JOURNAL OF MAGNETIC RESONANCE, Series B 106, 24–31 (1995)

A Relaxation-Matrix Analysis of Distance-Constraint Ranges for NOEs in Proteins at Long Mixing Times

ASIF K. SURI AND RONALD M. LEVY

Department of Chemistry, Rutgers University, Piscataway, New Jersey 08855-0939

Received December 27, 1993; revised April 4, 1994

Long-mixing-time data (τ_m > 200 ms) from NOE spectra have largely been ignored as a source of protein structural information due to the effects of spin diffusion on calculated interproton distances when using the two-spin approximation. An effective approach for incorporating spin-diffusion effects in an average way into refinements is to choose distance bounds based on distributions of distances observed in NOE back calculations on homologous proteins from a protein structure database. We have determined distributions of interproton distances characteristic of newly observed NOE cross peaks for the proteins crambin, PTI, and echistatin at long mixing times. A relaxation-matrix analysis was used to model the effects of spin diffusion. Constraint ranges were constructed from the interproton distance distributions which can be used in standard protein-refinement programs based on the two-spin approximation. Back calculations are also used to analyze constraint ranges typically used for protein structure determinations based on NOE spectra at shorter mixing times. © 1995 Academic Press, Inc.

INTRODUCTION

Two-dimensional nuclear magnetic resonance experiments have been extremely useful in the determination of the three-dimensional structures of proteins in solution (*1–5*). The most important of these is the NOE (nuclear Overhauser effect) experiment, from which distance constraints can be obtained between pairs of protons within a protein (*6–10*). NMR refinement algorithms use these distance constraints to generate three-dimensional protein structures (*11–17*). Improvements in the quality of the structures produced by these algorithms can be obtained by increasing the number and accuracy of such constraints. In principle, the best refinement procedure is to refine the calculated spectrum of a model structure against an experimental spectrum using a relaxation-matrix analysis (*18–21*), but in practice, refinements using empirically derived constraint ranges derived from calibrated NOE intensity–distance relations are often adequate.

With respect to increasing the number of constraints obtained, one avenue that has not been fully utilized is the use of data obtained at long mixing times. The mixing time is the period of time that cross relaxation is allowed to occur. Additional peaks observed at longer mixing times are often not used because, due to the effects of spin diffusion, inaccurate distance constraints would be obtained. We have been studying ways to optimize constraint algorithms when the two-spin approximation is retained (*22, 23*). In a previous paper, we proposed an empirical two-spin equation to estimate interatomic distances from NOE spectra at longer mixing times (*23*) which incorporates spin diffusion into the refinement in an average way. The primary focus of this paper is to determine the interproton distances that are associated with *newly observed* cross peaks at long mixing times (i.e., cross peaks whose intensities fall below threshold values at earlier mixing times) by a full simulation of NOE intensities including spin diffusion, using a known protein structure. The motivation for this study is to develop guidelines on how to use distance information obtained from NOE intensities at longer mixing times. In addition, we examined simulated NOE spectra at shorter mixing times in order to evaluate the spin-diffusion contribution to distance-constraint ranges frequently used in protein structure refinement. All NOE calculations are carried out using a full-relaxation-matrix approach in order to account for multispin effects.

METHODS

In order to simulate 2D NOE (NOESY) spectra, data are generated from the coordinates of a known protein structure. Cross-peak intensities created by the interaction among the nuclear spins of protons within a system are calculated by solving the relaxation-matrix equations. With the knowledge of both the protein structure and the spectra simulated from the structure, one can analyze the NOE intensities in terms of actual interproton distances. In this study, NOESY spectra were generated at several different mixing times (τ_m) for the same set of protein coordinates. The intensities of newly observed cross peaks at subsequent mixing times were recorded and analyzed.

Calculation of the NOE Intensities

The relaxation-matrix calculations and analysis of distance distributions were carried out for three proteins: crambin

1064-1866/95 $6.00

(*24*), pancreatic trypsin inhibitor (*PTI*) (*25*), and echistatin (*26*). Coordinates for both crambin and PTI were obtained from the Brookhaven protein data bank. Calculations were performed using the IMPACT modeling package (*17, 27, 28*). Protons were explicitly added to the molecule and then energy minimized. The NOE intensities were calculated from energy-minimized coordinates. In order to account for spin-diffusion effects, a full-relaxation-matrix approach was employed where the NOE intensity between proton pairs can be described by the Bloch equation (*29*)

$$\frac{d}{d\tau_m}\mathbf{N}(\tau_m) = -\boldsymbol{\sigma}\mathbf{N}(\tau_m), \qquad [1]$$

where $\mathbf{N}(\tau_m)$ is the matrix of cross-peak intensities. Magnetization can be exchanged between protons through cross relaxation. The relaxation matrix σ is constructed with the off-diagonal elements consisting of the cross-relaxation rates. The rate of cross relaxation σ_{ij} between protons i and j can be described by (*7, 30, 31*)

$$\sigma_{ij} = \left(\frac{\mu_0}{4\pi}\right)^2\left(\frac{\gamma^4\hbar^2}{10\langle r_{ij}^3\rangle^2}\right)[6J(\omega) - J(0)], \qquad [2]$$

where γ is the gyromagnetic ratio of a proton, r_{ij} is the internuclear vector between protons, $J(\omega)$ is the spectral-density function, and ω is the spectrometer frequency in rad/s. The spectrometer frequency used was 500 MHz. In the simulations, the protein is considered to be rigid and isotropically tumbling; thus the spectral-density function is

$$J(\omega) = \frac{\tau_c}{1 + \omega^2\tau_c^2}, \qquad [3]$$

where τ_c is the overall tumbling time.

The diagonal elements of the relaxation matrix σ_{ii} are the sum of the spin–lattice plus cross-relaxation rates of spin i. Assuming a constant external relaxation rate E_i for all spins, the direct relaxation rate for proton i can be written as (*30, 31*)

$$\sigma_{ii} = E_i + \sum_{i\neq j}\sigma_{ij}. \qquad [4]$$

Integration of the Bloch equation can then be performed analytically in order to obtain the intensity matrix at mixing time τ_m:

$$\mathbf{N}(\tau_m) = \exp(-\boldsymbol{\sigma}\tau_m)\mathbf{N}(0), \qquad [5a]$$

where

$$\exp(-\boldsymbol{\sigma}\tau_m) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^t. \qquad [5b]$$

$\mathbf{U}$ is the matrix containing the eigenvectors of $\boldsymbol{\sigma}$, and $\boldsymbol{\Lambda}$ is the diagonal matrix containing the exponential of the product of the eigenvalues of $\boldsymbol{\sigma}$ and τ_m. An intensity noise level was employed as in previous studies where cross peaks with intensities equal or above 0.003 were considered observable (*19, 26*).

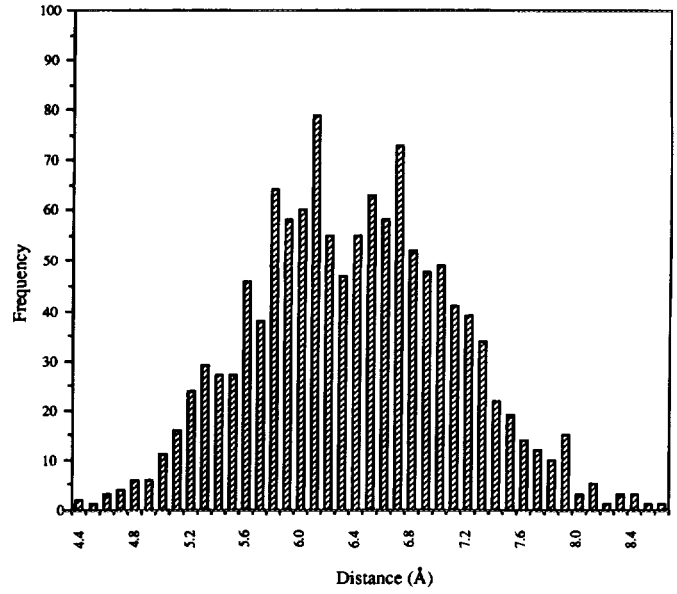

FIG. 1. Distribution of distances associated with new NOE cross peaks seen at 400 ms but not at 200 ms under group 1 conditions. The average distance calculated from the minimized crambin crystal structure was found to be 6.4 Å with a standard deviation of ±0.73 Å.

RESULTS AND DISCUSSION

Four groups of NOESY spectra were generated. Among the groups, the external relaxation rate and the tumbling time of the molecule were varied (Table 1). In order to analyze the additional cross peaks seen at progressively longer mixing times, four sets of spectra were generated for each group at mixing times τ_m of 50, 100, 200, and 400 ms. Cross peaks that appeared in spectra at a particular mixing time but were below the threshold of 0.003 at all earlier mixing times were recorded and analyzed. The sets of *newly* observed cross peaks were then correlated with actual distances. For example, cross peaks appearing above the threshold in a 400 ms spectrum but which are below the noise threshold at 200 ms are included in our data set.

TABLE 1
Groups Used to Generate NOESY Spectra

Group	E_i (s^{-1})	τ_c (ns)
1	0.5	5.0
2	0.5	2.0
3	1.0	5.0
4	1.0	2.0

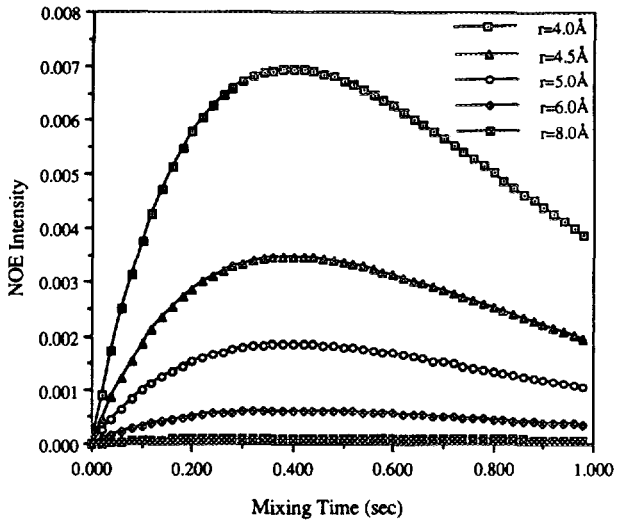

FIG. 2. NOE build-up curves for isolated spin pairs with various interproton distances and τ_c = 3.5 ns. Using a noise level of 0.003, only spin pairs with r = 4.0 and 4.5 Å can be seen.

New Cross Peaks at Long Mixing Times

It was found that the distribution of distances associated with new cross peaks approximates a Gaussian distribution, an example of which is shown in Fig. 1. Within each of the four groups defined in Table 1, the average, maximum, minimum, and standard deviation of the distances between proton pairs corresponding to new cross peaks were recorded from the minimized protein coordinates (Table 2). As one would expect, the average distances associated with sets of cross peaks appearing at longer mixing times were larger than sets seen at earlier mixing times. This can be attributed to the spin-diffusion effect. For isolated spin pairs that are approximately 5 Å apart or greater, the NOE intensity builds up slowly, is very weak, and, in fact, cannot be seen at any mixing time because it is always below the noise level (Fig. 2). At longer mixing times, spins coupled by multispin effects have more time to interact with each other through indirect magnetization transfer. Consequently, in proteins, longer mixing times allow spin pairs that are far apart to transfer magnetization not only directly, but also indirectly by relaying through intermediate spins that are nearby in space. These alternate pathways can augment the intensity of spin pairs separated by larger distances, allowing them to be "seen" at longer mixing times. Figure 3 illustrates this effect in echistatin (*26*), where protons that are 6.4 Å apart can be observed in a spectrum because of relay effects. If the protons were considered as an isolated spin pair, the NOE would not be observable.

Examples of the distributions of distances in crambin that can be observed at a long mixing time (400 ms) for several intensities are shown in Fig. 4. In this case there are 1224 additional cross peaks seen at 400 ms that are not seen at

TABLE 2
New Peaks That Were Not Observed at Previous Mixing Times

Mixing time (ms)	Distance (avg)	Distance (max)	Distance (min)	Distance (SD)	Number of new peaks
		Group 1 (τ_c = 5, E_i = 0.5)			
100	4.745	5.868	3.844	0.394	492
200	5.452	7.288	4.154	0.570	803
400	6.374	8.823	4.352	0.732	1224
		Group 2 (τ_c = 2, E_i = 0.5)			
100	3.854	4.610	3.415	0.308	187
200	4.464	5.699	3.809	0.357	321
400	5.091	7.210	3.984	0.503	608
		Group 3 (τ_c = 5, E_i = 1.0)			
100	4.716	5.832	3.844	0.393	461
200	5.396	7.288	4.154	0.551	719
400	6.178	8.259	4.490	0.704	913
		Group 4 (τ_c = 2, E_i = 1.0)			
100	3.835	4.610	3.415	0.314	175
200	4.392	5.556	3.743	0.354	279
400	4.956	6.241	4.025	0.448	465

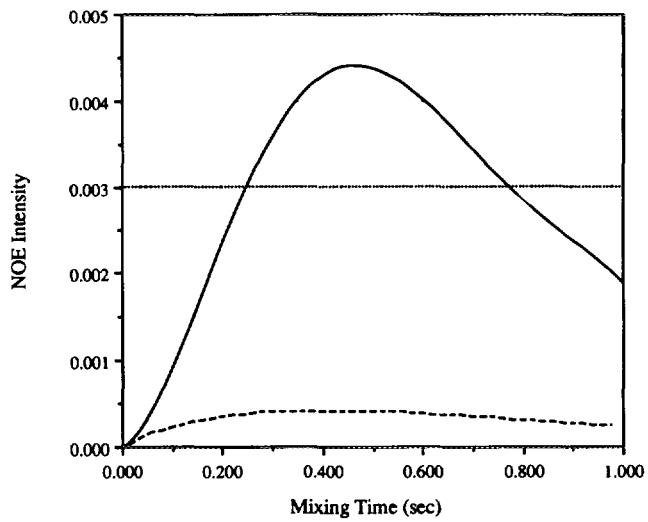

FIG. 3. NOE build-up curve for a proton pair 6.4 Å apart in the protein echistatin (solid line). The dashed line represents the NOE buildup of an isolated spin pair with the same interproton distance. The horizontal line represents the noise level. The isolated spin pair cannot be observed at any mixing time. In contrast, the NOE intensity of the spin pair in echistatin can be seen at longer mixing times due to the effects of spin diffusion from nearby protons which resulted in an increase in NOE intensity.

200 ms. A wide range of distances was seen for the weaker intensities, while a much narrower range was observed for stronger intensities. The solid line shows the NOE intensity as a function of interproton distance for isolated spin pairs. The majority of points are found to be above this line, which implies that the net effect of spin diffusion has been to increase the cross-peak intensities (*22, 23, 32, 33*). The few points that are below the solid line represent cross peaks that have slower build-up curves than the isolated spin pair reference, because the direct transfer of magnetization is "leaked" away due to cross relaxation to other spins. When examining the points which fall below or close to the solid line in Fig. 4, there is a decrease in the number of cross peaks corresponding to spin pairs at shorter distances as the intensity value increases. This is simply due to the fact that only cross peaks not observed at earlier mixing times are listed. Cross peaks corresponding to shorter distances are not present at the stronger intensity values because such distances have already been observed in spectra at earlier mixing times. The examination of the upper portion of the figure shows that there are fewer cross peaks associated with longer distances as the intensity increases. Visual examination of the few proton pairs observed at an intensity of 0.008 showed the existence of a network of spins scattered along the connecting interproton vector that could produce such an augmented intensity. In general, it is apparent from Fig. 4 that: (1) at longer mixing times the average distance between protons corresponding to newly observed NOEs is only weakly dependent on the NOE intensity, and (2) the variance in the distribution of the distances corresponding to a given intensity decreases as the intensity increases.

A possible approach for incorporating spin-diffusion effects in an average way into the refinement process is to choose distance bounds based on the distribution of distances observed at longer mixing times from NOE back calculations on homologous proteins from a protein structure database. By analyzing the average values and standard deviations of distance distributions associated with new cross peaks first observed at longer mixing times for each group (Table 2), constraint ranges for long-mixing-time data can be established and used in the refinement of proteins which are of similar size and shape to the reference molecule. Examining the spectra we generated for crambin, we find that, for new cross peaks seen at 400 ms under group 1 conditions, the average interproton distance is 6.4 Å (Table 2). Choosing twice the standard deviation of the distance distribution to be the constraint range, this value is calculated to be ~1.5 Å for the newly observed crambin cross peaks at 400 ms. Thus, additional cross peaks seen at 400 ms correspond to interproton distances between 4.9 and 7.9 Å. Based on our back calculations for crambin as well as the two other proteins studied (echistatin and PTI, see below), using these distance bounds is sufficient to account for essentially all of the spin-diffusion effects for small to medium-size proteins. For crambin, there are several hundred additional peaks seen at 400 ms which were not observed at 200 ms. Although only a fraction of these additional intensities will in practice be observed because of effects such as spectral overlap and motion, by accounting for spin-diffusion effects using back calculations to derive constraint ranges as we propose, one can incorporate the longer-mixing-time data in a simple and effective way into the refinement.

Influence of E_i and τ_c

It is apparent from the relaxation-matrix analysis that the value of the overall tumbling time affects the distance distributions associated with new cross peaks to a larger extent than the external relaxation rate. In groups 1 and 3, the overall tumbling time was kept constant at 5.0 ns while the external relaxation rates were set to 0.5 and 1.0 s^{-1} respectively. Between the two groups, the largest difference in the average distance was in the 400 ms set; in group 1 the average distance was 6.37 Å and in group 3 it was 6.18 Å. This is only a 0.19 Å deviation (Table 3a). Similarly, comparing group 2 and group 4 where the tumbling time was fixed, τ_c = 2.0 ns while the external relaxation rate was varied; the average distances were 5.09 and 4.96 Å, respectively, which corresponds to a deviation of only 0.13 Å.

In contrast, varying the tumbling time while keeping the external relaxation rate constant had a greater effect on the magnitude of the NOE intensities among spins (Table 3b). Looking at groups 1 and 2 where the tumbling times were 5.0 and 2.0 ns, respectively, and the external relaxation rate was 0.5 s^{-1}, at 400 ms the difference in the average distance was 1.28 Å, and even at 200 ms the average distances differed

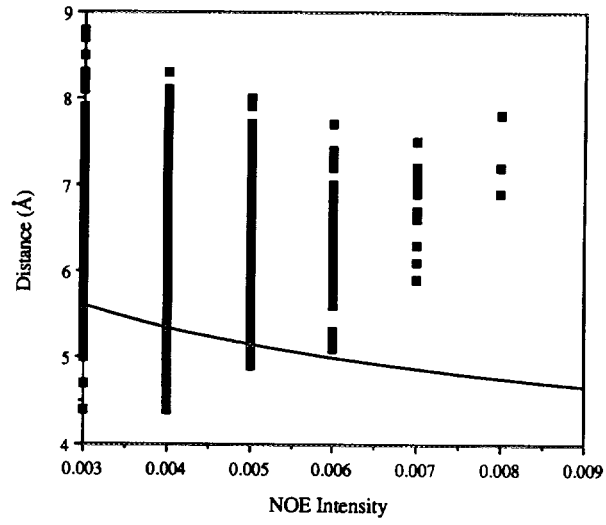

FIG. 4. Plot of distance vs NOE intensity for all additional cross peaks seen at 400 ms but not at 200 ms in group 1. There are over 1200 additional cross peaks plotted in the figure. The solid line shows the relationship between distance and intensity for isolated spin pairs.

TABLE 3a
Differences in Average Distances

Mixing time (ms)	τ_c = 5.0 (ns)	τ_c = 2.0 (ns)
	E_i varied from 0.5 to 1.0 s^{-1}	
100	0.029	0.019
200	0.056	0.072
400	0.196	0.135

by about 1 Å. A similar difference was found when examining groups 3 and 4, where the external relaxation rate was held constant at 1.0 s^{-1}. At 400 ms, the average distances differed by 1.22 Å between the slow-tumbling (τ_c = 5.0 ns) and fast-tumbling (τ_c = 2.0 ns) simulations, and at 200 ms they differed by 1 Å.

Changes in the overall tumbling time make a larger contribution to the spin-diffusion effect than variations in the external relaxation rate. However, more cross peaks can be seen at later mixing times with a slower external relaxation rate. Comparing groups 1 and 3 at a mixing time of 400 ms with a constant tumbling time of 5.0 ns, when E_i = 0.5 s^{-1} there are 1224 cross peaks with intensities greater than the threshold value 0.003. When E_i = 1.0 s^{-1}, this number is reduced to 913 cross peaks.

In order to check the sensitivity of these calculations to the details of the crambin structure, back calculations using the protein BPTI (*25*) were performed (Table 4). Cross peaks were generated at 200 and 400 ms for all four groups. In all cases, the average and standard deviation of distances associated with newly observed cross peaks at 400 ms closely resembled those observed with crambin. For example, with E_i = 1.0 s^{-1} and τ_c = 5.0 ns, new cross peaks corresponded to spin pairs separated by an average distance of 6.19 Å in BPTI as compared to 6.17 Å in crambin. In addition, the standard deviation of the average cross-peak distance was ±0.71 Å in BPTI, while in crambin it was ±0.70 Å. The results suggest that this analysis is independent of the detailed protein geometry for those of roughly similar size and shape.

Analysis of Short-Mixing-Time Data

Most refinement procedures make use of NOESY intensities obtained at mixing times of no more than 200 ms (*1, 4, 26, 34–36*). Constraints derived from these spectra are often divided into three categories: strong, medium, and weak. In general, a spectrum is obtained at approximately 50 ms, and the resultant cross peaks are divided into the strong and medium categories. Another spectrum at a longer mixing time, up to 200 ms, is then obtained; any newly observed peaks not previously seen are designated as weak constraints. A typical assignment of distances for strong, medium, and weak constraints is 2.0–2.7, 2.0–3.5, and 2.0–5.0 Å, respectively (*4, 37*).

In order to evaluate the validity of the distance ranges associated with different classes of cross-peak intensities discussed above, we completed a series of simulations to analyze the relationship between NOE intensities observed at short mixing times and the actual distances between the corresponding proton pairs. Crambin NOESY spectra were simulated at 50, 80, 100, and 200 ms for each of the four groups of parameters listed in Table 1. Spectra at 50, 80, and 100 ms were alternately used in determining the strong and medium constraints, while newly observed cross peaks in the 200 ms spectra were considered weak. In order to distinguish between strong and medium constraints for a given spectra, the two-spin equation was used to deduce the intensity value that should correspond to a distance of 2.7 Å (Eq. [5]), where the reference NOE intensity used, $N_{ref}(\tau_m)$, was the average of all glycine methylene-proton-pair intensities. The reference distance used, r_{ref}, was 1.78 Å.

$$N_{2.7\,\text{Å}}(\tau_m) = \frac{N_{ref}(\tau_m) r_{ref}^6}{(2.7\,\text{Å})^6}. \qquad [6]$$

Intensity values that were greater or equal to the cutoff intensity computed above were considered strong, while the rest were considered medium intensities. After classification of all cross peaks, each category was analyzed in terms of actual distances.

It is apparent that with a tumbling time of 2.0 ns, spectra at 50 ms yielded distances that fitted reasonably well into the predicted distance bounds. For those cross peaks considered to be strong (Table 5), when using 50 ms data, almost all corresponding distances were found to be below 2.7 Å. Of the few cross peaks with distances at or above 2.7 Å, none

TABLE 3b
Differences in Average Distances

Mixing time (ms)	E_i = 0.5 (s^{-1})	E_i = 1.0 (s^{-1})
	τ_c varied from 5.0 to 2.0 ns	
100	0.891	0.881
200	0.988	1.004
400	1.283	1.222

TABLE 4
New Peaks That Were Not Observed at Previous Mixing Times (BPTI, τ_m = 400 ms)

τ_c (ns)	E_i (s^{-1})	Distance (avg)	Distance (max)	Distance (min)	Distance (SD)
5.0	0.5	6.351	8.513	4.387	0.730
2.0	0.5	5.113	7.136	4.039	0.515
5.0	1.0	6.194	8.485	4.245	0.713
2.0	1.0	4.994	6.864	3.909	0.480

TABLE 5a
Strong Cross Peaks

Mixing time (ms)	Distance (avg)	Distance (max)	Distance (min)	Distance (SD)	Number of peaks
		Group 1 (τ_c = 5, E_i = 0.5)			
50	2.535	3.663	2.055	0.253	112
80	2.606	3.796	2.055	0.307	133
100	2.655	3.796	2.055	0.345	148
		Group 2 (τ_c = 2 E_i = 0.5)			
50	2.457	2.737	2.055	0.168	93
80	2.476	2.946	2.055	0.181	99
100	2.482	3.027	2.055	0.188	100
		Group 3 (τ_c = 5, E_i = 1.0)			
50	2.519	3.663	2.055	0.251	106
80	2.605	3.796	2.055	0.308	132
100	2.658	3.796	2.055	0.342	151
		Group 4 (τ_c = 2, E_i = 1.0)			
50	2.472	2.753	2.055	0.175	98
80	2.474	2.946	2.055	0.179	98
100	2.482	3.027	2.055	0.188	100

exceeded this value by more than 0.05 Å. Similarly, distance bounds generally used for medium and weak constraints seem to be adequate to bracket the true interproton distances corresponding to cross peaks that fall into the medium and weak categories, as shown in Tables 6 and 7. In contrast, using an 80 or 100 ms spectrum to determine the strong and medium cross peaks proved to be inadequate for having the corresponding distances fall into the correct range for each category. For those cross peaks categorized as having medium intensity at 80 ms, more than 40% are found to be above 3.5 Å. When a 100 ms spectrum is used, more than 50% of the cross peaks that are considered medium intensity are above 3.5 Å.

At a tumbling time of 5.0 ns, the majority of NOEs are found to correspond to interproton distances which are outside the expected distance bounds in both the medium and the weak categories for both external relaxation rates, regardless of the mixing time used to determine strong and medium constraints (Tables 5–7). As many as 57% of the spin pairs that first appear at 200 ms, classified as weak cross peaks, after using a 50 ms spectrum to determine strong and medium cross peaks, are found to correspond to interproton distances over 5.0 Å, and 11% are found above 6.0 Å. As was true for the long-mixing-time data, a longer tumbling time enhances the spin-diffusion effect and results in a large number of peaks observed outside the anticipated distance bounds. At shorter tumbling times, the standard distance bounds seem to be satisfactory for qualitatively predicting proton distances. It should be noted that for all medium and weak constraints, the minimum distances observed were well above 2.0 Å. This suggests that it is possible to use a lower bound that is greater than the sum of the van der Waals radii of two protons. Another point of interest was that the minimum distance observed for medium cross peaks in crambin, regardless of τ_c and E_i values, was always 2.73 Å. This was the distance between Phe 13 Hz and Arg 17 HD1. The NOE buildup of this proton pair was reduced because two other protons on the phenylalanine ring are close to Phe 13 Hz, resulting in an intensity weaker than expected for this distance.

CONCLUSION

Often, additional NOE peaks first observed at longer mixing times are not used for protein structure determination since it is believed that, because of spin diffusion, distances cannot be accurately estimated from intensities. As a result, potentially useful structural information is discarded. Additional distance constraints based on NOE data collected at longer mixing times can be obtained as long as one makes a judicious choice of the target distances and ranges to be used. This study offers guidance in choosing target distances and ranges appropriate for distances derived from NOEs observed at longer mixing times. It is interesting to note that

TABLE 5b
Strong Cross Peaks: Percentage and Number of Cross Peaks Above Cutoff

Mixing time (ms)	>2.7 Å	>3.0 Å	>3.3 Å
	Group 1 (τ_c = 5, E_i = 0.5)		
50	21.4 (24)	4.5 (5)	1.8 (2)
80	33.8 (45)	7.5 (10)	4.5 (6)
100	40.5 (60)	10.8 (16)	6.8 (10)
	Group 2 (τ_c = 2, E_i = 0.5)		
50	5.4 (5)	0.0 (0)	0.0 (0)
80	11.1 (11)	0.0 (0)	0.0 (0)
100	12.0 (12)	1.0 (1)	0.0 (0)
	Group 3 (τ_c = 5, E_i = 1.0)		
50	17.0 (18)	4.7 (5)	1.9 (2)
80	33.3 (44)	7.6 (10)	4.6 (6)
100	41.7 (63)	10.6 (16)	6.6 (10)
	Group 4 (τ_c = 2, E_i = 1.0)		
50	10.2 (10)	0.0 (0)	0.0 (0)
80	10.2 (10)	0.0 (0)	0.0 (0)
100	12.0 (12)	1.0 (1)	0.0 (0)

TABLE 6a
Medium Cross Peaks

Mixing time (ms)	Distance (avg)	Distance (max)	Distance (min)	Distance (SD)	Number of peaks
		Group 1 (τ_c = 5, E_i = 0.5)			
50	3.729	4.796	2.728	0.515	454
80	4.106	5.722	2.728	0.610	721
100	4.306	5.868	2.728	0.647	910
		Group 2 (τ_c = 2, E_i = 0.5)			
50	3.142	4.106	2.728	0.289	169
80	3.389	4.469	2.728	0.386	280
100	3.530	4.610	2.728	0.458	349
		Group 3 (τ_c = 5, E_i = 1.0)			
50	3.708	4.796	2.728	0.521	452
80	4.072	5.699	2.728	0.605	685
100	4.281	5.832	2.728	0.634	868
		Group 4 (τ_c = 2, E_i = 1.0)			
50	3.134	3.834	2.728	0.267	158
80	3.368	4.434	2.728	0.375	271
100	3.506	4.610	2.278	0.456	331

TABLE 7a
Weak Cross Peaks

Mixing time (ms)	Distance (avg)	Distance (max)	Distance (min)	Distance (SD)	Number of new peaks
		Group 1 (τ_c = 5, E_i = 0.5)			
50	5.184	7.288	3.844	0.615	1295
80	5.346	7.288	4.064	0.577	1007
100	5.452	7.288	4.154	0.570	803
		Group 2 (τ_c = 2, E_i = 0.5)			
50	4.239	5.699	3.415	0.450	508
80	4.396	5.699	3.685	0.372	391
100	4.464	5.699	3.809	0.357	321
		Group 3 (τ_c = 5, E_i = 1.0)			
50	5.130	7.288	3.844	0.596	1180
80	5.281	7.288	3.984	0.564	921
100	5.396	7.288	4.154	0.551	719
		Group 4 (τ_c = 2, E_i = 1.0)			
50	4.178	5.556	3.415	0.434	454
80	4.337	5.556	3.620	0.364	341
100	4.392	5.556	3.743	0.354	279

TABLE 6b
Medium Cross Peaks: Percentage and Number of Cross Peaks Above Cutoff

Mixing time (ms)	>3.5 Å	>4.0 Å	>4.5 Å
	Group 1 (τ_c = 5, E_i = 0.5)		
50	68.1 (309)	32.4 (147)	6.8 (31)
80	82.5 (595)	59.8 (431)	27.3 (197)
100	87.5 (796)	69.8 (635)	41.0 (373)
	Group 2 (τ_c = 2, E_i = 0.5)		
50	10.7 (18)	0.6 (1)	0.0 (0)
80	43.9 (123)	7.1 (20)	0.0 (0)
100	55.3 (193)	16.9 (59)	1.2 (4)
	Group 3 (τ_c = 5, E_i = 1.0)		
50	66.6 (301)	31.2 (141)	6.4 (29)
80	81.5 (558)	57.7 (395)	25.7 (176)
100	87.2 (757)	68.7 (596)	38.8 (337)
	Group 4 (τ_c = 2, E_i = 1.0)		
50	8.9 (14)	0.0 (0)	0.0 (0)
80	41.7 (113)	6.3 (17)	0.0 (0)
100	52.9 (175)	16.0 (53)	0.9 (3)

TABLE 7b
Weak Cross Peaks: Percentage and Number of Cross Peaks Above Cutoff

Mixing time (ms)	>5.0 Å	>5.5 Å	>6.0 Å
	Group 1 (τ_c = 5, E_i = 0.5)		
50	56.8 (736)	29.7 (384)	10.7 (139)
80	69.3 (698)	37.5 (378)	13.8 (139)
100	75.7 (608)	45.3 (364)	17.3 (139)
	Group 2 (τ_c = 2, E_i = 0.5)		
50	3.7 (19)	0.6 (3)	0.0 (0)
80	4.9 (19)	0.8 (3)	0.0 (0)
100	5.9 (19)	0.9 (3)	0.0 (0)
	Group 3 (τ_c = 5, E_i = 1.0)		
50	54.1 (638)	26.8 (316)	8.3 (98)
80	66.1 (609)	33.8 (311)	10.6 (98)
100	73.6 (529)	41.9 (301)	13.6 (98)
	Group 4 (τ_c = 2, E_i = 1.0)		
50	2.2 (10)	0.2 (1)	0.0 (0)
80	2.9 (10)	0.3 (1)	0.0 (0)
100	3.6 (10)	0.4 (1)	0.0 (0)

there were more new cross peaks observed above the threshold intensity seen between 200 and 400 ms than any of the other designated sets (Table 2). Thus, there is the potential for gaining more structural information by also using structural information derived from NOEs first observed at mixing times somewhat longer than those currently being used. Additional data should at least be consistent with models developed from earlier mixing times and, in favorable cases, may produce data which introduce constraints between portions of a molecule previously unconstrained. In addition, many labs use the sum of van der Waals radii as the lower bounds for all distance ranges regardless of mixing time or intensity. Clearly, protons which produce cross peaks that first appear at longer mixing times are much further apart.

Distance-constraint ranges can be parameterized from NOE back calculations carried out on a database of protein structures. A reasonable approach for incorporating spin-diffusion effects in an average way into refinements which use the two-spin approximation is to choose distance bounds based on the distribution of distances observed as a function of intensity obtained from NOE back calculations on homologous proteins selected from a database. Additional structural information at longer mixing times may be used to better define the tertiary structure of proteins when only small numbers of constraints are obtained at earlier mixing times. Based on our back calculations, in the analysis of shorter-mixing-time spectra (50 ms) normally used for the determination of protein structures, the distance bounds presently used are adequate when the protein tumbling time is short. As the tumbling time increases, adjustments in the distance bounds assigned to cross peaks may have to be made as an increasing fraction of the constraints fall outside these bounds. In another paper we have reported on the use of an empirical two-spin approximation in order to more accurately estimate interproton distances at longer mixing times (*23*). It is hoped that these studies will stimulate more extensive use of longer-mixing-time data in protein structure determination.

ACKNOWLEDGMENTS

This work was supported by NIH Grant GM-30580 and by allocations of computer time at The National Cancer Institute and The Pittsburgh Supercomputer Center. We thank Professor Jean Baum for discussions concerning spin-diffusion effects on protein structures.

REFERENCES

1. K. Wüthrich, "NMR of Proteins and Nucleic Acids," Wiley, New York, 1986.
2. K. Wüthrich, *Science* **243,** 45 (1989).
3. G. M. Clore and A. M. Gronenborn, *Crit. Rev. Biochem.* **24,** 479 (1989).
4. A. Bax, *Annu. Rev. Biochem.* **58,** 223 (1989).
5. P. E. Wright, *TIBS* **14,** 255 (1989).
6. I. Solomon, *Phys. Rev.* **99,** 559 (1955).
7. F. A. L. Anet and A. J. R. Bourn, *J. Am. Chem. Soc.* **87,** 5250 (1965).
8. J. Noggle and R. Schirmer, "The Nucleart Overhauser Effect—Chemical Applications," Academic Press, New York, 1971.
9. P. Balaram, A. A. Bothner-By, and J. Dadok, *J. Am. Chem. Soc.* **94,** 4015 (1972).
10. D. Newhaus and M. Williamson, "The Nuclear Overhauser Effect in Structural and Conformational Analysis," VCH, New York, 1989.
11. I. D. Kuntz, G. Crippen, P. A. Kollman, and D. Kimmelman, *J. Mol. Biol.* **106,** 983 (1976).
12. T. F. Havel and K. Wüthrich, *Bull. Math. Biol.* **46,** 673 (1984).
13. M. Levitt, *J. Mol. Biol.* **170,** 723 (1983).
14. R. Kaptein, E. R. P. Zuiderweg, R. M. Scheck, and W. F. Van Gunsteren, *J. Mol. Biol.* **182,** 179 (1985).
15. G. M. Clore, A. M. Gronenborn, A. T. Brünger, and M. Karplus, *J. Mol. Biol.* **186,** 435 (1985).
16. W. Braun and N. Go, *J. Mol. Biol.* **186,** 611 (1985).
17. D. A. Bassolino, F. Hirata, D. B. Kitchen, D. Kominos, A. Pardi, and R. M. Levy, *Int. J. Supercomput. Appl.* **2,** 41 (1988).
18. R. Boelens, T. M. G. Konig, G. A. van der Marel, J. H. van Boom, and R. Kaptein, *J. Magn. Reson.* **82,** 290 (1989).
19. B. A. Borgias and T. L. James, *J. Magn. Reson.* **79,** 493 (1988).
20. P. Yip and D. Case, *J. Magn. Reson.* **83,** 643 (1989).
21. J. D. Baleja, R. T. Pon, and B. D. Sykes, *Biochemistry* **29,** 4828 (1990).
22. D. Kominos, A. K. Suri, D. B. Kitchen, D. Bassolino, and R. M. Levy, *J. Magn. Reson.* **97,** 398 (1992).
23. A. K. Suri and R. M. Levy, *J. Magn. Reson. B* **101,** 320 (1993).
24. W. A. Hendrickson and M. M. Teeter, *Nature* **290,** 107 (1981).
25. A. Wlodawer, J. Deisenhofer, and R. Huber, *J. Mol. Biol.* **193,** 145 (1987).
26. Y. Chen, A. K. Suri, D. Kominos, A. M. Naylor, S. M. Pitzenberger, V. M. Garsky, R. M. Levy, and J. Baum, *J. Biomol. NMR* **4,** 307 (1994).
27. R. M. Levy, D. A. Bassolino, D. B. Kitchen, and A. Pardi, *Biochemistry* **28,** 9361 (1989).
28. D. B. Kitchen, F. Hirata, J. D. Westbrook, R. M. Levy, D. Ko f ke, and M. Yarmush, *J. Comp. Chem.* **11,** 1169 (1990).
29. A. Abragam, "The Principles of Nuclear Magnetism," Oxford Univ. Press, Oxford, 1978.
30. A. Kalk and H. J. C. Berendsen, *J. Magn. Reson.* **24,** 343 (1976).
31. G. M. Clore and A. M. Gronenborn, *J. Magn. Reson.* **61,** 158 (1985).
32. C. B. Post, R. P. Meadows, and D. G. Gorenstein, *J. Am. Chem. Soc.* **112,** 6796 (1990).
33. B. A. Borgias and T. L. James, *J. Magn. Reson.* **87,** 475 (1990).
34. M. Nilges, G. M. Clore, and A. M. Gronenborn, *FEBS* **239,** 129 (1988).
35. M. P. Williamson, T. F. Havel, and K. Wüthrich, *J. Mol. Biol.* **182,** 295 (1985).
36. A. D. Kline, W. Braun, and K. Wüthrich, *J. Mol. Biol.* **189,** 377 (1986).
37. G. M. Clore and A. M. Gronenborn, *Crit. Rev. Biochem.* **1,** 275 (1987).