J. Mol. Biol.* **190**, 455–479.

Post, C. B., Dobson, C. M. & Karplus, M. (1989). A molecular dynamics analysis of protein structural elements. *Proteins*, **5**, 337–354.

Post, C. B., Dobson, C. M. & Karplus, M. (1990*a*). Lysozyme hydrolysis of β-glycosides—A consensus between binding interactions and mechanism. In *Computer Modeling of Carbohydrate Molecules* (French, A. D. & Brady, J. W., eds), pp. 377–388, American Chemical Society, Washington DC.

Post, C. B., Meadows, R. P. & Gorenstein, D. G. (1990*b*). On the evaluation of interproton distances for three-dimensional structure determination by NMR using a relaxation rate matrix analysis. *J. Amer. Chem. Soc.* **112**, 6796–6803.

Poulsen, F. M., Hoch, J. C. & Dobson, C. M. (1980). A structural study of the hydrophobic box region of lysozyme in solution using nuclear Overhauser effects. *Biochemistry*, **19**, 2597–2607.

Smith, L. J., Sutcliffe, M. J., Redfield, C. & Dobson, C. M. (1991). Analysis of ϕ and χ_1 torsion angles for hen lysozyme in solution from ^{1}H NMR spin–spin coupling constants. *Biochemistry*, **30**, 986–996.

Swaminathan, S., Ichiye, T., van Gunsteren, W. & Karplus, M. (1982). Time dependence of atomic fluctuations in proteins: analysis of local and collective motions in bovine pancreatic trypsin inhibitor. *Biochemistry*, **21**, 5230–5142.

Thomas, P. D., Basus, V. J. & James, T. L. (1991). Protein solution structure determination using distances from 2D NOE experiments: effect of approximations on the accuracy of derived structures. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 1237–1241.

Torchia, D. A. (1984). Solid state NMR studies of protein internal dynamics. *Annu. Rev. Biophys. Bioeng.* **13**, 125–144.

Tropp, J. (1980). Dipolar relaxation and the nuclear Overhauser effects in nonrigid molecules: the effect of fluctuating internuclear distance. *J. Chem. Phys.* **72**, 6035–6043.

Wüthrich, K. (1989). Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science*, **243**, 45–50.

Wüthrich, K., Billeter, M. & Braun, W. (1983). Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949–961.

Wüthrich, K., Billeter, M. & Braun, W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. *J. Mol. Biol.* **180**, 715–740.

Edited by P. E. Wright