Journal of Biomolecular NMR, 9 (1997) 79–93
ESCOM

J-Bio NMR 411

Assessment of protein solution versus crystal structure determination using spin-diffusion-suppressed NOE and heteronuclear relaxation data

David M. LeMaster

Department of Biochemistry and National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, 420 Henry Mall, Madison, WI 53706, U.S.A.

Received 24 July 1996
Accepted 28 October 1996

Keywords: Protein deuteration; Spin diffusion; Heteronuclear relaxation; Structural analysis

Summary

A spin-diffusion-suppressed NOE buildup series has been measured for *E. coli* thioredoxin. The extensive ^{13}C and ^{15}N relaxation data previously reported for this protein allow for direct interpretation of dynamical contributions to the ^{1}H-^{1}H cross-relaxation rates for a large proportion of the NOE cross peaks. Estimates of the average accuracy for these derived NOE distances are bounded by 4% and 10%, based on a comparison to the corresponding X-ray distances. An independent fluctuation model is proposed for prediction of the dynamical corrections to ^{1}H-^{1}H cross-relaxation rates, based solely on experimental structural and heteronuclear relaxation data. This analysis is aided by the demonstration that heteronuclear order parameters greater than 0.6 depend only on the variance of the H-X bond orientation, independent of the motional model in either one- or two-dimensional diffusion (i.e., $1 - S^2 = 3/4 \sin^2 2\theta_\sigma$). The combination of spin-diffusion-suppressed NOE data and analysis of dynamical corrections to ^{1}H-^{1}H cross-relaxation rates based on heteronuclear relaxation data has allowed for a detailed interpretation of various discrepancies between the reported solution and crystal structures.

Introduction

Interproton distance constraints obtained from homonuclear cross relaxation provide the primary experimental data used to derive the structure of proteins in solution via NMR. In principle, the $1/r^6$ dependence of the cross-relaxation rate offers the potential for highly accurate distance estimates. However, in slowly tumbling macromolecules multistep magnetization transfers negate the simple pairwise distance dependence of the NOE buildup rates. A number of pulse sequences have been proposed which suppress the spin-diffusion effects for a subset of the ^{1}H-^{1}H pairs (Olejniczak et al., 1986; Massefski and Redfield, 1988; Fejzo et al., 1991,1992; Konrat et al., 1991; Boulat et al., 1992; Boulat and Bodenhausen, 1993; Hoogstraten et al., 1993,1995b). Data utilizing the network-edited NOESY technique have recently been incorporated into the solution structural analysis of turkey ovomucoid third domain (Hoogstraten et al., 1995a). Spin-diffusion effects can also potentially be estimated via relaxation matrix analysis (Boelens et al., 1988; Borgias and James, 1990; Post et al., 1990). Unfortunately, as the contribution of indirect transfer to cross-peak intensity becomes dominant in more slowly tumbling larger proteins, the structural determinacy of the relaxation matrix approach becomes less robust (Clore and Gronenborn, 1989).

Intramolecular dynamics provides a second potential major source of uncertainty in quantitative distance estimation from NOE data. Molecular dynamics simulations have been used to assess the effects of internal mobility on the observed cross-relaxation rates (Olejniczak et al., 1984; LeMaster et al., 1988; Bruschweiler et al., 1992; Post, 1992; Abseher et al., 1995). However, these calculations are limited to comparatively short simulation times for macromolecules. Furthermore, the derived dynamical properties for a given molecular system can depend significantly on the choice of force field used (Brunne et al., 1993; Schmidt et al., 1993). Heteronuclear relaxation measurements offer an independent assessment of the

Abbreviations: NOE, nuclear Overhauser enhancement; RFD, random fractional deuteration.

local dynamics. However, earlier studies have generally been limited to main-chain and a modest subset of side-chain positions.

In an attempt to circumvent the problem of NOE distance quantitation, the majority of solution structural analyses have assigned relatively loose ad hoc distance limits for each of the observed ^{1}H-^{1}H NOE interactions. Given a reasonably high density of NOE constraints, the precision among a family of model structures is commonly found to be 0.5 Å or better. Unfortunately, the more relevant question of experimental accuracy has remained problematic (Zhao and Jardetzky, 1994). A direct comparison between parallel crystal and solution structure determinations cannot unambiguously distinguish between experimental inaccuracies and actual differences between solution and crystal conformations. As solution structures are used increasingly for the interpretation of detailed structural features in enzyme mechanism studies and drug binding design, the assessment of the reliability of predicted local structural elements will gain enhanced significance.

The 108-residue protein *E. coli* thioredoxin provides an excellent experimental system in which to assess the effects of spin diffusion and molecular dynamics on the experimental determination of NOE distance constraints and on the resultant potential distortion of the derived solution structure. We have reported the X-ray structure of the oxidized form (the R-factor is 16.5% at 1.68 Å) (Katti et al., 1990). In addition, the solution structure of this protein (Jeng et al., 1994) is among the best reported to date in terms of the density of NOE distance constraints and the level of satisfaction of these constraints. Indeed, in contrast to most reported comparisons between analogous solution and crystal structures, the crystal structure lies within the estimated precision of the solution structure family. The rmsd for the backbone heavy atoms of residues 4–107 is 0.28 Å between the crystal structure and the mean solution structure (Jeng et al., 1994).

Recently, we (LeMaster and Kushlan, 1996) reported the relaxation analysis for 413 H-C and H-N bond vectors of *E. coli* thioredoxin, thus providing an extensive mapping of the spatial distribution of the internal mobility. The present manuscript presents a quantitative demonstration of the practical suppression of spin diffusion via the use of protein samples randomly deuterated at the carbon bound sites at a level of 75%. Deuteration has long been used to enhance the spectral quality of the remaining protein ^{1}H resonances (Crespi et al., 1968; Markley et al., 1968). Although deuteration as a means of suppressing spin diffusion in protein has been discussed for some time (LeMaster and Richards, 1988; Torchia et al., 1988; LeMaster, 1990a; Tsang et al., 1990), its practical use in the derivation of quantitative distance constraints has not been extensively explored.

The assessment of the accuracy of interproton distances derived from spin-diffusion-suppressed NOE data provides a means to interpret the detailed differences between solution structures obtained via the conventional qualitative distance bounding algorithms and the corresponding crystal structures. Specifically, the population of cross peaks for which both the solution structure and the random fractionally deuterated (RFD) NOE-derived distances are in closest agreement will characterize discrepancies which arise from either dynamical distortion of the cross-relaxation rates or else from real differences in the average conformation in the solution versus the crystal structure. Comparison with the extensive heteronuclear relaxation data (LeMaster and Kushlan, 1996) can serve to distinguish these alternatives. Conversely, those NOE distance estimates obtained from the 75% RFD sample which most closely agree with the corresponding X-ray structural distances define as a population those solution–crystal structure discrepancies which arise from errors in the NOE distance constraints. Needless to say, superimposed upon this division is noise arising from the inaccuracies in the quantitative NOE distance estimates as well as in the structure determination processes. However, the observed distinct spatial clustering of the NOE discrepancies according to this division argues strongly for its physical relevance.

Materials and Methods

[75% U-^{2}H] *E. coli* thioredoxin was prepared as previously described (LeMaster and Richards, 1988). 2D NOESY mix times of 25, 50, 75, 100, 150, 225, 375 and 600 ms were collected at 14.1 T on a 3 mM protein sample at pH 3.3 in 50 mM sodium [$^{2}H_{1}$]formate buffer at 25 °C. These solution conditions approximate those used to crystallize the protein from 2,4-methylpentanediol for the X-ray structure determination (Katti et al., 1990). The H_2O signal was suppressed by presaturation during the 1.5 s relaxation delay. Minimal attenuation of the amide cross peaks is anticipated as saturation transfer from the H_2O resonance is inefficient at this low pH (Li and Montelione, 1993). Processing and peak integrations were carried out using FELIX (Biosym Technologies Inc., San Diego, CA, U.S.A.) software with baseplane corrections using FLATT (Güntert and Wüthrich, 1992).

Main-chain (LeMaster and Richards, 1988) and side-chain (Dyson et al., 1989) ^{1}H resonance assignments have been previously reported. NOE cross-peak assignments were accepted only if both chemical shifts lay within 7 Hz of the reference value for each resonance (10 Hz for scalar coupled cross peaks). The ^{2}H-^{1}H differential isotope shifts for the two and three bond couplings (Hansen, 1988) presented no practical complication.

Linear least-square fits to the NOE buildup intensities were determined. The number of data points used as well

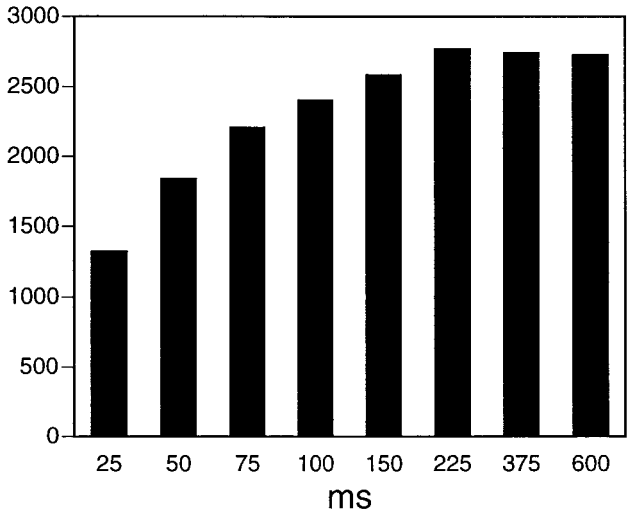

Fig. 1. Number of unique NOE cross peaks of [75% U-^{2}H] *E. coli* thioredoxin as a function of mix time. Cross peaks were selected on the 225 ms data set and these integration regions were then applied to the other spectra testing for cross peaks having volumes >4σ baseplane noise. The constancy in the number of observed cross peaks at longer mix times argues against a significant number of noise peaks being included in the analysis. Visual inspection of the longer mix times indicated that only a modest number of additional cross peaks appear in these spectra.

as the selection of linear models (i.e. ax versus ax + b) was based on maximization of the goodness-of-fit parameter Q (Press et al., 1989). Cross peaks were rejected if the optimal Q-value was below 0.05. The noise level estimation σ was based on the integration of resonance-free baseplane regions. Additionally linear fits were required to include at least two intensity values in excess of 3σ. For cross peaks predicting distances less than 2.7 Å, a linear fit to the 25 ms mix time intensity value was used.

Hydrogen atoms were introduced onto the crystallographic structural model using QUANTA (Molecular Simulation Inc.). Hydrogen positions were then optimized with heavy-atom positions constrained via energy minimization using CHARMM (Brooks et al., 1983). Comparison to the reported solution structures (Jeng et al., 1994) is based on the corresponding interproton distances averaged for 20 structures.

Results and Discussion

Distance calibration and accuracy estimation

Figure 1 summarizes the distribution of the number of NOESY cross peaks for 75% RFD *E. coli* thioredoxin as a function of mix time. The 225 ms data set was used to select 2774 unique cross peaks ($\omega_2 > \omega_1$ for ω_2 downfield of H_2O and $\omega_2 < \omega_1$ for ω_2 upfield of H_2O). The contour level used corresponded to minimal cross-peak volumes of 3–4 times the rmsd of the baseplane noise. These integration regions were then used to determine the number of cross peak volumes >4σ for each of the other mix times. Internal α-helical sequential αN connectivities were used to calibrate the ^{1}H-^{1}H distances. Distance estimates were adjusted to reflect the average 25% ^{1}H enrichment at the carbon bound sites. For this calibration set, the average deviation (mean absolute deviation) between the RFD NOE-derived distances and the corresponding interproton distances of the X-ray structure is 4.2%. In comparisons between the X-ray- and NMR-derived distances, the average deviation is used throughout this analysis as a robust estimator of dispersion for populations in which the tails are anticipated to be significantly non-Gaussian (Press et al., 1989). For all other cross peaks, apparent distances were derived assuming a $1/r^6$ dependence of the NOE buildup rates as predicted for a rigid isolated spin pair. When this distance calibration is applied to the β-strand sequential αN connectivities, the RFD NOE-derived distances are, on average, 0.23 Å (11%) too long. The occurrence and magnitude of the apparent lengthening of the NOE distance for these sequential connectivities was earlier predicted in molecular dynamics calculations (LeMaster et al., 1988).

Calibration using the α-helical sequential αN connectivities insures that the distance estimates are essentially model independent as both NMR and crystallographic analyses can readily identify these regions of secondary structure. More problematic is the assessment of the dispersion of the distance estimates as a means of assigning a measure of accuracy. Here a conservative approach is taken by using the corresponding interproton distances from the X-ray structure to estimate an upper bound for the uncertainty in the distance estimates obtained from the RFD NOE measurements. The advantage of this calibration is that obviously the X-ray data are immune to the effects of spin diffusion, and while internal dynamics affect the crystal structure determination, they are unlikely to manifest themselves in a structural distortion similar to that which would occur in the corresponding solution NMR analysis. On the other hand, such an estimate of dispersion also includes the real structural differences between crystal and solution as well as the errors that occur within the crystallographic determination. Hence such an estimate necessarily represents an upper bound to the inaccuracy of the NMR-derived distances.

For the purpose of assessing limitations in distance estimation, only well-resolved reliably integrated resonances were of interest. Cross peaks exhibiting appreciable overlap with other cross peaks or with spectral artifacts were excluded based on visual inspection. Furthermore, cross peaks were excluded if the crystal structure predicted potentially degenerate cross peaks accounting for over 15% of the observed intensity. As noted in the Materials and Methods section, only cross peaks exhibiting buildup rates which were linear to within the baseplane noise level were analyzed further. Only cross peaks involving methylene resonances of unambiguous chirality

TABLE 1
AVERAGE DEVIATION FOR NONVICINAL INTERPROTON DISTANCES OBTAINED FROM X-RAY COORDINATES VERSUS NOE BUILDUP RATES OF [75% U-^{2}H] *E. coli* THIOREDOXIN

Pair type	All (%)	S^2 [a] > 0.75	0.75 > S^2 > 0.5	S^2 < 0.5
NH-NH	9.0	8.7	9.6	21.7[b]
NH-CH	10.1	10.8	7.9	16.2[b]
NH-CH_2	12.3	14.6	10.6	12.8
NH-CH_3	10.9	10.1	9.9	13.9
CH-CH	7.7	9.2	5.7[b]	–
CH-CH_2	17.6	13.2[b]	17.4[b]	23.5
CH-CH_3	9.7	9.4	10.4	7.8
CH_2-CH_2	15.8	–	17.8	4.5[b]
CH_2-CH_3	10.1	–	11.3	6.5[b]
CH_3-CH_3	11.1	11.0	11.0[b]	11.7[b]
IS + I_3S	10.0	9.7	9.9	13.2
I_2S	13.2	14.4	11.9	14.9

[a] Order parameters for both attached heteroatoms observed (Stone et al., 1993; LeMaster and Kushlan, 1996). Categories are based on the smaller of the two S^2 values.
[b] Fewer than five pairs.

(LeMaster, 1987; Kushlan and LeMaster, 1993; Chandrasekhar et al., 1994) which the X-ray structure predicts to be spatially proximal to the NOE exchange partner are included. Due to efficient geminal cross relaxation, sidechain amides were excluded as were carbon bound protons vicinal to methyl groups due to factors discussed below. Finally, geminal and vicinal proton pairs were excluded due to potential zero quantum cross-peak distortions of the integrations.

For the remaining 1402 cross peaks, Table 1 presents the average deviation between the RFD NOE-derived and X-ray interproton distances as a function of heteroatom and multiplicity type. Table 1 also lists the corresponding average deviations for the cases in which the relaxation order parameters of both directly attached heteroatoms are known. Methyl groups are listed according to the rotation axis order parameter S^2_{axis}. Methyl distances are calculated according to the internal correlation function of a rapid threefold jump model (Tropp, 1980).

For the large majority of cross peaks involving IS and I_3S spin systems, an average deviation of 10% is obtained. A test of whether this dispersion might largely reflect real differences with the X-ray structure takes advantage of the fact that there are two nonequivalent monomers in the asymmetric unit of the crystal (Katti et al., 1990). The alignment of 90% of all heavy atoms yields an rmsd of 0.42 Å between the A (used for these comparisons) and B molecules. As the average interproton distance for the data of Table 1 is 3.6 Å, the average discrepancy between the two molecules in the asymmetric unit is consistent with the average deviation of 10% obtained between the RFD NOE-derived distances and the crystal structure. Hence it can be concluded that the average accuracy of the RFD NOE-derived distances is bounded by the 4.2% estimate obtained from the α-helical sequential connectivities and the 10% estimate derived from the data of Table 1. For the cross peaks involving IS and I_3S spins, only for the category of smallest order parameters does it appear that ps–ns internal dynamics, on average, give rise to an increased observable discrepancy between the RFD NOE-derived and X-ray interproton distances. Cross peaks involving methylene resonances exhibit a 3–4% higher uncertainty with a less apparent dependence on the order parameter. As discussed below, for both of these cases dynamical corrections to the apparent distances reduce the discrepancies with the X-ray interproton distances to ~10% as well.

For this 75% deuterated protein sample, the isolated spin pair approximation assumes that the effects of the vast set of isotopomers are operationally eliminated by a combination of statistical dilution and the more rapid relaxation of the ^{1}H-enriched isotopomers. In order to illustrate the suppression of most spin-diffusion effects, the NOE cross peaks between NH and CH_3 resonances were considered. Figure 2 plots the average deviations between the pairwise RFD NOE-derived distances and the crystallographic distances as a function of the crystallographic distance in 0.5 Å bins. The average deviations are approximately 10% throughout the range of 2.75–6.25 Å. This distance range corresponds to a 140-fold variation in the initial slope of the NOE buildup curves. In order to test for a systematic over/underestimation of distances, for the data of each 0.5 Å bin the average X-ray and average NOE distances were determined and these differences were also plotted in Fig. 2. Although for distances

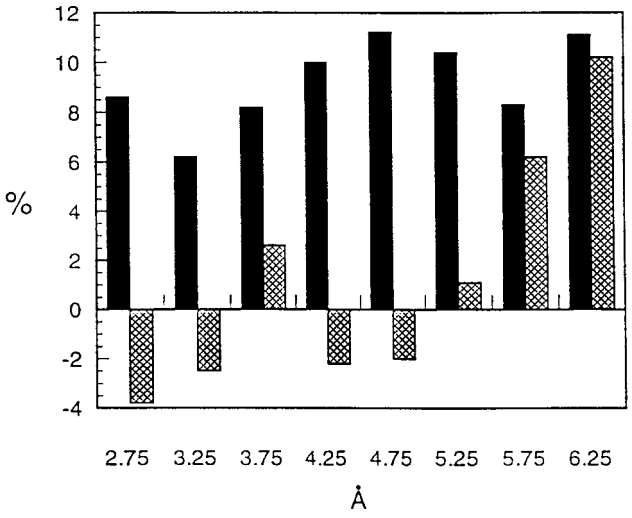

Fig. 2. Average deviation between the RFD NOE-derived distances compared to those obtained from the crystal structure of *E. coli* thioredoxin for the NH-CH_3 cross peaks as a function of X-ray distance in 0.5 Å bins (solid bars). In order to test for the systematic over/underestimation of NOE-derived distances, for the data of each 0.5 Å bin the average X-ray and average NOE distances were determined and the difference $[(d_{X\text{-}ray}-d_{NMR})/d_{X\text{-}ray}]$ was plotted (hatched bars).

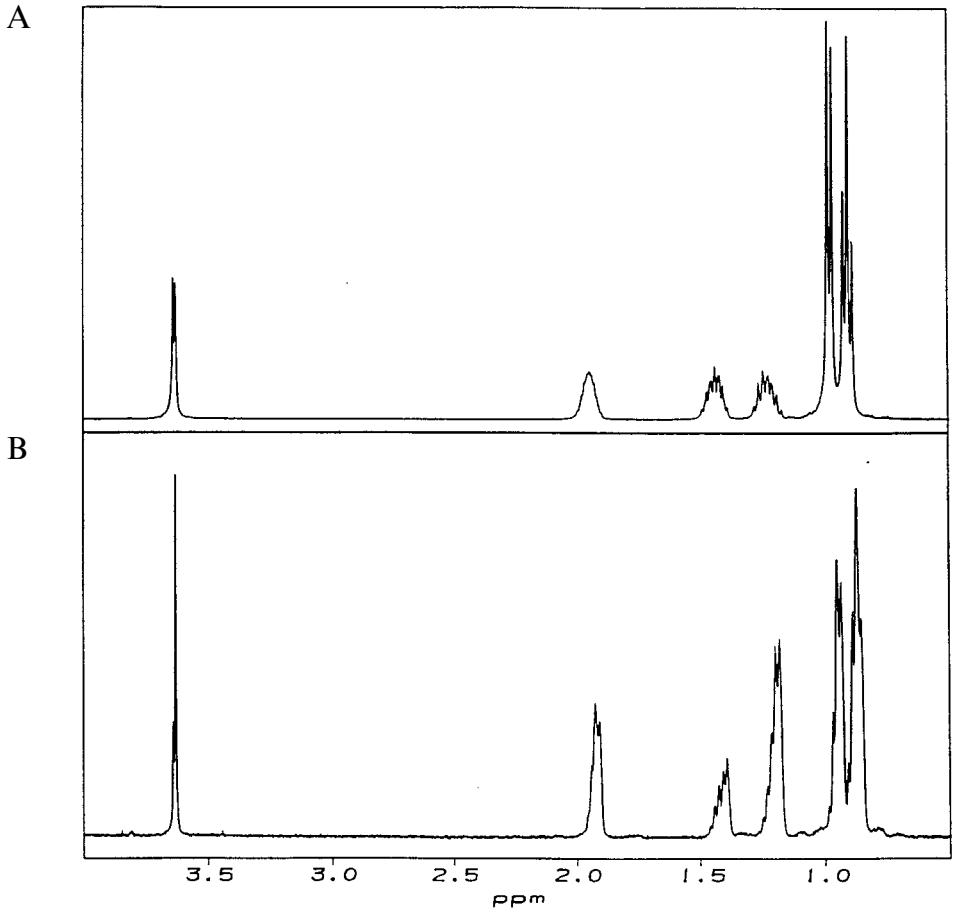

Fig. 3. ^{1}H spectrum of isoleucine derived from bacterial growth in an 80% deuterated medium. Panel A illustrates the natural abundance spectrum. For the deuterated sample of panel B, the H^{α}, $H^{\gamma 13}$ and methyl resonances have ^{1}H enrichment levels of approximately 25%. H^{β} (1.9 ppm) and $H^{\gamma 12}$ (1.2 ppm) exhibit isotopic selectivities of 1.5 and 2.3, respectively.

above 5.5 Å the average deviations are largely dominated by a systematic underestimation of the RFD NOE distance, the practical effect on the observed average deviation is modest over the entire distance range. The systematic underestimation at longer distances must exceed ~10% in order to dominate the other contributions, presumably mainly those due to deviations from the X-ray structure.

Residual spin diffusion is not the only cause for the systematic underestimation of NOE distances for weaker cross peaks observed at longer mix times. As the calculated initial slope of the NOE buildup must exceed a given threshold to be considered above statistical noise, those weak cross peaks for which noise variations enhance the apparent slope will be preferentially accepted (Liu et al., 1995). These data illustrate the practical suppression of spin diffusion over the entire distance range of Fig. 2. In particular, it should be noted that, despite the fourfold dilution of the methyl ^{1}H spins, the operational elimination of competing relaxation pathways allows for NH-CH_3 cross peaks to be reliably quantitated out to 6.0–6.5 Å.

Substantial discussion (Baleja et al., 1990; Summers et al., 1990; Thomas et al., 1991) has been directed toward defining a parameter analogous to the crystallographic R-factor for the purpose of monitoring structural refinement against the experimental NOE constraints. In marked contrast to crystallography for which the interpretation of the statistical significance of diffraction intensities is comparatively straightforward, spin diffusion effectively precludes a model-independent assessment of the statistical reliability of a structural constraint as a function of cross-peak intensity. For the random fractionally deuterated samples considered herein, the operational demonstration of a distance-independent fractional uncertainty provides a solution to this problem.

Another potential contribution to the observed uncertainty in the RFD NOE-derived distances is nonuniformity in the level of partial deuteration. Particularly for the hydrogen isotopes it is well known that biological incorporation exhibits isotopic selectivity (Galimov, 1985). Bulk protein from *E. coli* grown in the same growth mixture used for the RFD thioredoxin sample was acid hydrolyzed and separated into the individual amino acids according to published procedures (LeMaster and Richards, 1982). NMR analysis demonstrated that, for the large majority of sites, the observed level of deuterium

TABLE 2
NOE-PREDICTED DISTANCES FOR SEQUENTIAL H^N-H^N CONNECTIVITIES <2.4 Å IN THE X-RAY COORDINATES OF *E. coli* THIOREDOXIN

Residues	X-ray	NMR	NOE from [75%-^{2}H]	Structural position
Tyr^{49}–Gln^{50}	2.3	2.8	2.2	End of α_2-helix
Lys^{52}–Leu^{53}	2.0	2.1	1.8	Start of β_3-strand
Asp^{61}–Gln^{62}	2.2	2.6	2.1	End of highly distorted α_3-helix
Tyr^{70}–Gly^{71}	2.2	2.7	1.8	End of 3_{10}-helix
Arg^{73}–Gly^{74}	2.0	2.6	2.1	Start of active Pro^{76} hairpin turn
Ala^{87}–Ala^{88}	2.1	2.5	2.0	β-Bulge in β_5-strand
Asn^{106}–Leu^{107}	2.2	2.6	2.1	End of α_4-helix
Leu^{107}–Ala^{108}	2.2	2.7	2.2	Terminal dipeptide

Interproton distances based on the X-ray (Katti et al., 1990) and NMR (Jeng et al., 1994) structure determinations.

incorporation deviated from the average by less than 10%. Several other sites differ from the average by less than 20%. Due to the $1/r^6$ dependence of the NOE cross-relaxation rate, a 20% deviation from the assumed average deuterium level would give rise to a systematic bias in the derived distance of only 3%.

The two cases of substantial differential isotopic selectivity both occur at positions vicinal to methyl groups as illustrated in Fig. 3 for partially deuterated isoleucine. For this amino acid, the β-position (1.9 ppm) exhibits a ^{1}H enrichment 1.5-fold higher than that observed at the H^α, $H^{\gamma 13}$ and methyl positions. Such an enrichment was anticipated based on the deuterium kinetic isotope effect of 1.39 reported for the reaction of 3-hydroxy-3-methyl-2-oxobutanoate with NADPH during the biosynthesis of isoleucine (Chunduru et al., 1989). The same ketol acid reductoisomerase serves to introduce hydride reducing equivalents to the H^β of valine and the H^γ of leucine as well. More unexpected was the large isotopic selectivity of 2.3 at the $H^{\gamma 12}$ site (1.2 ppm). This hydrogen is introduced by protonation from the *re* face of the enamine intermediate formed during the conversion of L-threonine into 2-oxobutanoate catalyzed by the L-threonine dehydratase (Crout et al., 1980). Particularly when combined with ^{13}C enrichment (LeMaster and Kushlan, 1996), this differential isotopic enrichment provides a facile means of assigning the chirality at the γ-position of isoleucine, which is otherwise comparatively problematic to determine. Although these differential enrichment levels can be directly incorporated into the distance estimates derived from the RFD NOE data, as discussed below there are additional considerations which limit the utility of cross-relaxation rates estimated for positions vicinal to methyl groups.

To assess the relative reliability of a more limited NOE series, the precisions of the distance estimates based on the individual mix times were compared to those deduced from the entire set. At short mix times, the estimates are comparatively precise since only cross peaks that were linear from the initial mix times were included in the analysis of Table 1. On the other hand, at longer mix times the precision diverges markedly (6.3% at 150 ms, 8.5% at 225 ms, 13.5% at 375 ms) as the cross peaks exhibiting the most rapid buildup rates deviate substantially from linear. As a smaller number of cross peaks are observed at short mix times, an effective compromise is obtained with the 50 and 225 ms data sets. The cross peaks with volumes $>4\sigma$ at 50 ms provide distance estimates with a precision of 3.4% relative to those of the full mix time set. The additional cross peaks which become observable in the 225 ms data set yield distances with a precision of 3.8%. Comparable performance is anticipated for a pair of mix times scaled inversely to the correlation time of the protein under study ($\tau_c \sim 7$ ns for *E. coli* thioredoxin).

This analysis of spin diffusion in random fractionally deuterated proteins should be largely applicable to [U-^{13}C,^{15}N]-labeled samples as well. The median ^{13}C R_1 relaxation rate for *E. coli* thioredoxin is 1.2 s^{-1} with less than 3% of the nuclei relaxing more than twice as fast (LeMaster and Kushlan, 1996). At 225 ms the additional decay during the mix time due to the ^{1}H-^{13}C dipole interaction would yield an average distance overestimate of ~4%. As modest systematic distance underestimation occurs for the weaker peaks monitored at this mix time (Fig. 2), a partial cancellation of errors is anticipated. Unfortunately, ^{13}C-induced relaxation during the evolution and acquisition periods will further reduce the sensitivity of these experiments, particularly for cross peaks between carbon bound protons as the statistical dilution already yields an ~3-fold reduction in sensitivity for the 75% RFD samples (LeMaster and Richards, 1988). Amide ^{1}H observed heteronuclear NOESY experiments will be more favorable as the reduced ^{1}H-^{1}H dipole relaxation of the partially deuterated samples yields markedly reduced decay during the indirect ^{1}H and ^{13}C evolution periods utilizing multiquantum coherence (Griffey and Redfield, 1987; Seip et al., 1992; Grzesiek and Bax, 1995). As has been previously discussed (LeMaster, 1990a,b,1994), selective enrichment patterns creating a multiple isolated spin pair distribution will overcome the statistical dilution of intensity while retaining the benefit of the reduced spin diffusion.

It should be stressed that the accuracy estimates obtained herein are intrinsically statistical in character and their implementation into solution structural refinement need reflect that fact. They are not readily compatible with the rigid distance boundaries commonly used in structure determination. In that regard, it should be emphasized that, even in the operational absence of spin diffusion, this study clearly demonstrates the observation of NOE interactions extending significantly beyond the upper distance bounds commonly assumed in the standard protocols.

Assessing structural accuracy using spin-diffusion-suppressed NOE distances

The practical benefit of accurate NOE-derived distance constraints can be illustrated in the conformational analysis of tight turns which are notoriously difficult to characterize using local distance constraints based on qualitative NOE intensity assessments. In the crystal structure, there are nine sequential H^N-H^N distances in *E. coli* thioredoxin less than 2.4 Å. These constitute the majority of residues found in the 'bridge' region between the α-helical and extended conformational regions of the (ϕ,ψ) dihedral angle plot. RFD NOE-derived distances have been obtained for eight of these. The Val[16] H^N - Leu[17] H^N cross peak is severely overlapped with the geminal side-chain amide cross peak of Asn[83]. Table 2 lists these sequential connectivities with their interproton distances as assessed from the reported crystal and solution structures and from the RFD NOE analysis. The agreement between the crystal structure distances and the RFD NOE-derived distances is quite good, while in all but one case the solution structure exhibits the longer interproton distances more characteristic of an α-helical conformation. It has been argued that the conformational assignment of such turn sequences should in general await the overall tertiary structural analysis (Wüthrich, 1986). However, for the solution structural analysis of *E. coli* thioredoxin which incorporated an average of 25 NOE constraints per residue, the longer range constraints were largely unsuccessful in providing definition to these local conformational elements.

The effective suppression of spin diffusion for the majority of NOE cross peaks provides the opportunity to assess the presence of residual efficient spin-diffusion pathways. Two specific cases became apparent during the analysis of the RFD NOE data. Particularly in α-helical regions, the NOE distances for H^N_n,H^α_{n+1} and H^α_n,H^N_{n+2} were systematically foreshortened. As the amide hydrogen positions are nearly fully enriched with ^1H and the sequential H^N-H^N distances are relatively short, this foreshortening almost surely results from spin diffusion via the intervening amide proton. However, given the number of additional structural constraints normally present along the main chain in these regions, such a systematic distance error would be unlikely to substantially bias the resultant structure.

A more interesting case occurs for cross peaks involving protons vicinal to methyl groups which yield substantially disparate distance estimates more than 2 times as often as for the remainder of the NOE cross-peak set. These distance discrepancies were larger than could be accounted for by the differential isotope selectivity discussed above, and both under- and overestimation of distance frequently occurred. Efficient cross relaxation between these protons and the adjacent methyl group will tend to yield equivalent distance estimates to a common distal NOE exchange partner. The threefold higher proton density at the methyl position renders this group dominant in the interactions. As illustrated in Fig. 4, the apparent angle defined by the interproton vectors to the methyl group will tend to decrease if the vicinal proton is farther from the distal proton, while the apparent angle will tend to increase if the vicinal proton is closer to that exchange partner.

An analysis was carried out for the cases in which the distance between the vicinal proton and the distal exchange partner differed between the X-ray and solution structures by more than 3 times the rmsd between the backbone atoms of the solution and crystal structure (i.e. 0.84 Å). When the X-ray structure predicts a value less than 60° for the angle between the vectors to the methyl position, the angles obtained from the solution structure were, on average, 18° larger. Conversely, for those cases in which the X-ray structure predicted angles greater than 75°, the angles from the solution structure were, on average, 26° smaller. These average angular variations correspond to a change in the distance between the vicinal proton and the distal exchange partner of approximately +0.8 Å and –1.1 Å, respectively. This correlation indicates that vicinal NOE equilibration can result in appreciable distortion in the structure surrounding methyl groups.

Analysis of sites in E. coli thioredoxin exhibiting significant differences between solution and crystal structures

For 10% of the over 1400 NOE cross peaks selected for the quantitative distance analysis in Table 1, the corresponding interproton distances of the crystal and solution structures differ by more than 3 times the main-chain rmsd between those structures. After removing the NOE cross peaks involving protons vicinal to methyl groups as discussed above, these distance discrepancies are illustrated in Fig. 5. Panel A indicates that for 73% of those distance discrepancies the RFD NOE distance estimates more closely agree with the crystal structure. In only 27% of the cases are the RFD NOE distances closer to those of the solution structure (panel B). Nearly half of the proton pairs highlighted in panel A are in the up-

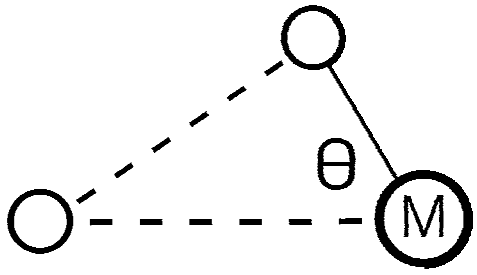

Fig. 4. Apparent NOE-derived distances involving protons vicinal to methyl positions. Efficient cross relaxation between the vicinal protons and the methyl protons causes both groups to appear equidistant from a distal NOE exchange partner. The angle defining the equidistant orientation is a function of the distance to the distal exchange partner ranging from 60° at 2.5 Å to 75° at 5 Å.

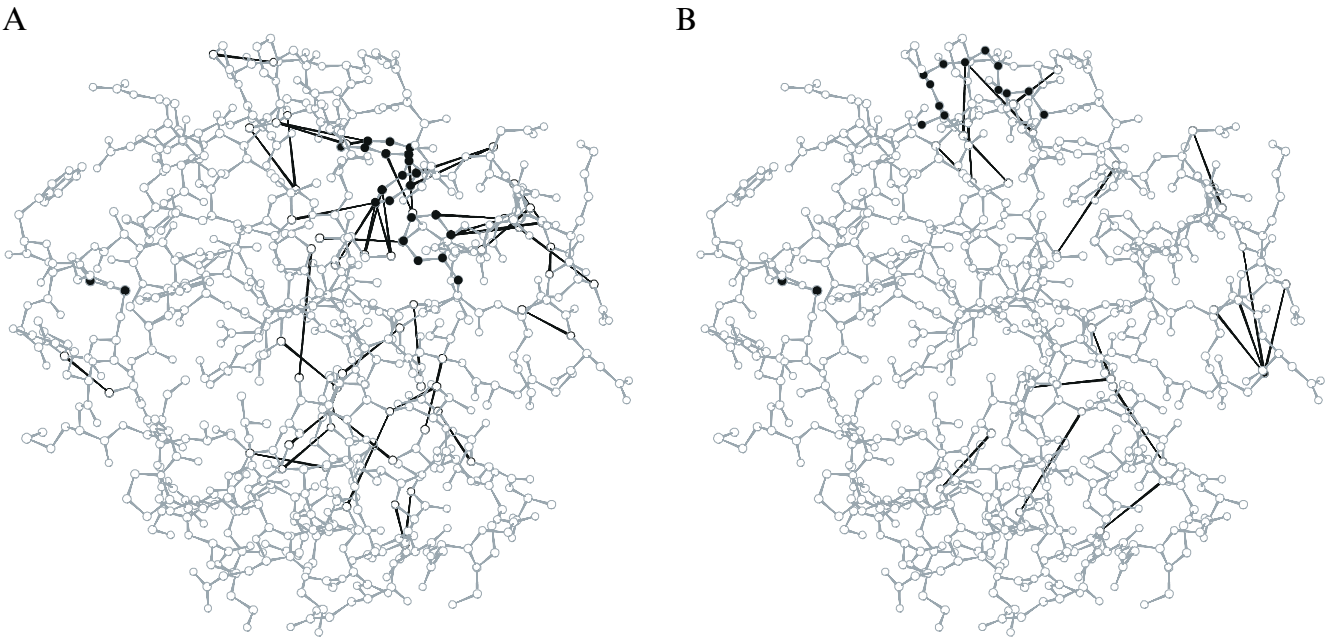

Fig. 5. RFD NOE connectivities (from Table 1) for which the solution (Jeng et al., 1994) and crystal (Katti et al., 1990) structures predict differences in interproton distances greater than 0.84 Å. Cross peaks involving side-chain protons vicinal to methyl groups are excluded. Panel A shows the connectivities for which the RFD NOE-derived distances agree most closely with the X-ray structure. The cluster of aromatic rings from Phe^{12}, Tyr^{70} and Phe^{81} are marked in black. Panel B shows the connectivities for which the agreement for the RFD NOE-derived distances is closest to those of the solution structure. In this case, Gly^{21} N is marked in black at the top of the figure, while the backbone atoms extending from Asn^{63} to Ala^{67} are marked in black at the left side of the figure. In both panels, the active site cystine sulfurs are marked in black near the bottom of the figure.

per left corner concentrated around the aromatic rings of Phe^{12}, Tyr^{70} and Phe^{81} which form the core of the major hydrophobic cluster.

The region surrounding Phe^{12}, Tyr^{70} and Phe^{81} is shown in Fig. 6. For every distance discrepancy >0.84 Å involving the aromatic rings, the RFD NOE-derived distances agree most closely with the crystal structure. There are 11 additional observed RFD NOE cross peaks to these aromatic rings for which the discrepancy between solution and crystal structure is >0.42 Å. For all but one of these cross peaks, the RFD NOE-derived distances agree most closely with the crystal structure. These aromatic rings have an average S^2 of 0.88, consistent with a minimal dynamical correction to the derived internuclear distances. Hence the structural discrepancy appears to arise from errors in the treatment of spin-diffusionally perturbed distance constraints in the solution structural analysis in a region having a particularly high density of NOE constraints.

A clustering of distance discrepancies emanate from Gly^{21} H^N at the top of Fig. 5B. For each of these, the interproton vector is approximately perpendicular to the peptide plane containing Gly^{21} H^N. The H^α's of Gly^{21} were excluded from the data in Fig. 5 due to the degeneracy of their chemical shifts. However, using the structure to predict the closest of this geminal pair as well as including an additional NOE involving Gly^{21} H^N which was excluded

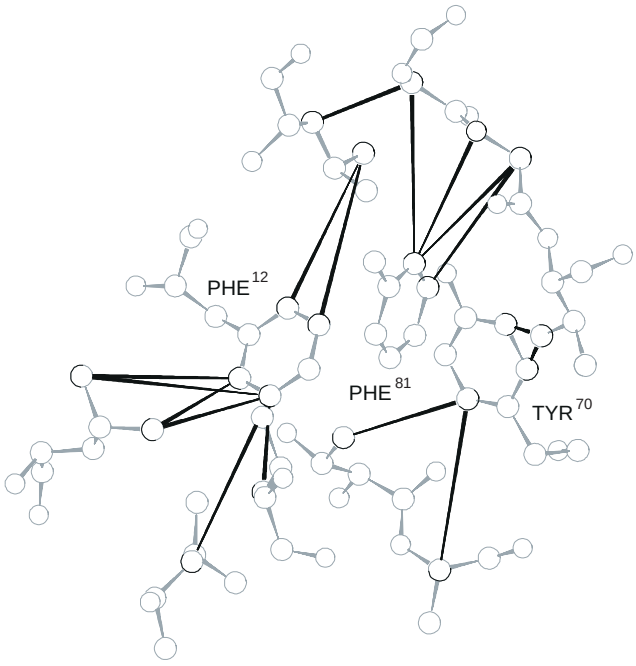

Fig. 6. NOE connectivities for which the solution (Jeng et al., 1994) and crystal (Katti et al., 1990) structures predict differences in interproton distances greater than 0.84 Å for the aromatic rings of Phe^{12}, Tyr^{70} and Phe^{81}. In all cases, the RFD NOE-derived distances agree most closely with the X-ray distances.

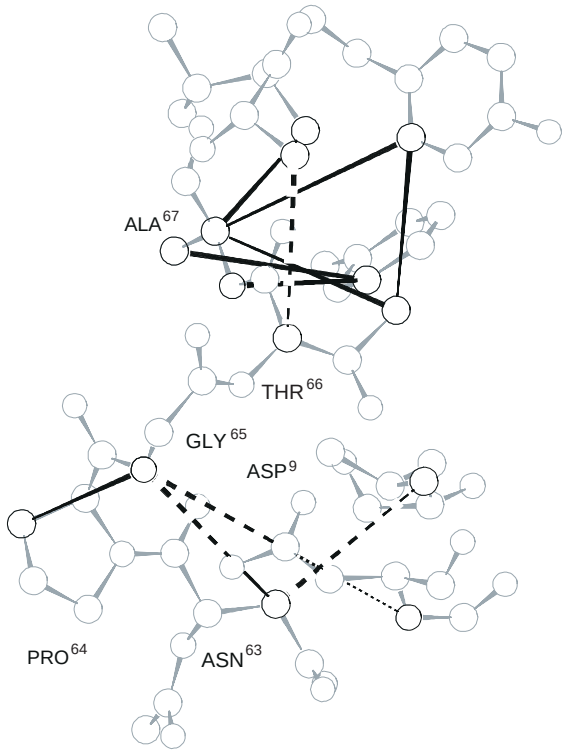

Fig. 7. NOE connectivities for which the solution (Jeng et al., 1994) and crystal (Katti et al., 1990) structures predict differences in interproton distances greater than 0.84 Å for the residues surrounding Gly^{65}. The solid lines indicate NOE connectivities for which the RFD NOE-derived distances agree most closely with the X-ray distances. The dashed lines indicate NOE connectivities for which the RFD NOE-derived distances agree most closely with the solution structure distances.

due to a partial resonance overlap, there are seven NOE cross peaks between Gly^{21} and the residues Ala^{19}, Asn^{83} and Gly^{84} H^N. The buildup rates for all of these cross peaks correspond to interproton distances for which the predicted X-ray value is at least 0.7 Å longer than the NOE-derived distances.

Molecular dynamics calculations (Go and Gō, 1976; McCammon et al., 1976) have predicted the occurrence of 'crankshaft' correlated oscillation of the ψ_{i-1} and ϕ_i dihedral angles which serve to wobble the intervening peptide plane. The observed S^2 value of 0.26 for Gly^{21} H^N (Stone et al., 1993) can be used to predict an angular diffusion of this peptide plane with the two adjacent C^α atoms fixed. For protons perpendicular to this plane, dynamic foreshortenings are calculated to range from 1.15 Å for a mean internuclear separation of 3.5 Å, to 0.6 Å for a mean internuclear separation of 6.0 Å. The agreement between these simple model calculations and the observed apparent foreshortening of the NOE-derived distances strongly suggests the presence of a substantial dynamical correction in this segment of the structure.

It should be noted that a concomitant error in the X-ray structure may also be present. Despite an overall excellent electron density map for the A molecule of the asymmetric unit, the electron density is broken at the position of Asp^{20} C^α while the side chain is not well positioned (Katti et al., 1990; LeMaster, 1990a). The implicit distortion may be propagated to the adjacent residues as well. In this regard, it should be noted that the crystallographic Debye–Waller (B-) factor for Asp^{20} C^α is a quite typical 17 $Å^2$, which serves to illustrate the need to avoid the overinterpretation of temperature factors as local monitors of structural uncertainty.

In both panels of Fig. 5, there is a cluster of distance discrepancies involving the irregular structure spanning the sequence between the α_3-helix and the 3_{10}-helix. As illustrated in the expanded view of Fig. 7, this segment is stabilized by the multiple hydrogen bonds formed with Asp^9. Although the two closely positioned clusters of distance discrepancies may arise from separate structural perturbations, their position in the tertiary structure indicates an alternative possibility. The analysis of ^{13}C and ^{15}N relaxation data indicates the presence of chemical exchange dynamics in this segment of the structure (LeMaster and Kushlan, 1996). For chemical exchange occurring near the fast exchange limit, the effective distance derived from NOE analysis reflects a weighted average of the various interchanging conformers. As no independent information is available for modeling these conformers, the estimation of a dynamical correction to the NOE-derived distances is not possible. However, it should be noted that the B-factors for the backbone of this segment are among the highest in the protein (Katti et al., 1990). Hence the most straightforward interpretation of these data is that the conformational transition reflected in the chemical exchange line broadening gives rise to dynamical distortions of both the NOE-derived distance constraints and the time-averaged X-ray structure.

Dynamical contributions to the ^1H-^1H cross-relaxation rate

As heteronuclear relaxation provides an experimental monitor for ps–ns motions, it is of interest to consider if a 'model-free' interpretation of the dynamics derived from heteronuclear relaxation can be used to estimate the dynamical contributions to the ^1H-^1H cross-relaxation rates. Such an estimation is not straightforward since the modulation of the ^1H-^1H vector can be viewed as a convolution of the reorientation of the two ^1H-X bond vectors and the relative translational motion of the two directly attached heavy atoms. The analysis is additionally complicated by the fact that heteronuclear relaxation and homonuclear cross relaxation have differing frequency responses to the local dynamics. Nevertheless, as argued below, a direct prediction of the dynamical contributions to ^1H-^1H cross relaxation can be obtained from experimentally accessible data using heteronuclear relaxation combined with the crystallographic B-factors as a monitor of the translational disorder of the heavy atoms.

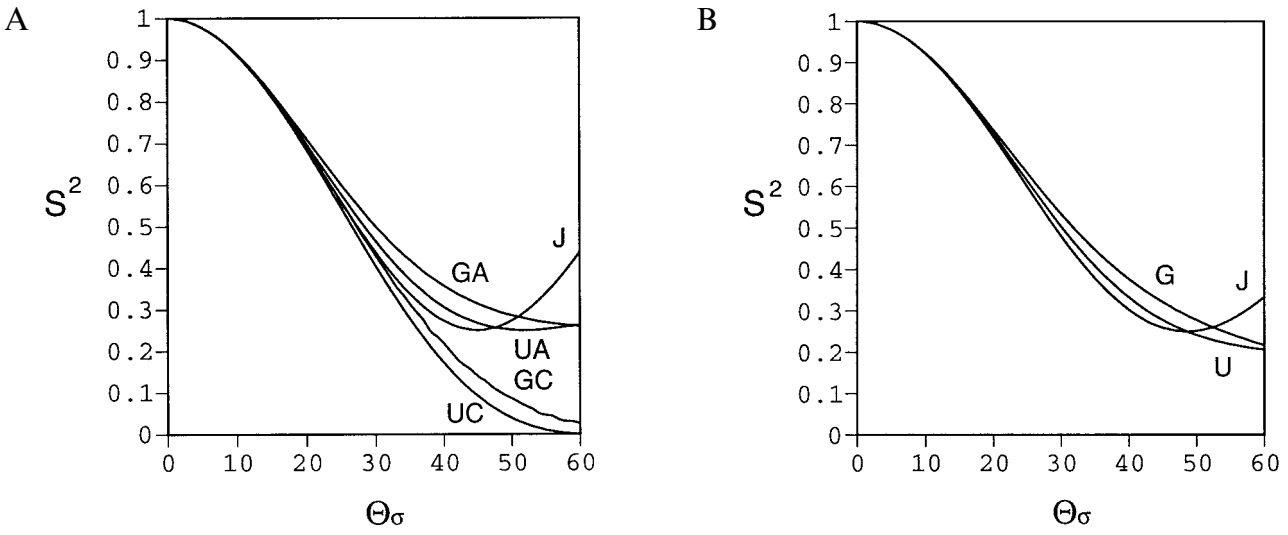

Fig. 8. Angular dependence of the heteronuclear order parameter S^2. Panel A illustrates the two-state jump model ($J = 1/4[1 + 3\cos^2(2\Theta_\sigma)]$) (Lipari and Szabo, 1982b), uniform diffusion in a planar arc ($UA = 1/4[1 + 3\cos^2(\sqrt{3}\,\Theta_\sigma)(\sin\sqrt{3}\,\Theta_\sigma)/\sqrt{3}\,\Theta_\sigma)^2]$) (Lipari and Szabo, 1982b), Gaussian diffusion in a planar arc ($GA = 1/4[1 + 3\exp(-2\Theta_\sigma)^2]$) (Bruschweiler and Wright, 1994) and uniform diffusion in a cone ($UC \sim 1/4\,[\cos\sqrt{2}\,\Theta_\sigma(1 + \cos\sqrt{2}\,\Theta_\sigma)]^2$) (Woessner, 1962). Also included are simulations of Gaussian diffusion in a cone as no analytical formulation has been proposed. The exact calculation of the uniform diffusion in a cone shown in the figure differs slightly from the simplified formula above as $\sqrt{2}\,\Theta_\sigma = \Theta_{max}$ at small angles and differs by less than 0.8% at $\Theta_\sigma = 30°$. For torsional diffusion in which the H-X vector spans the arc of a cone (i.e. the angle to the rotation axis is <90°), S^2 is standardly given in terms of the angle around the rotation axis rather than that spanned by the H-X vectors. Panel B illustrates uniform (Woessner, 1962) and Gaussian (Bruschweiler and Wright, 1994) torsional diffusion around a tetrahedral axis. Included as well are the calculations for the two-state jump model (Lipari and Szabo, 1982b), which are identical to that of panel A but recalibrated to the rotation axis angle by the law of cosines.

The dipole–dipole cross-relaxation rate constant Γ for spins i and j is given in terms of the spectral densities $J_{ij}(\omega)$ as (Ernst et al., 1987; Bruschweiler and Case, 1994):

$$\Gamma_{ij} = q[-1/2\, J_{ij}(0) + 3\, J_{ij}(2\omega_H)] \tag{1}$$

in which $q = (1/10)(\mu_0/4\pi)^2\gamma^4(h/2\pi)^2$, where μ_0 is the magnetic field constant, γ is the gyromagnetic ratio and ω_H is the ^{1}H Larmor frequency. In the case of an internally rigid ^{1}H-^{1}H vector attached to an isotropically tumbling macromolecule, the cross-relaxation rate exhibits a simple $1/r^6$ dependence. For macromolecules which exhibit molecular tumbling rates $1/\tau_c$ far slower than ω_H, there is a negligible dependence of Γ_{ij} on the $J_{ij}(2\omega_H)$ term even for rather small (e.g. 0.2) order parameters of the r_{ij} vector (Bruschweiler et al., 1992). Even for the case of methyl groups which exhibit rotation frequencies in the range of the ^{1}H Larmor frequency, for macromolecules the simple rapid threefold rotation model quite closely matches the cross-relaxation rates predicted from more detailed time-dependent models (Edmondson, 1994). As a result, for internal motions significantly more rapid than molecular tumbling, the internal correlation function characterizing those motions collapses to a static distribution of the r_{ij} vectors. As the spectral density function is the Fourier transform of the correlation function, the time independence of the summation over the r_{ij} vectors carries over directly. For the assumptions of isotropic molecular tumbling and the dynamical independence of the internal motions, the corresponding spectral density function is (Yip and Case, 1991):

$$J_{ij}(\omega) = \frac{2\tau_c}{5(1+\omega^2\tau_c^2)} \times \sum_{\alpha,\beta} p_\alpha p_\beta \sum_{m=-2}^{2} \left\langle \frac{Y_{2m}(\Theta_\alpha,\phi_\alpha)\, Y_{2m}^*(\Theta_\beta,\phi_\beta)}{r_{ij,\alpha}^3\, r_{ij,\beta}^3} \right\rangle \tag{2}$$

in which p_α is the probability for the interproton vector $r_{ij,\alpha}$, and the angles Θ and ϕ of the spherical harmonics Y_{2m} are defined with respect to the molecular frame for the interproton vector between spins i and j.

Given the effectively rigid bond length, the dynamics of heteronuclear relaxation is characterized by the angular order parameter (Lipari and Szabo, 1982a):

$$S_\Omega^2 = \frac{4\pi}{5} \sum_{m=-2}^{2} |\langle Y_{2m}(\Theta,\phi)\rangle|^2 = \langle P_2(\chi)\rangle \tag{3}$$

where the equality to the second-rank Legendre polynomial utilizes the addition theorem of spherical harmonics. Much of the conceptual appeal of S^2 comes from its relationship to the variance of the probability-weighted spherical harmonics characterizing the orientational disorder of the heteronuclear vector (Bruschweiler and Wright, 1994):

$$S^2 = 1 - \frac{4\pi}{5} \sum_{m=-2}^{2} \sigma_{Y_{2m}}^2 \tag{4}$$

In fact, for moderate levels of internal dynamics (e.g. >0.6), the S^2 value is effectively determined solely by the angular variance of the H-X vector as illustrated in Fig. 8. Panel A shows the angular dependence of S^2 for the two-state jump model, uniform and Gaussian diffusion in a planar arc, and uniform and Gaussian diffusion in a cone. In contrast to the traditional representations, each of the models are expressed in terms of θ_σ, the standard deviation of the angular distribution. The two-state jump model ($1-S^2 = 3/4\ \sin^2 2\theta_\sigma$) provides an excellent representation for the other continuous distributions for limited to moderate levels of mobility ($\pm 1.5°$ at $S^2 = 0.6$). Likewise, as seen in panel B for torsional fluctuations at a tetrahedral site, θ_σ reliably characterizes the relaxation effects.

For long-range NOE interactions, justification of an assumption of uncorrelated motion of the cross-relaxing protons has been obtained via molecular dynamics simulations on bovine pancreatic trypsin inhibitor (LeMaster et al., 1988). Combined with the model independence of the heteronuclear S^2, it is plausible that the dynamics of the H-X vectors might likewise affect the ^{1}H-^{1}H cross relaxation in a model-independent fashion. To test this prediction, two ^{1}H-X bond vectors initially oriented antiparallel were allowed to reorient according to either a conical diffusion model or a torsional diffusion model as a function of the ^{1}H-X order parameter. Figure 9 shows the correlation function (normalized to $1/r^6$ for the static orientation) for these calculations. Indeed the correlation function, and far more so the sixth root apparent distance estimate, is largely insensitive to the motional model assumed for the H-X vectors. Although an approximately antiparallel orientation is to be expected in most cases for long-range NOEs near the van der Waals contact distance, it should be noted that the dynamics of H-X vectors perpendicular to the interproton vector are model dependent.

Although for ^{1}H-^{1}H cross relaxation the radial and angular components of the internal correlation function cannot be rigorously separated, given the conceptual benefit of separability, a radial order parameter $S^2_{r^6}$ has been proposed (Olejniczak et al., 1984):

$$S^2_{r^6} S^2_\Omega \sim S^2 = \frac{4\pi}{5} \left\langle \frac{1}{r_{ij}^6} \right\rangle^{-1} \sum_{m=-2}^{2} \left| \left\langle \frac{Y_{2m}(\Theta,\phi)}{r^3} \right\rangle \right|^2 \qquad (5)$$

Molecular simulations (Bruschweiler et al., 1992; Post, 1992) have demonstrated that generally the assumption of separability is well justified. On the other hand, the factoring out of a $\langle 1/r_{ij}^6 \rangle$ term in this order parameter formalism can give rise to potentially misleading results. As illustration, order parameter calculations have been reported for an interproton vector with the hydrogen positions drawn from two nonoverlapping uniform spherical distributions (Bruschweiler et al., 1992). The product $S^2_\Omega S^2_{r^6} \sim S^2$ is strongly dependent on the distance between the mean positions of the protons, despite the fact that the correlation function for any pair of spherically sym-

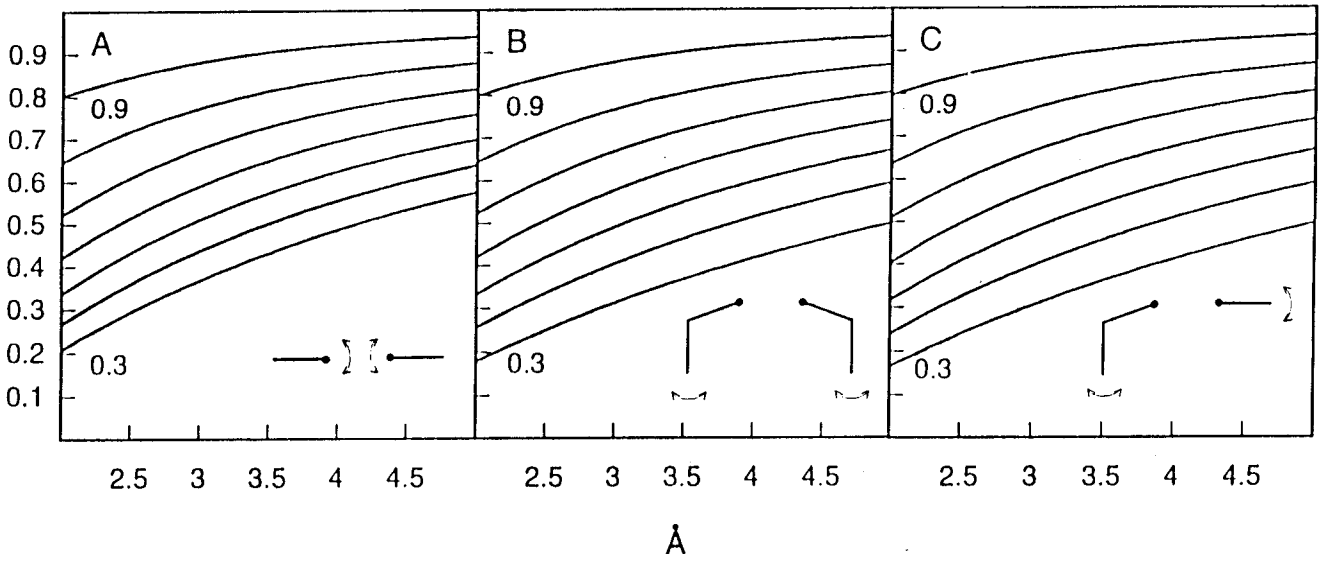

Fig. 9. Normalized ^{1}H-^{1}H correlation function values as a function of internuclear distance and of the order parameters of the ^{1}H-X vectors (assuming a 1.09 Å bond length). The correlation function is normalized to the interproton distance in the antiparallel orientation. Panel A illustrates the case of two ^{1}H-X vectors undergoing equivalent uniform conical diffusion (S^2 varied from 0.3 to 0.9 in 0.1 increments) in which the axes of diffusion are directed toward each other. Panels B and C illustrate torsional diffusion around axes perpendicular to the interproton vector for the antiparallel orientation. The axes of rotation are assumed parallel in panel B and perpendicular in panel C. The similarity between these results indicates, for the antiparallel H-X orientation, negligible dependence on the detailed motional model assumed for the H-X vector motion. It should be noted that variations in the normalized correlation function illustrated in Fig. 10 primarily reflect the fact that in all dynamical orientations r_{ij} exceeds that of the antiparallel orientation used for normalization.

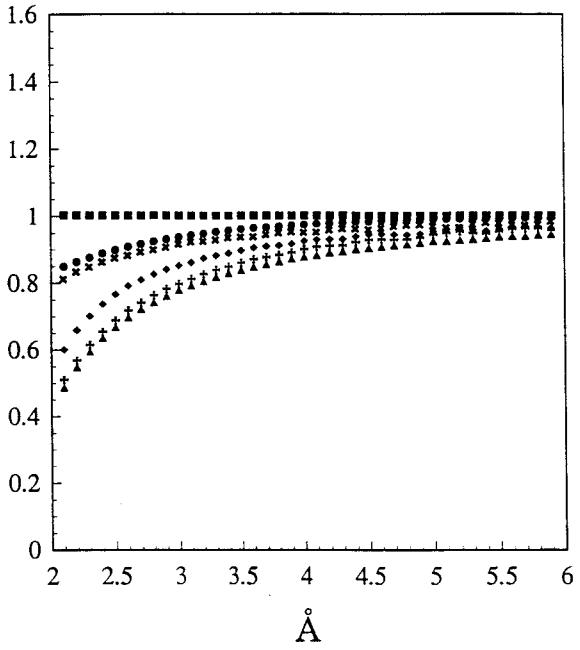

Fig. 10. ^{1}H-^{1}H order parameters for two independent isotropic Gaussian distributed nuclei. Numerical simulations utilizing Eq. 2 were carried out for 10^7 interproton vector pairs at each average distance. The σ for the distributions is assumed to be 0.5 Å (consistent with typical crystallographic temperature factors (Ringe and Petsko, 1985)) truncated at 2σ. Values are plotted for the correlation function normalized to $1/r_m^6$ (■), angular order parameter S_Ω^2 (●), radial order parameters $S_{r_m}^2$ (✕) and $S_{r^6}^2$ (◆) as well as the products $1/(r_m^6\langle 1/r_{ij}^6\rangle)$ (✚) and $S_\Omega^2\, S_{r^6}^2$ (▲).

metric nonoverlapping uncorrelated distributions when normalized to $1/r_{mean}^6$ is known to be analytically independent of the separation distance (LeMaster et al., 1988). Expressed equivalently, in this case the decrease in cross-relaxation rate resulting from angular disorder is perfectly balanced by the enhancement of the cross-relaxation rate due to the radial dispersion. An analogous calculation is illustrated in Fig. 10. As is readily apparent, $S_\Omega^2 S_{r^6}^2 \sim S^2 \sim 1/(r_m^6\langle 1/r_{ij}^6\rangle)$. Hence in this order parameter formalism S^2 merely serves to extract out the inverse of the $\langle 1/r_{ij}^6\rangle$ normalization factor.

Herein we consider an alternate radial order parameter $S_{r_m}^2$ defined as:

$$S_{r_m}^2 = 1 / \left(r_m^6 \langle 1/r_{ij}^3\rangle^2\right) \tag{6}$$

This radial order parameter representation has the obvious benefit that, in marked contrast to the $\langle 1/r_{ij}^6\rangle$ normalization factor, the separation between the means r_m is generally the desired geometric parameter and is often independently experimentally accessible. Furthermore, when combined with S_Ω^2, $S_{r_m}^2$ serves to define spherically symmetric atomic fluctuations as a natural reference state for assessing dynamical effects in cross relaxation.

Under the assumptions of internal motion significantly faster than and dynamically independent of the (isotropic) molecular tumbling and that of the separability of the radial and angular components, the spectral densities depend on the ratio of order parameters and the distance between the mean nuclear positions:

$$J_{ij}(\omega) \sim \frac{\tau_c}{2\pi\left(1+\omega^2\tau_c^2\right)} \frac{S_\Omega^2}{S_{r_m}^2\, r_m^6} \tag{7}$$

The individual radial and angular order parameters are not experimentally accessible. They enter the spectral density as a ratio reflecting the opposite effects of angular and radial disorder on the cross-relaxation rate. This ratio presents an asymmetry factor which characterizes the degree to which the distribution of r_{ij} vectors differs from that predicted for the independent spherically symmetrically distributed nuclei. The calculations illustrated in Fig. 10 are readily extended to the use of anisotropic thermal ellipsoids. This approach offers the appeal that, in principle, the anisotropic thermal ellipsoids are experimentally accessible from X-ray diffraction analysis, although in practice only data from the very best diffracting protein crystals have yielded these parameters (Teeter et al., 1993). Previous analyses of the effect of anisotropic motion on the predicted cross-relaxation rates have emphasized that this effect attenuates rather slowly with increased internuclear separation (LeMaster et al., 1988).

The model of independent Gaussian fluctuations for the two cross-relaxing protons allows interproton distances well less than the sum of the van der Waals radii of the atoms. These occurrences can be readily eliminated in simulations by imposing a minimal distance criterion on the pairs of interproton vectors used to calculate the correlation function. However, determination of the corresponding mean positions of each nucleus then becomes more complicated. Both anharmonicity in the potential energy functions and correlated motion of the nuclei will shift the relevant mean positions away from what would be observed experimentally as the average structure. Since in practice the desired information is usually how much the introduction of dynamical disorder shifts the apparent distance relative to the 'static' model, an additional 'order parameter' can then be defined:

$$S_{r_s}^2 = 1 / \left(r_s^6 \langle 1/r_{ij}^3\rangle^2\right) \tag{8}$$

The behavior of $S_{r_s}^2$ is illustrated in Fig. 11, in which the calculations from Fig. 10 were repeated with a 1.8 Å exclusion distance to mimic the hard-sphere repulsion between the protons. The earlier 2σ cutoff was also removed. As expected, the effects of the steric exclusion are comparatively negligible above 3.0 Å and become rapidly more severe at shorter distances. Most notably, the value of $S_{r_s}^2$ rises above 1.0 at shorter distances in contrast to $S_{r_m}^2$ which is mathematically constrained to the [0,1] interval as is generally the case in the definition of correlation

functions and order parameters. Values of $S^2_{r_s}$ greater than 1.0 clearly signal a shifting of the mean positions away from those of the static model. Indeed at 2.0 Å, the radial contribution represents approximately 2/3 of the apparent lengthening of the interproton distance. In contrast, S^2_Ω is comparatively unaffected by the introduction of the steric exclusion correction. Given the acute inverse sixth power dependence on distance, even modest shifts in the normalizing distance can substantially alter the resultant interpretation in terms of radial versus angular dispersion.

Modeling ^{1}H-^{1}H cross-relaxation rates from heteronuclear relaxation data

The dynamical corrections to the ^{1}H-^{1}H cross-relaxation rates may be estimated by a fluctuation model assuming the independence of the heavy-atom translational and ^{1}H-X orientational motion. This analysis has been applied to *E. coli* thioredoxin utilizing our recently reported dynamical analysis of 413 independent H-C and H-N vectors (LeMaster and Kushlan, 1996). As the X-ray refinement of this protein utilized isotropic Debye–Waller factors (Katti et al., 1990), the effect of asymmetry in the heavy-atom motion cannot be represented. Furthermore, given the additional complexities of the quantitative dynamical interpretation of experimental crystallographic B-factors (Ringe and Petsko, 1985), the present analysis merely assumes a uniform 0.5 Å Gaussian dispersion consistent with average crystallographic and molecular dynamic results.

The effects of dynamics on the correlation function for sequential αN connectivities of *E. coli* thioredoxin were calculated assuming that uniform rotational diffusion around ψ is responsible for the S^2 values of the H^α-C^α and H^N-N vectors. For the main-chain sites lacking relaxation data (20% of all sites), an S^2 value of 0.85 was assumed. For the α-helical sequential αN connectivities, the average apparent distance was essentially unaffected (net decrease by 0.5%). In addition to the moderate distance (3.55 Å), there is cancellation between the angular and radial contributions since in this conformation the H-X bonds are in a trans orientation along the C^α-C' bond. This predicted insensitivity to torsional fluctuations further supports the utility of the α-helical sequential αN connectivities for defining the NOE distance calibration. In contrast, for the eclipsed conformation of the β-strand sequential αN connectivities, both the radial and angular fluctuations lead to increased apparent interproton distances and hence further contribute to the overestimation of the apparent NOE distance. As noted above, the RFD NOE-derived distances for the β-strand sequential connectivities overestimated the corresponding X-ray distances by an average of 0.23 Å. Incorporation of the calculated dynamical correction reduced this average discrepancy to 0.09 Å.

For proton pairs separated by at least two flexible dihedral angles for which relaxation data were available at both sites, ^{1}H-^{1}H correlation functions were calculated assuming uniform diffusion in a cone for the ^{1}H-^{13}C and ^{1}H-^{15}N vectors. Use of the conical diffusion model is anticipated to minimize artifacts arising from the residual dependence of ^{1}H-^{1}H cross-relaxation rates on the details of the ^{1}H-X bond reorientation. If in addition there were at least two flexible dihedral angles separating the corresponding heteroatoms, the ^{1}H positions were additionally displaced according to an isotropic Gaussian distribution of σ equal to 0.5 Å. Steric exclusion was applied assuming hard-sphere radii of 0.9 Å, 1.2 Å and 1.5 Å for ^{1}H, ^{15}N and ^{13}C, respectively.

When this simple independent fluctuation model of conical H-X bond motion superimposed on isotropic heavy-atom displacement was applied to the NOE pairs of Table 1, the average deviation of the dynamically corrected RFD NOE-derived distances for the more generally mobile methylene resonances decreased to 10.7% from the 13.2% estimated from the static analysis. Similarly, for the IS + I_3S cross peaks containing a heteroatom with $S^2 < 0.5$ the average deviation decreased from 13.2% to 11.3%. In contrast, for the IS + I_3S cross peaks for less mobile positions dynamical corrections did not significantly lower the average distance discrepancy. These results reinforce the interpretation that the consistent ~10% average discrepancy is not dominated by dynamical distortions of the distance estimates.

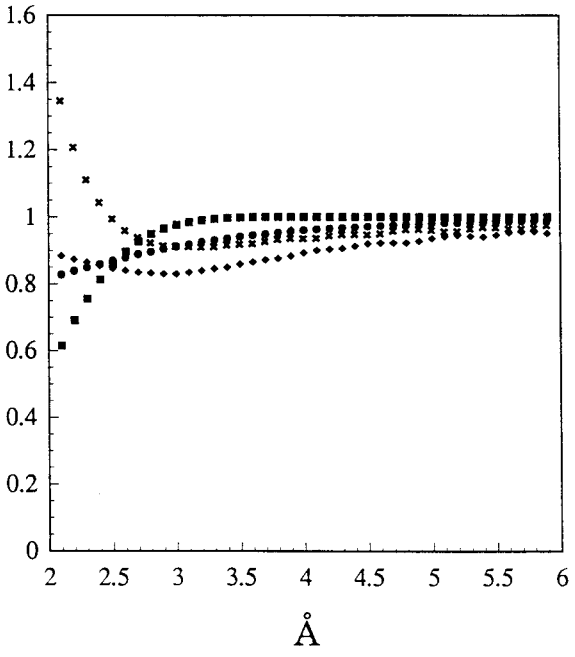

Fig. 11. ^{1}H-^{1}H order parameters for two independent isotropic Gaussian distributed nuclei assuming steric exclusion. Calculations were analogous to those of Fig. 10 with the inclusion of a hard-sphere radius of 0.9 Å and the elimination of the 2σ cutoff for the Gaussian distributions. Values are plotted for the correlation function normalized to $1/r_s^6$ (■), angular order parameter S^2_Ω (●), and radial order parameters $S^2_{r_s}$ (✕) and $S^2_{r^6}$ (◆).

TABLE 3
AVERAGE VALUES OF NOE CROSS-RELAXATION PARAMETERS MODELED FROM HETERONUCLEAR DATA OF *E. coli* THIOREDOXIN COMPARED TO THOSE OBTAINED VIA MOLECULAR DYNAMICS SIMULATION

	$S_\Omega^2/S_{r_s}^2$		S_Ω^2		$S_{r_s}^2$	
	MD	Relaxation	MD	Relaxation	MD	Relaxation
Interior						
Interresidue, long range[a]	0.94 (0.51)b	0.95 (0.09)	0.91 (0.08)	0.93 (0.03)	0.96 (0.57)	0.98 (0.08)
Interresidue, short range	0.97 (0.26)	0.90 (0.16)	0.93 (0.06)	0.92 (0.05)	0.95 (0.27)	1.03 (0.16)
Intraresidue, nonvicinal	0.89 (0.23)	0.90 (0.21)	0.86 (0.12)	0.93 (0.05)	0.95 (0.30)	1.05 (0.21)
Exterior						
Interresidue, long range	0.86 (0.49)	0.95 (0.11)	0.86 (0.09)	0.93 (0.04)	1.01 (0.54)	0.98 (0.09)
Interresidue, short range	1.01 (0.34)	0.97 (0.14)	0.91 (0.06)	0.92 (0.05)	0.90 (0.30)	0.95 (0.13)
Intraresidue, nonvicinal	0.84 (0.23)	0.89 (0.23)	0.80 (0.14)	0.92 (0.05)	0.95 (0.23)	0.96 (0.21)

Molecular simulation data calculated for hen lysozyme (Post, 1992).
[a] Long range is defined as residues more than three apart.
[b] Rmsd.

For a large proportion of the NOE cross peaks of *E. coli* thioredoxin, the dynamical correction to the derived NOE distance is predicted to be appreciably smaller than the average deviation for the data of Table 1. Hence the calculations can only be effectively compared for the more mobile positions for which the analyses of the dynamical corrections are least reliable. In order to assess whether, on average, these calculations reasonably estimate the dynamical corrections for the less mobile positions, comparison was made to previous molecular dynamics simulations. Table 3 gives the average values of the motional averaging parameters from the simulation on hen lysozyme (Post, 1992). Included as well are the corresponding population averages obtained by applying the independent fluctuation model to the RFD NOE cross peaks of *E. coli* thioredoxin for which relaxation data were available for both heteronuclei. It is anticipated that the assumptions of independent proton motion and limited mobility should be best met for the long-range interior proton pairs. Indeed for the mean values for S_Ω^2, $S_{r_s}^2$ and the asymmetry parameter $S_\Omega^2/S_{r_s}^2$, the agreement between the two calculations is excellent. For the solvent-exposed long-range interactions, the independent fluctuation model underestimates the dynamical corrections from the molecular dynamics simulations by approximately a factor of 2. This underestimation presumably reflects the limitations of both the assumption of independence of heavy-atom displacement and H-X bond oscillation as well as the inapplicability of the conical diffusion model for the more highly mobile surface side chains.

The lysozyme molecular dynamics results quoted in Table 3 were based on referencing the interproton distances to an energy-minimized structure rather than by the use of the mean proton positions obtained directly from the trajectories. Post (1992) noted that when the radial order parameters were normalized to the mean proton positions their distribution was one-sided and much narrower, as anticipated theoretically. The exceedingly broad lysozyme $S_{r_s}^2$ distribution reflects the acute sensitivity of this order parameter to the choice of normalization distance (most atoms shifted less than 0.25 Å during minimization) (Post, 1992). This broad distribution in turn tends to obscure the NOE attenuating effect of asymmetric dynamical radial fluctuations manifested in the $S_{r_s}^2$ values greater than 1.0 and its relevance to the interpretation of the radial–angular compensation effect.

For the short-range and nonvicinal intraresidue interactions, the predictions of the mean asymmetry factors are reasonably good. These are the most challenging cases as they are not realistically represented by either the assumption of independent fluctuations or of strictly correlated dihedral angle oscillations as considered above. Nevertheless, despite the simplifying assumptions of the proposed fluctuation model, there is substantial general agreement between the predictions of molecular dynamics simulations and the predictions based on experimental relaxation and structural data.

Acknowledgements

This work was supported by the National Institutes of Health Grant GM 38779. The NMR Facility at Madison is acknowledged for collection of the ^{1}H spectrum of partially deuterated isoleucine. H.J. Dyson is thanked for making available the coordinates of the solution structure of *E. coli* thioredoxin previous to their becoming available through the Brookhaven Databank.

References

Abseher, R., Ludemann, S., Schreiber, H. and Steinhauser, O. (1995) *J. Mol. Biol.*, **249**, 604–624.

Baleja, J.D., Moult, J. and Sykes, B.D. (1990) *J. Magn. Reson.*, **87**, 375–384.

Boelens, R., Koning, T.M.G. and Kaptein, R. (1988) *J. Mol. Struct.*, **173**, 299–311.

Borgias, B.A. and James, T.L. (1990) *J. Magn. Reson.*, **87**, 475–487.

Boulat, B., Konrat, R., Burghardt, I. and Bodenhausen, G. (1992) *J. Am. Chem. Soc.*, **114**, 5412–5414.
Boulat, B. and Bodenhausen, G. (1993) *J. Biomol. NMR*, **3**, 335–348.
Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.
Brunne, R.M., van Gunsteren, W.F., Bruschweiler, R. and Ernst, R.R. (1993) *J. Am. Chem. Soc.*, **115**, 4764–4768.
Bruschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M. and Ernst, R.R. (1992) *J. Am. Chem. Soc.*, **114**, 2289–2302.
Bruschweiler, R. and Case, D.A. (1994) *Prog. NMR Spectrosc.*, **26**, 27–58.
Bruschweiler, R. and Wright, P.E. (1994) *J. Am. Chem. Soc.*, **116**, 8426–8427.
Chandrasekhar, K., Campbell, A.P., Jeng, M.F., Holmgren, A. and Dyson, H.J. (1994) *J. Biomol. NMR*, **4**, 411–432.
Chunduru, S.K., Mrachko, G.T. and Calvo, K.C. (1989) *Biochemistry*, **28**, 486–493.
Clore, G.M. and Gronenborn, A.M. (1989) *J. Magn. Reson.*, **84**, 398–409.
Crespi, H.L., Rosenberg, R.M. and Katz, J.J. (1968) *Science*, **161**, 795–796.
Crout, D.H.G., Gregorio, M.V.M., Muller, U.S., Komatsubara, S., Kisumi, M. and Chibata, I. (1980) *Eur. J. Biochem.*, **106**, 97–105.
Dyson, H.J., Holmgren, A. and Wright, P.E. (1989) *Biochemistry*, **28**, 7074–7087.
Edmondson, S.P. (1994) *J. Magn. Reson.*, **B103**, 222–233.
Ernst, R.R., Bodenhausen, G. and Wokaun, A. (1987) *Principles of NMR in One and Two Dimensions*, Clarendon Press, Oxford, U.K.
Fejzo, J., Westler, W.M., Macura, S. and Markley, J.L. (1991) *J. Magn. Reson.*, **92**, 195–202.
Fejzo, J., Westler, W.M., Markley, J.L. and Macura, S. (1992) *J. Am. Chem. Soc.*, **114**, 1523–1524.
Galimov, E.M. (1985) *The Biological Fractionation of Isotopes*, Academic Press, New York, NY, U.S.A.
Go, M. and Gō, N. (1976) *Biopolymers*, **15**, 1119–1127.
Griffey, R.H. and Redfield, A.G. (1987) *Q. Rev. Biophys.*, **19**, 51–82.
Grzesiek, S. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 335–339.
Güntert, P. and Wüthrich, K. (1992) *J. Magn. Reson.*, **96**, 403–407.
Hansen, P.E. (1988) *Prog. NMR Spectrosc.*, **20**, 207–255.
Hoogstraten, C.G., Westler, W.M., Macura, S. and Markley, J.L. (1993) *J. Magn. Reson.*, **B102**, 232–235.
Hoogstraten, C.G., Choe, S., Westler, W.M. and Markley, J.L. (1995a) *Protein Sci.*, **4**, 2289–2299.
Hoogstraten, C.G., Westler, W.M., Macura, S. and Markley, J.L. (1995b) *J. Am. Chem. Soc.*, **117**, 5610–5611.
Jeng, M.-F., Campbell, A.P., Begley, T., Holmgren, A., Case, D.A., Wright, P.E. and Dyson, H.J. (1994) *Structure*, **2**, 853–868.
Katti, S., LeMaster, D.M. and Eklund, H. (1990) *J. Mol. Biol.*, **212**, 167–184.
Konrat, R., Burghardt, I. and Bodenhausen, G. (1991) *J. Am. Chem. Soc.*, **113**, 9135–9140.
Kushlan, D.M. and LeMaster, D.M. (1993) *J. Biomol. NMR*, **3**, 701–708.
LeMaster, D.M. and Richards, F.M. (1982) *Anal. Biochem.*, **122**, 238–247.
LeMaster, D.M. (1987) *FEBS Lett.*, **223**, 191–196.
LeMaster, D.M., Kay, L.E., Brünger, A.T. and Prestegard, J.H. (1988) *FEBS Lett.*, **236**, 71–76.
LeMaster, D.M. and Richards, F.M. (1988) *Biochemistry*, **27**, 142–150.
LeMaster, D.M. (1990a) *Q. Rev. Biophys.*, **23**, 133–174.
LeMaster, D.M. (1990b) *Annu. Rev. Biophys. Biophys. Chem.*, **19**, 243–266.
LeMaster, D.M. (1994) *Prog. NMR Spectrosc.*, **26**, 371–419.
LeMaster, D.M. and Kushlan, D.M. (1996) *J. Am. Chem. Soc.*, **118**, 9255–9264.
Li, Y.C. and Montelione, G.T. (1993) *J. Magn. Reson.*, **B101**, 315–319.
Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4559.
Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.
Liu, H., Spielmann, P., Ulyanov, N.B., Wemmer, D.E. and James, T.L. (1995) *J. Biomol. NMR*, **6**, 390–402.
Markley, J.L., Putter, I. and Jardetzky, O. (1968) *Science*, **161**, 1249–1251.
Massefski, W. and Redfield, A.G. (1988) *J. Magn. Reson.*, **78**, 150–155.
McCammon, J.A., Gelin, B.R., Karplus, M. and Wolynes, P.G. (1976) *Nature*, **262**, 325–326.
Olejniczak, E.T., Dobson, C.M., Karplus, M. and Levy, R.M. (1984) *J. Am. Chem. Soc.*, **106**, 1923–1930.
Olejniczak, E.T., Gampe, R.T. and Fesik, S.W. (1986) *J. Magn. Reson.*, **67**, 28–41.
Post, C.B., Meadows, R.P. and Gorenstein, D.G. (1990) *J. Am. Chem. Soc.*, **112**, 6796–6803.
Post, C.B. (1992) *J. Mol. Biol.*, **224**, 1087–1101.
Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1989) *Numerical Recipes*, Cambridge University Press, Cambridge, U.K.
Ringe, D. and Petskö, G.A. (1985) *Prog. Biophys. Mol. Biol.*, **45**, 197–235.
Schmidt, J.M., Bruschweiler, R., Ernst, R.R., Dunbrack, R.L., Joseph, D. and Karplus, M. (1993) *J. Am. Chem. Soc.*, **115**, 8747–8756.
Seip, S., Balbach, J. and Kessler, H. (1992) *J. Magn. Reson.*, **100**, 406–410.
Stone, M.J., Chandrasekhar, K., Holmgren, A., Wright, P.E. and Dyson, H.J. (1993) *Biochemistry*, **32**, 426–435.
Summers, M.F., South, T.L., Kim, B. and Hare, D.R. (1990) *Biochemistry*, **29**, 329–340.
Teeter, M.M., Roe, S.M. and Heo, N.H. (1993) *J. Mol. Biol.*, **230**, 292–311.
Thomas, P.D., Basus, V.J. and James, T.L. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1237–1241.
Torchia, D.A., Sparks, S.W. and Bax, A. (1988) *J. Am. Chem. Soc.*, **110**, 2320–2321.
Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.
Tsang, P., Wright, P.E. and Rance, M. (1990) *J. Am. Chem. Soc.*, **112**, 8183–8185.
Woessner, D.E. (1962) *J. Chem. Phys.*, **36**, 1–4.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.
Yip, P. and Case, D.A. (1991) *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy*, NATO ASI Series, Vol. A225, Plenum Press, New York, NY, U.S.A., pp. 317–330.
Zhao, D. and Jardetzky, O. (1994) *J. Mol. Biol.*, **239**, 601–607.