Journal of Biomolecular NMR, **12:** 259–276, 1998.
KLUWER/ESCOM

Some NMR experiments and a structure determination employing a {^{15}N,^{2}H} enriched protein

T.K. Mal, S.J. Matthews[a], H. Kovacs, I.D. Campbell & J. Boyd
Oxford Centre for Molecular Sciences and Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, U.K.
[a] *Department of Biochemistry, Imperial College of Science and Technology and Medicine, Exhibition Road, South Kensington, London, SW7 2AY, U.K.*

Received 26 November 1997; Accepted 26 March 1998

Key words: isotopic enrichment {^{15}N,^{2}H}, ^{1}H relaxation, structure calculations, isotope shifts, solvent interactions

Abstract

We present the results of studies of an aqueous sample of a highly {^{15}N,^{2}H} enriched protein, the SH3 domain from Fyn. Measurements of ^{1}H relaxation and interactions between H_2O solvent and exchangeable protons are given, as well as a method for increasing the effective longitudinal relaxation of solvent exchangeable proton resonances. The long-range isotope shifts are measured, for ^{1}H and ^{15}N, which arise due to perdeuteration. Simulations, which employed a 7 or 8 spin relaxation matrix analysis, were compared to the experimental data from a time series of 2D NOESY datasets for some resonances. The agreement between experiment and simulation suggest that, with this ^{1}H dilute sample, relatively long mixing times (up to 1.2 s) can be used to detect specific dipolar interactions between amide protons up to about 7Å apart. A set of 155 inter-amide NOEs and 7 side chain NOEs were thus identified in a series of 3D HSQC-NOESY-HSQC experiments. These data, alone and in combination with previously collected restraints, were used to calculate sets of structures using X-PLOR. These results are compared to the available X-ray and NMR structures of the Fyn SH3 domain.

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; NOESY, nuclear Overhauser enhancement spectroscopy; HSQC, heteronuclear single quantum correlation; SA, simulated annealing; Tris-HCl, tris(hydroxymethyl)aminomethane-hydrochloride; SDS-PAGE, sodium dodecylsulphate-polyacrylamide gel eletrophoresis; PBS, buffered saline; CSA, chemical shift anisotropy; DSSP, define secondary structure of proteins.

Introduction

Several groups have explored the use of fully or partially ^{2}H isotopically enriched proteins (Crespi et al., 1968; Markley et al., 1968; Torchia et al., 1988; LeMaster et al., 1988; Tsang et al., 1990) in combination with ^{15}N enrichment (Grzesiek et al., 1995) and ^{15}N and^{13}C enrichment (Grzesiek et al., 1993; Markus et al., 1994; Yamazaki et al., 1994; Farmer and Venters, 1995; Venters et al., 1995; Venters et al., 1996; Nietlispach et al., 1996; Gardner et al., 1997; Yamazaki et al., 1997; Sattler et al., 1997, Yu et al., 1997). The primary focus of these papers was to establish a route to the structure determination of larger proteins than is possible with ^{15}N and ^{13}C isotopic enrichment alone. Examples are a Trp repressor-DNA complex (Zhang et al., 1994), Shc phosphotyrosine binding domain (Zhou et al., 1995), and the protein Bcl-X_L (Muchmore et al., 1996).

In this paper we use an aqueous sample of Fyn SH3 domain, where some effort was made to obtain as high a level of {^{15}N,^{2}H} enrichment as possible. The structure of the Fyn SH3 domain from the Fyn tyrosine kinase (Cooke and Perlmutter, 1989) has been determined by both X-ray crystallography (Noble et al., 1993) and NMR (Morton et al., 1996) and shown to

have two 3-stranded antiparallel β sheets with a single turn of 3_{10} helix. The aim of the current work was to characterise the effects of complete deuteration on various parameters, including the ^{1}H amide longitudinal and transverse relaxation rates. We have also measured the long-range isotope shifts which occur with deuteration and report some experiments which probe the interaction of water with the solvent-exposed amides and hydroxyl groups. An important feature of this study was prior knowledge of the amide ^{1}H and ^{15}N resonance assignments from previous studies with a separate protonated ^{15}N enriched sample (Morton et al., 1996).

It has previously been suggested that a restricted NOE dataset from a deuterated sample might provide a method for establishing the global fold of a protein (Venters et al., 1995; Smith et al., 1996). Here we exploit the relatively slow ^{1}H longitudinal relaxation rate between the isotopically dilute ^{1}H spins to record 2D homonuclear NOESY (Wüthrich, 1986) and 3D ^{1}H-^{15}N HSQC-NOESY-HSQC (Frenkiel et al., 1990) datasets over a wide range of mixing times. These 2D and 3D NOESY experiments have made possible the detection of specific dipolar interactions over significantly larger distances than can be obtained from a sample without ^{2}H labelling. The observed NOE intensities were divided into three broad classes which were converted into distance upper-limit constraints of 3 Å, 4 Å and 6.8 Å. 155 amide-amide and 7 side chain-amide NOEs were identified. Structure calculations were carried out using either this set of NOE restraints alone, or with the inclusion of previously determined ^3J(HN,Hα) vicinal coupling constants (Morton et al., 1996). The resulting family of backbone NMR structures with their relative statistics are compared to the X-ray structure as well as a new family of NMR structures, computed using the previous restraints list augmented with the new NOE restraints.

Materials and Methods

Protein preparation

The pGEX-2T plasmid containing the gene for Fyn SH3 domain was transformed in the *E. Coli* cell line BL21 and then grown in vitamin supplemented M9 minimal medium (6 g/l Na_2HPO_4, 3 g/l KH_2PO_4, 0.5 g/l NaCl, 1 ml of 1M $MgSO_4$, 1 ml of 0.1M $CaCl_2$, 1 ml of 1M thiamine, 25 ml 20% glucose and 10 ml of *E. coli* trace elements) with $^{15}NH_4Cl$ as the sole nitrogen source. The resulting bacterial cells were recultured in M9 minimal medium containing sodium acetate and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources, respectively. The bacterial cells obtained from this cell growth were transferred to 50% deuterated M9 minimal medium. Finally these cells were grown, at 37 °C with vigorous aeration, in 3 l of fully deuterated M9 minimal medium, which contained CD_3COONa and$^{15}ND_4Cl$. All the components used in this M9 minimal medium had been fully exchanged with deuterium before use. The bacterial cells were found to grow approximately one quarter as fast as those grown under similar conditions in aqueous medium. Bacterial cell growth was monitored until the optical density at 600 nm was 0.5. The cells were induced with isopropyl-β-D-thiogalactopyranoside (final concentraition, 120 mg/l) for 6 h and harvested by centrifugation at 4 °C before suspending in ice-cooled PBS lysis buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 and 0.5 mM phenylmethylsulfonylfluoride). The bacterial cell suspension was sonicated for 1 min with an interval of 2 min for 6 cycles at 0 °C. The resulting cell debris was pelletted using centrifugation for 45 min at 15k rpm and 4 °C. The supernatant was passed through a Glutathione Sepharose 4B affinity column (Pharmacia Biotech) which had been pre-equilibrated with PBS buffer. After extensive washing with PBS (8 bed volumes) the protein was eluted with buffer (50 mM Tris-HCl and 10 mM reduced glutathione, pH 8.4). The protein-containing solution was then incubated with thrombin at room temperature and the cleavage was monitored by SDS-PAGE electrophoresis. The cleaved proteins were directly loaded onto a Sephacryl s-100 gel filtration column (2.6 cm × 90 cm) which had been pre-equilibrated with elution buffer 50 mM Tris-HCl and 150 mM NaCl, pH 8.4. The protein was collected, dialysed against water and freeze-dried. At the end of the purification procedure all the exchangeable deuterons were found to have completely exchanged to protons as judged by mass spectroscopy. The D_2O, 99.8%, was purchased from Isotec Inc.

As judged by SDS-PAGE the {^{15}N,^{2}H} enriched Fyn SH3 sample was pure. The molecular weight of 8919.0 daltons, experimentally determined using electrospray mass spectroscopy, was consistent with that calculated from the amino acid sequence assuming complete ^{15}N and ^{2}H labelling (8914.6 daltons).

NMR methods

For NMR the {^{15}N,^{2}H} labelled SH3 sample was dissolved to 1 mM, pH 6 in 95%/5% H_2O/D_2O and the solution transferred to a Shigemi NMR tube (Shigemi Co., Ltd.).

NMR spectra were obtained on a home built spectrometer interfaced to an Oxford Instruments 750 MHz superconducting magnet. A homemade RF probehead (Soffe et al., 1995) with a 3 axis gradient set was used. All NMR spectra were obtained using a probehead temperature of 283 K.

Thirteen 2D NOESY spectra were collected with variable mixing times (25 ms, 100 ms, 200 ms, 300 ms, 400 ms, 600 ms, 800 ms, 1 s, 1.4 s, 2 s, 2.6 s, 3.2 s, 3.8 s) employing 128 and 2048 complex points with dwell times of 250 μs and 80 μs for t_1 and t_2, respectively, four scans/point, a recycle delay of 15 s and the pulse scheme of Huang and Shaka (Huang and Shaka, 1995) followed the longitudinal relaxation period. Five 3D HSQC-NOESY-HSQC spectra were collected with variable mixing times (10 ms, 300 ms, 600 ms, 900 ms, 1.2 s) employing 32, 32 and 1024 complex points with dwell times of 500 μs, 500 μs and 80 μs for t_1(^{15}N), t_2 (^{15}N) and t_3 (^{1}H) respectively, four scans/point, a recycle delay of 4.4 s and using gradient coherence selection with sensitivity enhancement (Kay et al., 1992) following the t_2 evolution period. 2D ^{15}N-^{1}H HSQC spectra (Bodenhausen and Ruben, 1980) were collected employing 256 and 4096 complex points with dwell times of 250 μs and 80 μs for t_1 and t_2, respectively, 64 scans/point, a recycle delay of 5 s and using gradient coherence selection with sensitivity enhancement (Kay et al., 1992).

The NMR data were processed with FELIX (FELIX2.3, Biosym Technologies, Inc.) and spectral analysis was carried out using the program XEASY (Bartels et al., 1995).

Structure calculations

The structure calculations reported here were done within the programme X-PLORv. 3.1 (Brünger, 1992) employing the *ab initio* simulated annealing protocol. The calculation starts from an extended structure with ideal covalent geometry and randomised side chains with sum averaging applied to the NOE constraints. The energy mimimisation routine employed 'floating chirality', in which prochiral groups without stereospecific assignment, such as valine and leucine methyl groups and methylene hydrogens, were allowed to interchange positions during the course of the calculation. In each case the 50 structures were further refined by an additional cycle of simulated annealing with an extended cooling phase. The refined structures were analysed using the programme MOLMOL (Koradi et al., 1996).

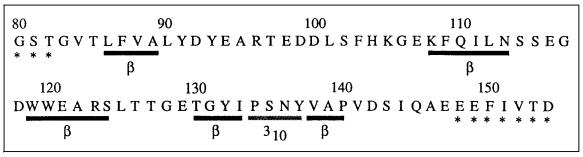

Figure 1. The amino acid sequence of the Fyn SH3 domain employed in this study. The position and identity of each of the 5 β strands and the 3_{10} helix are shown beneath the sequence. An asterisk below a residue indicates that no ^{1}H or ^{15}N assignment was available.

Simulations

A set of Six NMR structures were used for some of the calulations and simulations reported here and these were the six lowest energy structures from the family displayed in Figure 12[4]

Results and Discussion

The amino acid sequence of the Fyn SH3 construct used is shown in Figure 1. This sequence comprises residues 82–148 of the SH3 domain from human Fyn tyrosine kinase plus a two-residue N-terminal extension (GS) and a six-residue C-terminal extension (EFIVTD). The X-ray data (Noble et al., 1993) have provided structural information for residues 84–142 and the previous NMR study (Morton et al., 1996) provided structural information and assignments for residues 85–141.

^{1}H and ^{15}N long-range isotope shifts

Figure 2 shows a 2D ^{15}N-^{1}H HSQC spectrum recorded from a mixture of separately prepared {^{15}N} and {^{15}N,^{2}H} enriched Fyn SH3 samples. This spectrum clearly shows the long-range isotope shifts for the amide resonances, both ^{1}H and ^{15}N, which occur upon deuteration. The resonances from the deuterated sample are easily identified as those with the much decreased ^{1}H linewidth; this arises from the fact that the ^{3}J(^{1}H,^{2}H) vicinal coupling constants are greatly reduced due to the rapid relaxation of the ^{2}H. We have assumed that all the factors contributing to the observed chemical shifts retain the same magnitude in

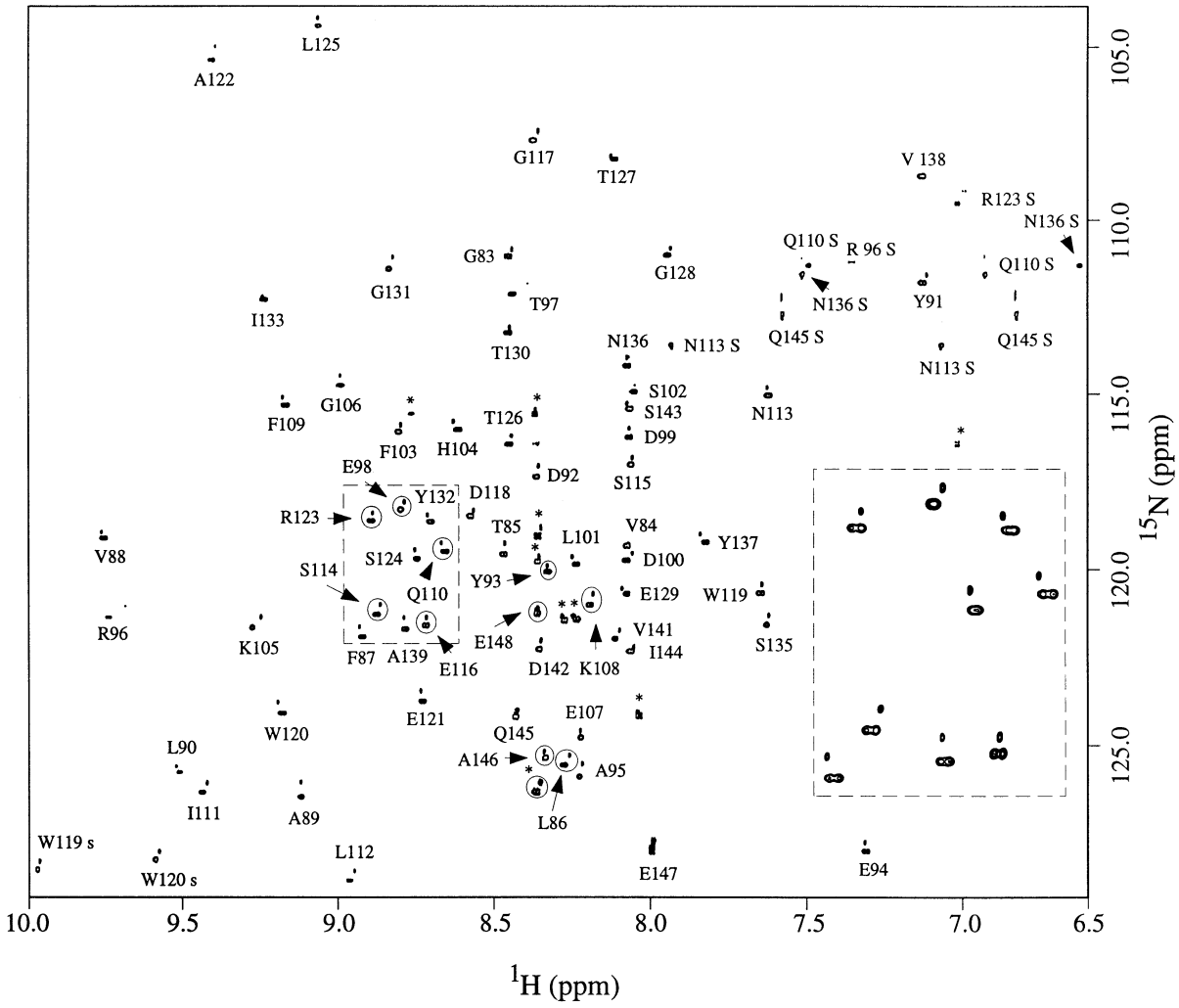

Figure 2. A 2D ^{15}N–^{1}H HSQC spectrum from a mixture of {^{15}N} and {^{15}N,^{2}H} enriched Fyn SH3 samples recorded at 750 MHz. The inset in the right corner shows an expansion of the region enclosed by the dashed lines.

each species and that any differences in chemical shift may be attributed solely to isotope effects. Collecting NMR data from a mixture of labelled species overcomes the difficulty of making two separate samples with identical pH and ionic composition

In all cases the $^{n}\Delta^{15}$N(^{1}H,^{2}H) long-range isotope shift is observed to be towards lower frequency (Figure 3a). The maximum observed isotope shift is 0.42 ppm for the ^{15}N resonance of A89. Whereas, for the ^{1}H long-range isotope shifts, $^{n}\Delta^{1}$H(^{1}H,^{2}H) Figure 3b, both low and high frequency shifts are observed. The maximum and minimum observed ^{1}H isotope shifts are 5.24×10^{-2} ppm for R96 and -1.73×10^{-2} ppm for Y137, respectively. It has been proposed (Jameson, 1996) that long-range isotope shifts are primarily a measure of the sensitivity of the shielding of the nucleus, $\sigma(^{1}H_i)$ or $\sigma(^{15}N_i)$, to a change in length at the bond which incorporates the isotopic substitution. Here these substitutions will be at the C-^{2}H bonds. This contribution to the long-range isotope shift is mediated through-bond rather than through-space and is an electronic property of the system. Therefore, the long-range shifts may correlate with factors contributing to the ^{3}J(HN,Hα) vicinal coupling constants. A correlation of the $^{n}\Delta^{15}$N(^{1}H,^{2}H) isotope shifts with the dihedral angles ϕ and ψ can be established (Figure 3c). However, no statistically significant correlation of the $^{n}\Delta^{1}$H(^{1}H,^{2}H) isotope shifts was established. Interestingly the two resonances showing the largest ^{1}H isotope shift, R96 and T97, are associated

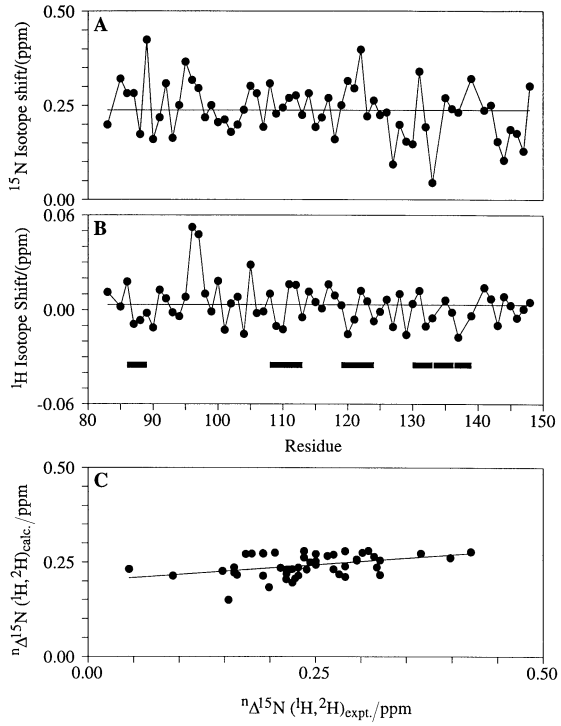

Figure 3. (a) The ^{15}N long-range isotope shifts versus residue number. The average shift of 0.237 ppm is shown by the solid line. (b) The ^{1}H long-range isotope shifts as a function of residue number. The average shift of 3.47×10^{-3} ppm is shown by the solid line. The residues involved in secondary structure elements are shown by black bars. (c) A plot of the calculated and predicted $^{n}\Delta^{15}$N(^{1}H,^{2}H) isotope shifts. The calculated data were derived from a best fit to an equation of the form (Ottiger and Bax, 1997) $^{n}\Delta^{15}\mathrm{N}(^{1}\mathrm{H},^{2}\mathrm{H})_{(i)} = 0.191 + 0.065\sin(\phi_i + 156) - 0.035\sin(\psi_i + 61)$ ppm where ϕ_i and ψ_i are the dihedral angles, which were obtained using the X-ray structure. Fifty data points were used (residues 85–141 excluding all G, P and V138) and the correlation coefficient is 0.32. For this number of data points, statistical significance at the 97.5% confidence level requires a correlation coefficient > 0.28 (Devore, 1991).

with much increased ^{1}H linewidths suggesting that the isotope shifts can be influenced by factors such as chemical exchange. We did not observe a correlation of long-range isotope shift with amino acid type.

^{1}H amide transverse relaxation rates

There are several relaxation mechanisms expected to contribute to amide ^{1}H transverse relaxation rates, $1/T_2$ ($= R_2$), in an aqueous sample of a {^{15}N,^{2}H} enriched protein. These are dipole–dipole interactions amongst all the solvent exchangeable ^{1}H–^{1}H (R_2(H$_i$–H$_j$)) and ^{1}H-^{15}N (R_2(H$_i$–^{15}N$_i$)) spin pairs, anisotropic chemical shift of the ^{1}H amide nuclei (R_2(H$_i$ CSA)) and a contribution from dipole–dipole relaxation amongst the ^{1}H–^{2}H spin pairs (R_2(H$_i$–^{2}H$_j$)). For a sample at pH 6.0 and 283 K, intermolecular chemical exchange with solvent may contribute significantly to relaxation of those amide protons which are: (i) exposed to the solvent; (ii) have appropriate near neighbours (Bai et al., 1993); and (iii) are not involved in long-lived hydrogen bonds in elements of secondary structure. A further contribution to ^{1}H line broadening may arise from exchange between several well populated sites at rates that only partially average any ^{1}H chemical shift differences.

The ^{1}H transverse relaxation rate of an individual amide, R_2(H$_i$), can be expressed as the sum of the contributions from these individual relaxation processes. For this direct summation (Equation 1), ^{1}H-^{15}N decoupling is assumed. Various possible cross correlations, transverse cross-relaxation from partially overlapped resonances (Anet and O'Leary, 1990), intramolecular chemical exchange, the influence of any remaining unresolved ^{n}J(^{1}H,^{2}H) scalar couplings via scalar relaxation and the dynamic frequency shift (Werbelow, 1996; Tjandra et al., 1996) have been neglected.

$$\begin{aligned} R_2(\mathrm{H}_i) \;=\; & R_2(\mathrm{H}_i - \mathrm{H}_j) + R_2(\mathrm{H}_i - {}^{2}\mathrm{H}_j) \\ & + R_2(\mathrm{H}_i - {}^{15}\mathrm{N}_i) + R_2(\mathrm{H}_i\mathrm{CSA}) \\ & + k_{\mathrm{ex}} + R_2^{*} \end{aligned} \tag{1}$$

k_{ex} is a pseudo-first-order exchange rate with the solvent; R_2^{*} represents a contribution to the lineshape from B_0 inhomogeneities and incomplete ^{15}N decoupling (Shaka et al., 1983). The individual relaxation terms can be defined in terms of spectral densities, Equations 2–5 (Abragam, 1960; McConnell, 1987),

$$\begin{aligned} R_2(\mathrm{H}_i - \mathrm{H}_j) \;=\; & \left\{(6\pi/5)(\mu_0/4\pi)^2\gamma_{\mathrm{H}}^4\hbar^2\right\} \\ & \{5/6J(0) + 3/2J(\omega_{\mathrm{H}}) \\ & + J(2\omega_{\mathrm{H}})\}\sum_{i\neq j} 1/r_{ij}^6 \end{aligned} \tag{2}$$

$$\begin{aligned} R_2(\mathrm{H}_i - {}^{2}\mathrm{H}_j) \;=\; & \left\{(16\pi/5)(\mu_0/4\pi)^2\gamma_{\mathrm{H}}^2\gamma_{\mathrm{D}}^2\hbar^2\right\} \\ & \{2/3J(0) + 1/6J(\omega_{\mathrm{H}} - \omega_D) \\ & + J(\omega_{\mathrm{D}}) + 0.5J(\omega_{\mathrm{H}}) + J(\omega_{\mathrm{H}} \\ & + \omega_{\mathrm{D}})\}\sum_{j} 1/r_{ij}^6 \end{aligned} \tag{3}$$

$$R_2(\mathrm{H}_i - {}^{15}\mathrm{N}_i) = \left\{(6\pi/5)(\mu_0/4\pi)^2\gamma_\mathrm{H}^2\gamma_\mathrm{N}^2 \hbar^2/r_\mathrm{NH}^6\right\}\{2/3J(0) + 1/6J(\omega_\mathrm{H}-\omega_\mathrm{N}) + J(\omega_\mathrm{N}) + 0.5J(\omega_\mathrm{H}) + J(\omega_\mathrm{H}+\omega_\mathrm{N})\} \quad (4)$$

$$R_2(\mathrm{H}_i\ \mathrm{CSA}) = (8\pi/15)\gamma_\mathrm{H}^2 B_0^2 \Delta\sigma(H_i)^2 \{2/3J(0) + 0.5J(\omega_\mathrm{H})\}. \quad (5)$$

In the equations above, and for all the following calculations and simulations, we neglect any internal motions and assume a rigid protein undergoing isotropic motion. In this case the real part of the spectral density is $J(a\omega) = (1/4\pi)\ \tau_r/(1+(a\omega\tau_r)^2)$; Equation 5 assumes axial symmetry for the ^{1}H chemical shift tensor.

The summations in Equations 2 and 3 were performed using coordinates from the six NMR structures (see methods), assuming that backbone amide groups and residues with solvent exchangeable atoms in their side chains (T,S,Y,R,K,N,Q,W,H) were the only atoms protonated (there are no cysteines in Fyn SH3). The side chains of aspartic and glutamic acids were assumed to be ionised at pH 6. The averaged transverse relaxation rate, $R_2(\mathrm{H}_i)$, from the six structures, assuming similar motional properties for each residue, is shown in Figure 4a. An average value for $R_2(\mathrm{H}_i)$, taken over the six NMR structures, reflects variations of internuclear separation and the possible extent of intraresidue motion. The overall mean transverse relaxation rate is 17.4 s^{-1} with a maximum value of 30.1 s^{-1} for S114 and a minimum value of 10.9 s^{-1} for V141. The relative magnitudes of the relaxation contributions were $R_2(\mathrm{H}_i{-}{}^{15}\mathrm{N}_i) \geq R_2(\mathrm{H}_i\text{-}\mathrm{H}_j) >> R_2(\mathrm{H}_i{-}{}^{2}\mathrm{H}_j) = R_2(\mathrm{H}_i\ \mathrm{CSA})$ for those residues with calculated relaxation rates $< 16\ s^{-1}$. Whereas for the remainder, with a relaxation rate $> 16\ s^{-1}$, $R_2(\mathrm{H}_i{-}\mathrm{H}_j) > R_2(\mathrm{H}_i{-}{}^{15}\mathrm{N}_i)$. For the majority of amide protons with a near neighbour (< 3.0 Å) the summation in Equation 2 is found to reach >95% of its maximum value when using a cut-off distance of 6.8 Å and for those protons with more than one near neighbour the 95% level is attained with a shorter cut-off distance than 6.8 Å. This point has implications for ^{1}H–^{1}H NOE measurements and we return to it below. The frequency dependence of these calculations is small

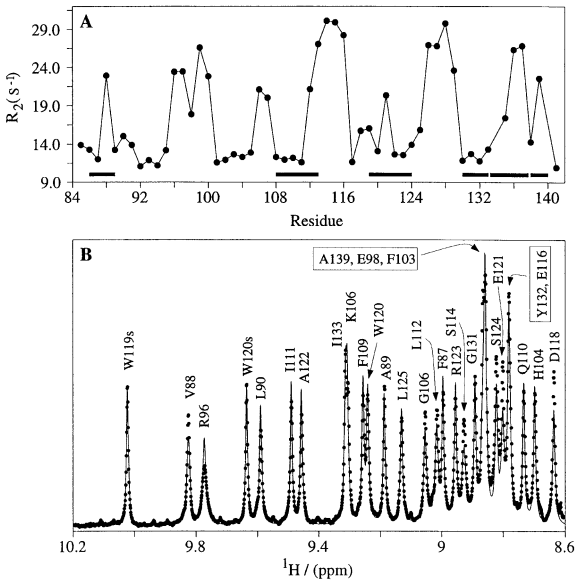

Figure 4. (a) The calculated ^{1}H amide transverse relaxation rate, using Equation 1, with $k_{ex} = R_2{}^* = 0$, as a function of residue number. The following parameters were used, for all residues, in the calculation $r_\mathrm{nh} = 0.102$ nm, $B_0 = 17.6$ T, $\tau_r = 6.6$ ns and $\Delta\sigma(\mathrm{H}) = 10$ ppm. (b) Part of the fully relaxed 1D ^{1}H 750 MHz spectrum from the {^{15}N,^{2}H} sample recorded with ^{15}N broadband decoupling and using a relaxation delay of 15 s. The solid line represents a simulated 1D spectrum, calculated using the ^{1}H amide transverse relaxation rates shown in Figure 4a, with $R_2{}^* = 2$ Hz and assuming each resonance was a singlet. Both the simulated and experimental spectra used a dwell time of 80 μs and were zero-filled once prior to Fourier transformation to give a total of 16384 complex points.

and the ratio of $R_2(\mathrm{H}_i$ 750 MHz)/$R_2(\mathrm{H}_i$ 500 MHz) is only about 0.98.

The dipolar interaction, represented by $R_2(\mathrm{H}_i{-}\mathrm{H}_j)$, typically contributes between 10–70% of the total ^{1}H amide transverse relaxation rate. This means that in those cases where $R_2(\mathrm{H}_i{-}\mathrm{H}_j) < R_2(\mathrm{H}_i{-}{}^{15}\mathrm{N}_i)$ the transverse relaxation rates are not particularly sensitive to the global geometry and in these cases have a significant contribution from purely local factors. A simulated ^{1}H 1D spectrum, which uses the calculated transverse relaxation rates, is compared to the experimental 750 MHz ^{1}H 1D spectrum in Figure 4b. This simulation takes no account of differential internuclear motion. However, for those resolved resonances where a direct comparison is possible, the agreement is reasonably good. Only relatively small changes to the calculated amide transverse relaxation rates would be needed to give a significant improvement of the agreement for individual resonances. Assuming that these changes would result from the relaxation term $R_2(\mathrm{H}_i{-}\mathrm{H}_j)$, they could easily occur through local motion or

because the set of NMR structures used for the calculations is not quite representative of local geometry. Relatively small changes to the internuclear distances, of about 10%, would be sufficient to improve the fits significantly. The rest of the 1D spectrum is not simulated because of the dual problems of unassigned resonances and because some of the amide resonances have the same chemical shift as broad resonances from side chains.

^1H *amide longitudinal relaxation rates*

The longitudinal relaxation of a system of dipolar coupled spins, assuming no dipolar cross correlations or chemical exchange, is governed by the system of coupled differential equations, Equation 7, (Solomon, 1955; Abragam, 1961)

$$\mathrm{d}\left[\mathbf{I}_z^i(t)\right]/\mathrm{d}t = -\rho_i(\mathbf{I}_z^i(t) - \mathbf{I}_z^i(\infty)) - \sum_{i\neq j}\sigma_{ij}(\mathbf{I}_z^j(t) - \mathbf{I}_z^j(\infty)) \qquad (7)$$

where $\mathbf{I}_z^i$(t) and $\mathbf{I}_z^j$(t) are the expectation values of the Zeeman operators $\mathbf{I}_z^i$ and $\mathbf{I}_z^j$; $\mathbf{I}_z(\infty)$ refers to thermal equilibrium, ρ_i is the self-relaxation rate of spin i and σ_{ij} is the cross-relaxation rate. Both ρ_i and σ_{ij} can be expressed in terms of spectral densities, Equations 8 and 9,

$$\rho_i = \left\{(6\pi/5)(\mu_0/4\pi)^2\gamma_{\mathrm{H}}^4\hbar^2\right\}\{1/3J(0) + J(\omega_{\mathrm{H}}) + 2J(2\omega_{\mathrm{H}})\}\sum_{i\neq j}1/r_{ij}^6 \qquad (8)$$

$$\sigma_{ij} = \left\{(6\pi/5)(\mu_0/4\pi)^2\gamma_{\mathrm{H}}^4\hbar^2\right\}\{-1/3J(0) + 2J(2\omega_{\mathrm{H}})\}\,1/r_{ij}^6. \qquad (9)$$

Employing a similar approach to that used for the calculation of the ^{1}H amide transverse relaxation rates, and with the same coordinate sets, the average calculated value of ρ_i for the amide protons is shown in Figure 5a. Additional small contributions to the self-relaxation rate, ρ_i, of each amide have been added from the dipolar ^{1}H–^{15}N interaction, from ^{1}H chemical shift anisotropy and from a summation over the ^{1}H–^{2}H spin pairs. In contrast to the ^{1}H amide transverse relaxation rates, the dipolar interactions between amide protons now contributes approximately 95% to the total amide self-relaxation rate. It is therefore expected that both the self-relaxation rate and the cross-relaxation rate will be sensitive to global geometry, with 4–10 neighbouring protons contributing significantly to the longitudinal dipolar relaxation. Also shown in Figure 5a is a value, averaged over the six coordinate sets, for the cross-relaxation rate, σ_{ij}; this calculation used only the nearest neighbour to the amide proton from each coordinate set to perform the average. A summation of these two quantities is also shown. Some experimental data, recorded using a non-selective inversion recovery sequence, are shown in Figure 5b. These examples are from amide protons known to be involved in long-lived secondary structure. Consequently there is negligible exchange between the amide proton and the solvent over the duration of these experiments. These decay curves are expected to be non-exponential (Campbell and Freeman, 1973) and the solid lines are shown only to aid the visualisation of each dataset. The decay of ^{1}H longitudinal magnetisation for these resonances is seen to be very slow compared to a fully protonated protein. The longitudinal relaxation times can be further increased simply by changing the solvent composition. In a 40%/60% H_2O/D_2O solvent mixture full longitudinal recovery for these protons would not be complete until after about 20 s (Figure 5b), and with this solvent composition the relative importance of the ^{1}H–^{1}H dipolar interactions is reduced compared to the other relaxation mechanisms.

For a fully protonated protein, cross-relaxation between protons is an important mechanism for longitudinal relaxation and this eventually leads to similar relaxation rates being observed throughout the sample (Kalk and Berendsen, 1976). In contrast, the data from the non-selective inversion recovery experiment (95%/5% H_2O/D_2O) reported here show a wide range of ^{1}H relaxation rates. Complete proton recovery was observed to fall within the range of 2 s to 15 s. This wide range of longitudinal relaxation rates is consistent with slower relaxation processes, which, because of fewer spins at increased distances, do not give sufficient time during the relaxation period for an equalisation of these rates. Consequently the longitudinal relaxation of an amide proton is expected to be dominated by the distance to the near neighbours and their relative motion. Before presenting data from homonuclear 2D NOESY experiments in support of this proposal we highlight how chemical exchange processes with the solvent can significantly reduce the longitudinal relaxation time of some amide ^{1}H resonances and demonstrate how to increase this time.

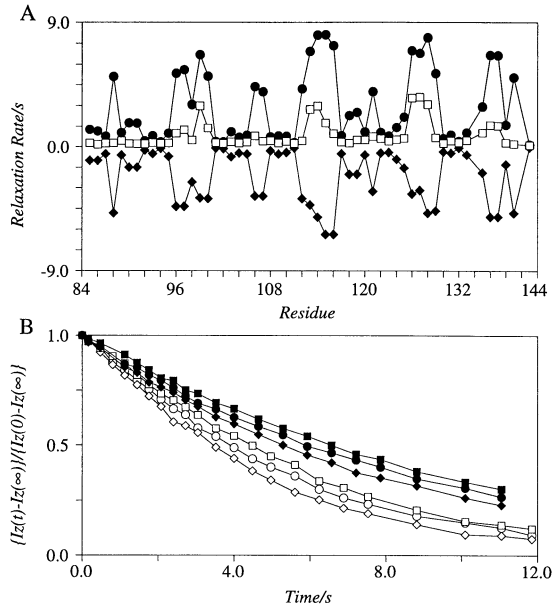

Figure 5. (a) The calculated ^{1}H amide self relaxation rate, ρ_i, as a function of residue number, ●. A calculated value for the ^{1}H–^{1}H cross-relaxation rate, σ_{ij}, as a function of residue number, ♦. A summation of ρ_i and σ_{ij} as a function of residue number, □. The calculations used the same parameters as described in the legend to Figure 4a. (b) Non-selective inversion-recovery datasets for the ^{1}H amide resonances of residues V88, □, A89, ○ and R123, ◇, in 95%/5% H_2O/D_2O and in 40%/60% H_2O/D_2O filled points. The pulse sequence used for these measurements is scheme 2 of Figure 6a. The ^{1}H amide longitudinal magnetisations I_z(t) and $I_z(\infty)$ were acquired on alternate scans and subtracted so that the difference longitudinal magnetisation, I_z(t) – $I_z(\infty)$, was collected directly (Freeman and Hill, 1971). A total of 32 accumulations were collected with a recycle delay of 38s for each datapoint.

^{1}*H amide direct and indirect interactions with the solvent*

The longitudinal relaxation properties of the solvent can have two extremes. For the first case, radiation damping is dominant and the longitudinal recovery, after the solvent magnetisation is inverted, has a time constant of tens of ms, with the precise value depending on various instrumental parameters (Abragam, 1961). In the second case, when radiation damping is suppressed (Sklenar, 1995), the longitudinal relaxation time of the solvent is significantly increased. Measurement of the water resonance in the {^{15}N,^{2}H} sample gave a longitudinal relaxation time of 1.97 s (283 K, pH 6) when radiation damping was suppressed. In contrast, when radiation damping was active, longitudinal recovery of the solvent was complete 60–70 ms after inversion.

In the next few paragraphs we demonstrate widely differing longitudinal relaxation behaviour for those amide protons which are undergoing direct chemical exchange and/or are close enough for cross-relaxation to a proton, undergoing rapid exchange with the solvent. Specifically, we consider the influence on ^{1}H longitudinal relaxation of the situation where the solvent magnetisation is either $+Z$ or $-Z$ at the beginning of the relaxation period, with radiation damping suppressed.

Three inversion recovery schemes, used for the longitudinal 1D measurements, are shown in Figure 6a. In schemes 1 and 2, either the solvent resonance or amide resonances are selectively inverted, respectively, and scheme 3 is a non-selective inversion recovery experiment. The principles of these experiments in a system with chemical exchange have been described previously (Forsen and Hoffman, 1963; Campbell et al., 1978).

No resonances which could be assigned to hydroxyl protons from serine, threonine or tyrosine residues have been identified in the {^{15}N,^{2}H} sample. We make the assumption, therefore, that intermolecular exchange of these hydroxyl protons with the solvent is rapid, leading to a broad ^{1}H resonance in each case. For a sample at pH 6 and 283 K there are data (Liepinsh et al., 1992) to support this assumption, especially as the sample also contains the buffer Tris-HCl, which catalyses hydroxyl/water proton exchange (Englander and Kallenbach, 1983).

The intrinsic amide proton exchange rates of the {^{15}N,^{2}H} sample (283 K, pH 6) were estimated using published data (Bai et al., 1993). In most cases the predicted intrinsic amide proton exchange rate constants are < 1 s^{-1}. An exception is formed by the pair of residues S114 and S115, which are part of a loop, where the predicted amide poton exchange rates are > 4s^{-1}. It was not possible to significantly reduce the amide proton exchange rates by decreasing the pH because the sample is not stable below pH 6.

A series of difference spectra is shown in Figure 6b; these display the influence of transient NOE/exchange with the solvent upon the amide resonances using pulse scheme 1 in Figure 6a. The difference spectra are displayed as negative peaks to indicate that in all cases they are consistent either with a direct exchange or with an indirect transient negative NOE via some intermediate such as a hydroxyl proton (Spera et. al., 1991). No positive peaks, which might be interpreted as an NOE between the protein and a water molecule (Otting et al., 1991), were seen

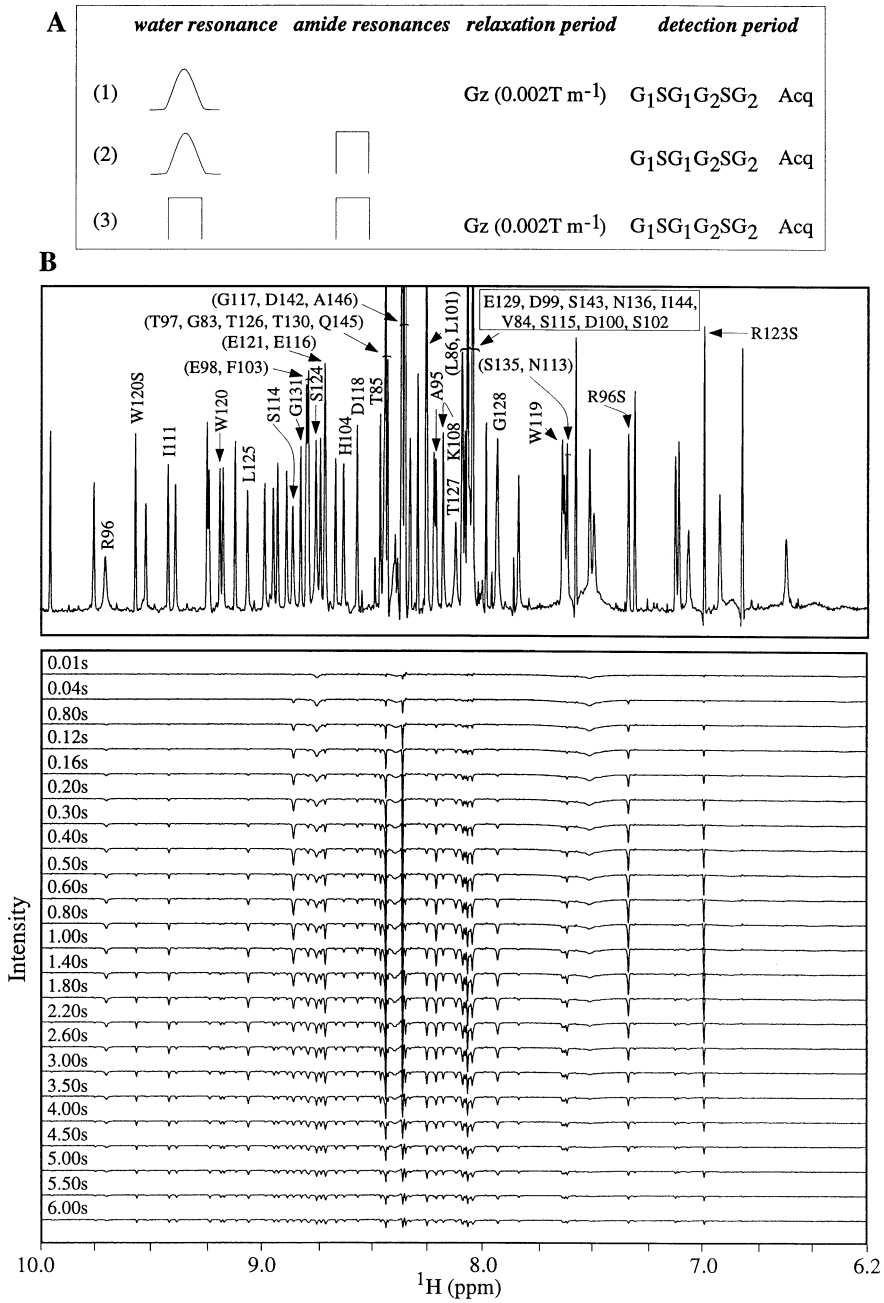

Figure 6. (a) The 3 pulse sequences used to monitor the longitudinal relaxation properties of the amide ^{1}H resonances. The first two columns indicate which resonances are to be inverted prior to the longitudinal relaxation period. ⩘ represents a selective 180°, 2 ms, Gaussian-shaped RF waveform centred at the water frequency, and ⊓ is a 180° non-selective RF pulse which is assumed to affect the water and amide resonances identically. The third column indicates the longitudinal relaxation period, which immediately follows the RF pulses and the fourth column indicates the sequence of RF pulses applied after the longitudinal relaxation period prior to data collection. For the inversion recovery sequence shown in scheme 3 the non-selective RF pulse is shown twice purely to indicate that both the water and amide resonances are inverted. Radiation damping of the water magnetisation was suppressed, in those experiments where the water magnetisation can be described as $-\mathbf{I}_z$ just prior to the longitudinal relaxation period, by applying a weak single axis field gradient, Gz (0.002 Tm^{-1}) (Sklenar 1995). $G_1SG_1G_2SG_2$ is the double pulse field gradient spin echo sequence of Huang and Shaka (Huang and Shaka, 1995) and included CYCLOPS (Hoult and Richards, 1975) phase cycling. The acquisition period, Acq, includes broadband ^{15}N waltz16 decoupling. The *Z* magnetisation of the solvent after the RF pulses and just prior to the longitudinal relaxation period was measured to be $-0.95M_0$, $0.93M_0$ and $-0.95M_0$ for pulse sequences 1, 2 and 3 respectively. (b) A time series showing the amide proton resonances as a function of the longitudinal relaxation period recorded using pulse sequence 1 of Figure 6a. The resolution-enhanced 1D spectrum is also shown with the annotation indicating the assignment of the resonances appearing in this time series (a S following a residue label indicates a resonance from a side chain). The ^{1}H amide longitudinal magnetisations were acquired on alternate scans either with or without selective inversion of the water resonance and subtracted. A total of 160 accumulations were collected for each dataset with a recycle delay of 15 s and the duration of the longitudinal relaxation period is indicated.

during the time course. The ^{1}H resonance assigned to S114 at 8.93ppm can be seen to have a large perturbation. These data also show that many resonances are unperturbed, even with mixing times of several seconds.

The amide protons predicted to be within 5.0 Å of a hydroxyl proton, from serine, threonine and tyrosine, are indicated in Figure 7a. An improvement to the resolution is afforded by recording 2D ^{15}N–^{1}H HSQC spectra at different longitudinal relaxation periods after selective inversion of the solvent (Grzesiek and Bax, 1993; Wang et al., 1996). Two 2D ^{15}N–^{1}H HSQC datasets are shown in Figure 7b. Both S114 and S115 give quite intense peaks after only a short longitudinal relaxation delay. All of the assigned peaks in these spectra, with two exceptions (H104 and R123), are either from amides contained within loops where the protons are presumably exposed to the solvent, or are from amide protons predicted to be within 5.0 Å of a hydroxyl proton, from which significant cross-relaxation could occur. For H104 $r_{\mathrm{NH,OH}}$ is, on average, calculated to be > 5 Å in 86% of the structures and a resonance was observed, whereas, for R123, $r_{\mathrm{NH,OH}}$ was calculated to be < 5 Å in 57% of the structures and no resonance was observed. For this analysis, we have only considered those amide resonances which have NMR or X-ray coordinates, which excludes some assigned resonances from the N- and C-termini which give rise to peaks in these spectra.

Figures 8a and 8b show inversion recovery time courses for the amide protons of residues L125 and S114 as a function of the three separate pulse schemes. These data show that amide longitudinal recovery is slowed when both sets of resonances are inverted and solvent radiation damping is supressed. Amide longitudinal magnetisation recovery is faster if the solvent magnetisation is aligned along $+Z$ during the time course. In each case the solid lines are simulations which used different initial conditions and assumed similar motional properties for all the nuclei involved. For the simulations a longitudinal relaxation matrix was constructed employing five of the nearest proton neighbours and the solvent. For S114 the five near neighbours are the amide protons of N113, S115 and E116 with the side chain hydroxyl protons of S114 and S115. For L125 the five neighbours are the amide protons of S124, T126, and T127 with the side chain hydroxyl protons of S124 and T126. In most of the six NMR structures these near neighbours were found to make the primary contribution to ^{1}H-^{1}H dipolar relaxation. When averaged over the six NMR and X-ray coordinate sets these five near-neighbour protons are calculated to contribute about 57% and 88% to the total ^{1}H–^{1}H dipolar interactions of S114 and L125, respectively. For the simulations only the X-ray structure was used to generate the inter-proton distances for the calculation of the self- and cross-relaxation rates, via Equations 8 and 9. The other parameters used in the simulations are shown in the legend to Figure 8. It was found that if the value used in the simulations for the hydroxyl/proton exchange rate, k_{oh}, was $> 25\ \mathrm{s}^{-1}$, the results were independent of whether the particular hydroxyl resonance is considered to have a chemical shift in the amide region or at the solvent frequency. Simply by changing the initial conditions of the simulations, the same parameters for either L125 or S114 can simultaneously provide the fit for each experimental dataset. A rather large number of parameters are required for these simulations, some of which have not been experimentally measured, so we do not propose that the precise values are important, rather that the three separate datasets can be modelled satisfactorily in each case.

Those amide protons, from residues such as V88 or A89 which do not exchange with the solvent and are further than 5.0 Å from a hydroxyl group show no significant difference in longitudinal relaxation behaviour when using either pulse scheme 2 or 3 of Figure 6a (data not shown). We have thus demonstrated that the amide ^{1}H longitudinal relaxation rate can be very sensitive to the relaxation properties of the solvent.

2D NOESY spectra recorded with variable mixing time

The data in Figure 9 illustrate the effect of varying the mixing time on the time dependence of the transient inter-amide NOE effects for two residues, V88 and A89. The mixing times range from 25ms to 3.8s, leading to identification of six and seven cross peaks for V88 and A89, respectively. The numbers in brackets represent the distance, calculated from the X-ray structure after the addition of hydrogens (using X-PLOR), from the amide proton of either V88 or A89 to the particular amide proton giving rise to the cross peak. The observed cross peaks are localised to the amides within a distance of about 6.1 Å. Even at the long mixing times these spectra retain the specificity of the local environment. Employing the NMR and X-ray coordinate sets shows that the indicated cross peaks for V88 and A89 account for 99% and 93% of the total ^{1}H–^{1}H dipolar interactions, respectively.

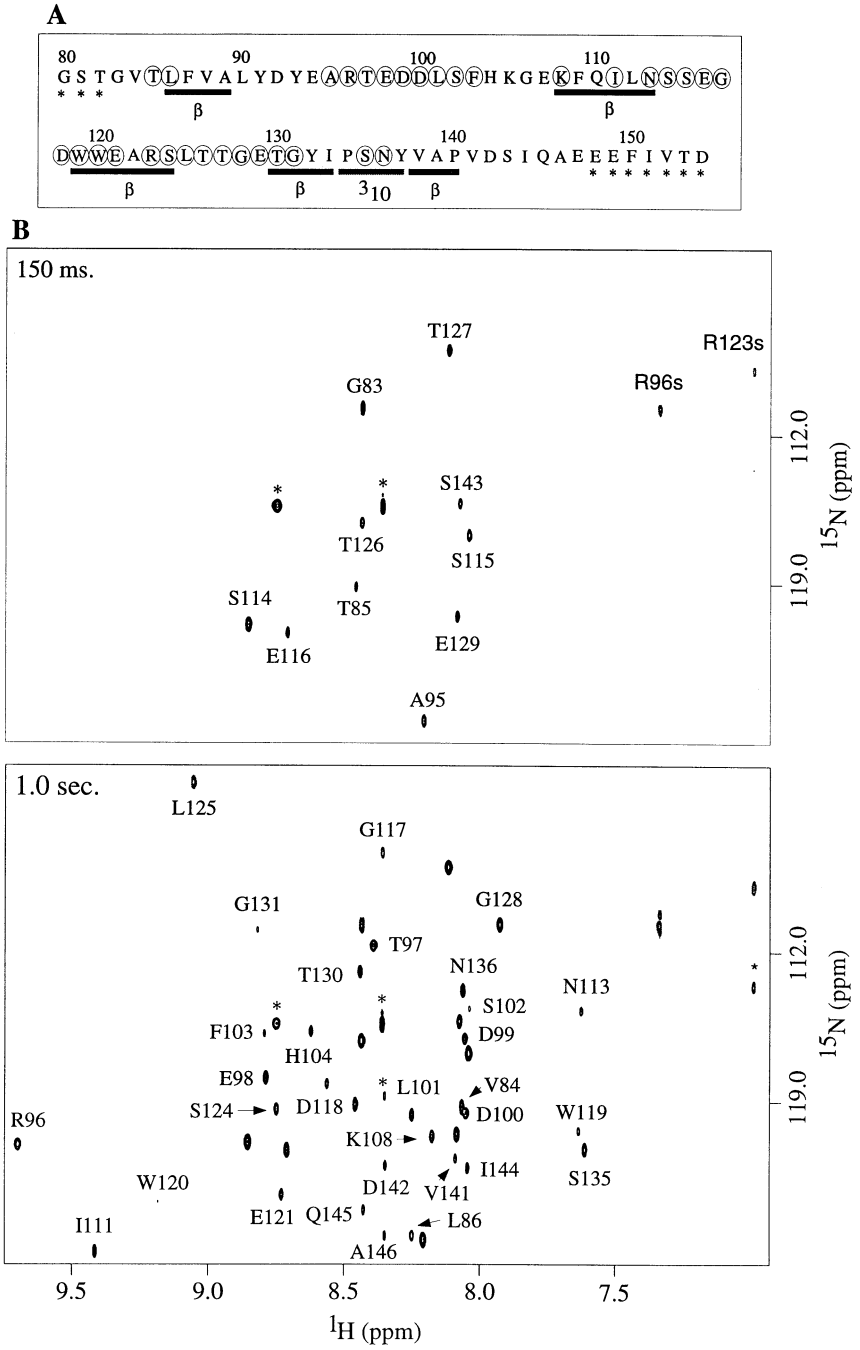

Figure 7. (a) The circles indicate the amide protons which are predicted to be within 5.0 Å of the hydroxyl proton of a serine, threonine or tyrosine residue. In this particular case the distances were estimated using all the NMR structures shown in Figure 12[4] and the X-ray structure, after the addition of hydrogens. For a residue to be included the distance to a hydroxyl proton had to be less than 5.0 Å in more than 50% of these structures. (b) The data were recorded as difference spectra. The selective inversion of the solvent was achieved using the double pulse field gradient spin echo sequence of Huang and Shaka (Huang and Shaka, 1995) in which the first selective pulse was phase cycled using EXORCYCLE (Stott et al., 1997), followed by a relaxation period in which radiation damping was suppressed (Sklenar, 1995) and a HSQC employing gradient coherence selection.

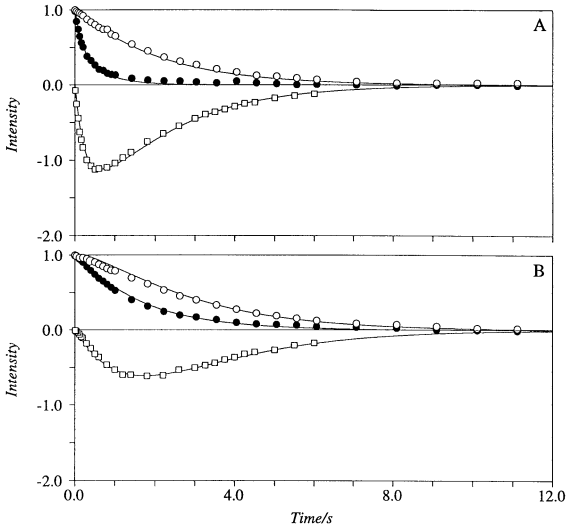

Figure 8. (a) Inversion recovery measurements for the amide proton of S114. □, selective inversion of the solvent using scheme 1, ●, selective inversion of the amides using scheme 2, ○, non-selective inversion using scheme 3. The solid lines are simulations to the experimental data which used the following parameters $r_{NH} = 0.102$ nm, $B_0 = 17.6$ T, $\tau_r = 6.6$ ns, $\Delta\sigma(H) = 10$ ppm, $\rho_{hoh} = 0.507s^{-1}$, $\rho_{oh} = 0.5s^{-1}$, $k_{oh} = 35s^{-1}$, $k_{114} = 2s^{-1}$, $k_{115} = 1s^{-1}$, $k_{116} = 0.2s^{-1}$ and $k_{113} = 0.4s^{-1}$. For this simulation and that of Figure 8b the longitudinal relaxation matrix was not symmetric; however, all the eigenvalues and eigenvectors were real in both cases. In order to improve the agreement between the simulated and experimental data it was necessary to change the self relaxation rates from the values calculated using Equation 8. (b) Inversion recovery measurements for the amide proton of L125. The labelling of the separate datasets is similar to that above. The solid lines are simulations to the experimental data with parameters similar to above except, $k_{125} = k_{126} = k_{127} = 0.1s^{-1}$ and $k_{124} = 0s^{-1}$.

The two sets of spectra are a representative selection of the complete dataset. From other cross sections a few cross peaks were noted to occur between amides with predicted separations up to about 7.0 Å. This is consistent with the summations used to generate the data for Figure 5, where it was noted that for the majority of the amide protons the summations had reached at least 95% of their maximum value when using a cut-off distance of 6.8 Å. Cross peaks are not expected between amides with separations much greater than 7.0 Å. From these and other cross sections it is clear that, when using a mixing time of about 300 ms, the strong intensity cross peaks are from neighbours with $r_{ij} < 3$ Å. Medium intensity cross peaks, corresponding to distances in the range 3 Å $< r_{ij} < 4$ Å, become apparent with mixing times of about 600ms and weak cross peaks, corresponding to distances in the range 4 Å $< r_{ij} < 7$ Å, are detectable at mixing times of about 1.2 s.

Figure 10 shows the diagonal and cross peak intensities, for some of the resonances of Figure 9, as a function of mixing time. Also shown in this figure are data from a non selective inversion recovery measurement. The solid lines are simulations employing 7- and 8-spin longitudinal relaxation matrices incorporating the amide protons from {V88, F87, A89, L90, F109, A139, V141} and {A89, V88, L90, Y91, G106, E107, K108, A139} for V88 and A89, respectively, and similar motional properties were assumed for all the spins involved. For the calculation of the self and cross-relaxation rates with Equations 8 and 9, the X-ray coordinates were used to generate the necessary inter-amide proton distances. All cross-correlation and chemical exchange effects were ignored. A contribution to the self-relaxation rate was added from ^{1}H chemical shift anisotropy, ^{1}H–^{15}N dipolar interaction and from a summation over all possible ^{1}H–^{2}H spin pairs. To improve the fits between the simulations and the experimental data it was found necessary, for some residues, to adjust the value of the self-relaxation rates in the relaxation matrix from the value calculated using Equation 8. The values for the cross-relaxation rates were not altered from those calculated using equation 9. These changes arise in part because the data are being modelled as a rigid sphere undergoing isotropic motion. Once a set of parameters was found, both the transient NOE time development and the inversion recovery measurements could be fitted simultaneously, merely by changing the initial conditions in the simulation.

The reasonable agreement between the data and the simulations indicates that even when employing long mixing times the data can be interpreted satisfactorily using only relatively few protons from the local enviroment. This is quite unlike the situation which holds for a fully protonated protein, where for mixing times $>$ 200 ms, the spectra become dominated by cross relaxation processes which would require much larger groupings of spins for their analysis (Kalk and Berendesen, 1976; James, 1994). We conclude that it is possible to record NOESY spectra from a highly enriched {^{15}N,^{2}H} labelled protein with long mixing times.

Collection of conformational restraints

3D HSQC-NOESY-HSQC datasets were collected with mixing times ranging from 10 ms to 1.2 s. A

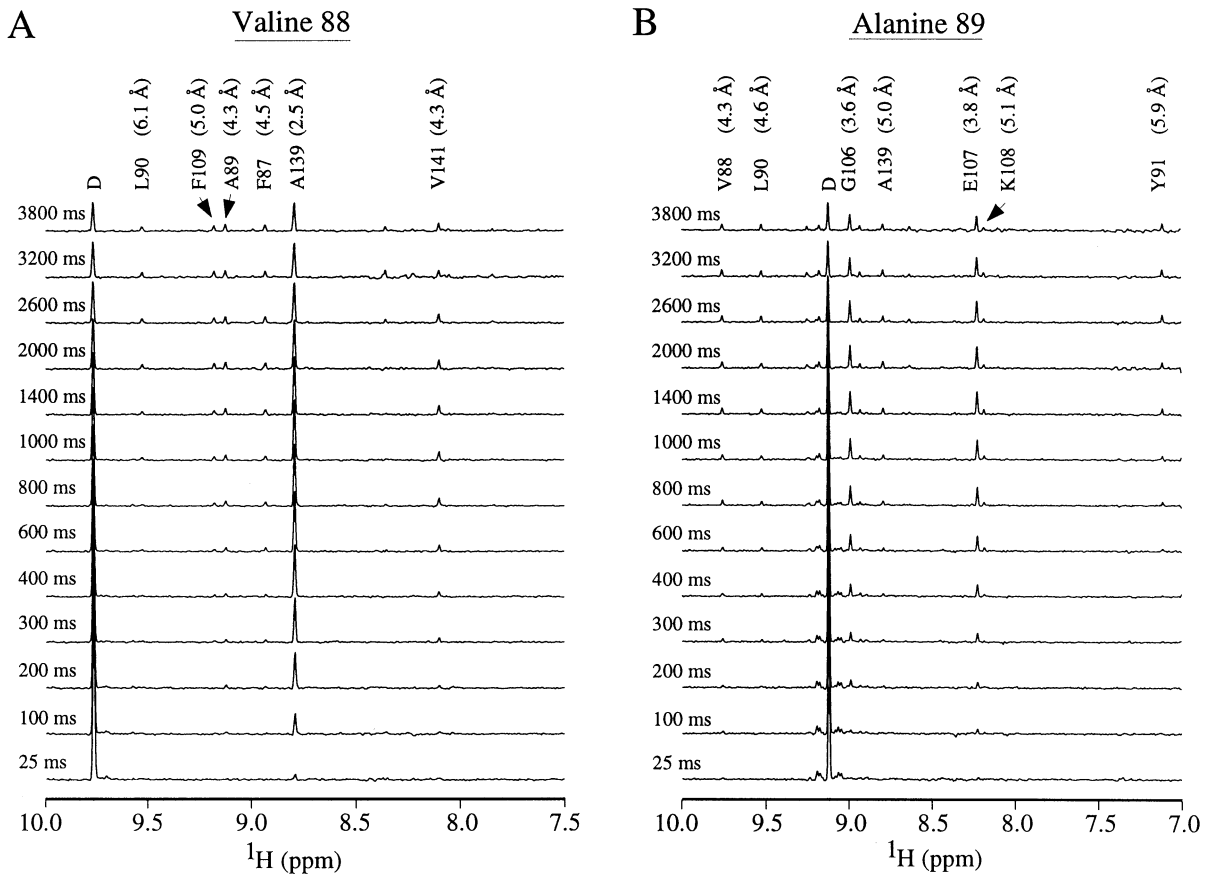

Figure 9. A stacked plot of cross sections, parallel to F2, from 2D homonuclear NOESY datasets recorded at 750MHz, showing the development of the inter amide ^{1}H–^{1}H NOE effects as a function of mixing time. The data are from residues (a) V88, (b) A89.

diagonal matrix plot of the 155 inter-amide NOEs, identified from an analysis of these datasets, is shown in Figure 11. They could be analysed into short-range, which are mainly sequential $d_{NN}(i, i+1)$ and medium range $d_{NN}(1 < |\ i-j\ | \leq 4)$ NOEs and $d_{NN}(|\ i-j\ | > 4)$ long-range NOEs from the β sheet regions. A summary of these NOEs is shown in Table 1. The previous NMR study reported 31 inter-amide NOEs, all of which were identified in this study. In addition, a total of seven NOEs, from tryptophan residues W119 (1) and W120 (6), were observed between the side chain solvent exchangeable indole proton and other amide protons, which had not been observed previously. No NOEs were observed between an amide proton and any of the assigned side chain solvent exchangeable protons of asparagine, glutamine or arginine residues. The 155 inter-amide and seven side chain-amide NOEs were broadly characterised into strong, medium and weak, using the criteria outlined earlier, and converted into distance upper limit constraints of 3 Å, 4 Å and 6.8 Å respectively.

Structure calculation and analysis

Four independent sets of structure calculations, denoted as [1], [2], [3], and [4], were performed. For [1] the input restraints list employed only the 162 NOEs while for [2] 35 ^3J(HN,Hα) torsion angle restraints were used in addition to the 162 NOEs observed in this work. Calculation [3] employed the NOE restraint list previously reported (Morton et al., 1996) with the 35 torsion angle restraints, but hydrogen bond restraints were excluded. The distance upper limits for this set of NMR restraints were 2.8 Å, 3.5 Å and 5 Å (Morton et al., 1996). For calculation [4] the additional 131 NH–NH NOEs were added, giving a total of 754 restraints as input data. Fifty structures were calculated for each set of restraints [1], [2], [3] and [4]. From the results of these calculations a family of 20 structures

Table 1. A summary of input data and statistics from the structure calculations

Input data	**[1]**	**[2]**	**[3]**	**[4]**
Sequential: $d_{NN}(i, i+1)$	$48''$	$48''$	18^*	$18^* + 30''$
Medium-range: $d_{NN}(2 \leq i-j \leq 4)$	$40''$	$40''$	2^*	$2^* + 38''$
Long-range: $d_{NN}(i-j \geq 5)$	$67''$	$67''$	11^*	$11^* + 56''$
Additional NOE from Tryptophan side-chains	$7''$	$7''$	–	$7''$
Additional NOE restraints	–	–	557*	557*
$^3J_{HN-H\alpha}$ torsion angle restraints	–	35	35	35
Total number of restraints	162	197	623	754
Structural statistics for 20 NMR derived structures				
Average rmsd from experimental distance restraints (Å)[†]				
Sequential	0.012 ± 0.011	0.014 ± 0.014	0.005 ± 0.003	0.021 ± 0.005
Medium-range	0.014 ± 0.009	0.012 ± 0.007	0.004 ± 0.001	0.012 ± 0.004
Long-range	0.027 ± 0.010	0.028 ± 0.012	0.002 ± 0.001	0.031 ± 0.003
Average rmsd from experimental torsion angle restraints (deg)[#]				
ϕ angles (35)	–	0.51 ± 0.04	0.48 ± 0.01	0.51 ± 0.01
X-PLOR potential energies (kcal mol^{-1})				
E_{total}	80 ± 8	83 ± 13	68 ± 1	93 ± 6
E_{noe}	3 ± 2	4 ± 3	2 ± 1	13 ± 2
Cartesian coordinate rmsd (Å) of the backbone heavy atoms (N, Cα, C′)				
<SA> versus <$\overline{\text{SA}}$>[$]	2.02 ± 0.26	1.76 ± 0.28	0.75 ± 0.12	0.45 ± 0.05
<SA> versus <$\overline{\text{SA}}$>[**]	1.52 ± 0.22	1.25 ± 0.27	0.45 ± 0.07	0.35 ± 0.04
<SA> versus X-ray structure[$]	2.86 ± 0.33	2.47 ± 0.30	1.27 ± 0.21	1.17 ± 0.07
<SA> versus X-ray structure[**]	2.01 ± 0.28	1.57 ± 0.32	0.76 ± 0.13	0.61 ± 0.05
Ramachandran plot[⊥]				
most favourable region[$](%)	24	47	69	69
most favourable region[**](%)	28	52	80	82

<SA> is the ensemble of 20 simulated annealing structures; <$\overline{\text{SA}}$> is the mean atomic structure obtained by averaging the coordinates of the individual <SA> structure following a least-square superposition of the backbone heavy atoms (N, Cα, C′).

[†]NOE-derived distance restraints were used with a soft square-well potential ($F_{noe} = 50$ kcal mol^{-1} Å^2). No distance restraints was violated by more than 0.3 Å in any of the 20 final structures.

[#]Torsion angle restraints were applied to 35 residues (Morton et al., 1996). A force constant of 200 kcal mol^{-1} rad^{-2} was applied to each torsion restraint.

[″]These NOE restraints were classified as strong (3.0 Å), medium (4.0 Å) and weak (6.8 Å).

[*]These NOE restraints were employed as classified previously (Morton et al., 1996).

[**]for residues 86–89, 108–113, 119–124, 130–133, 138–140.

[$]for residues 85–141.

[⊥]using PROCHECK_nmr.

was selected for analysis in each case. A statistical summary of both the input and output data from the four calculations is shown in Table 1. For the analysis of the structures which resulted from calculations [1] and [2] the selection criteria made no reference to the known structure.

From the results of calculation [1] 12 structures were found to have a significantly higher NOE energy than the remainder and each of these was also found to have large NOE violations (> 0.3 Å). These 12 structures were excluded from any further analysis. Inspection of the pairwise backbone rmsd values between each of the remaining 38 structures indicated that they could be separated into two sets with 20 and 18 members, respectively. The larger and smaller set were found to have a mean global pairwise backbone rmsd value of 2.93 ± 0.35 Å and 2.97 ± 0.37 Å respectively. However, the mean global pairwise backbone

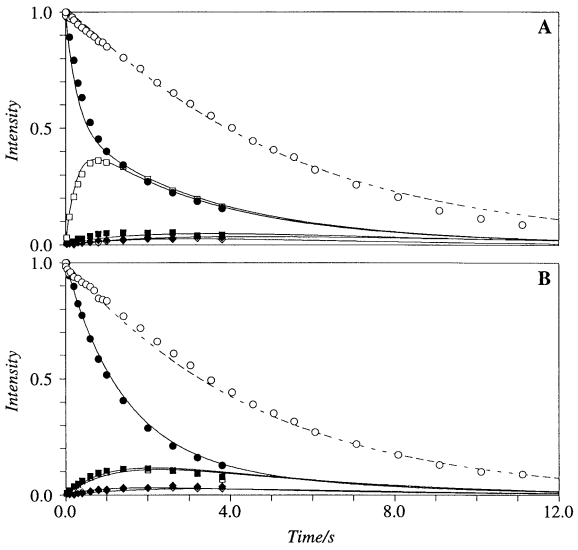

Figure 10. (a) Dependence of the diagonal and some cross peak intensities for V88, using the data from Figure 9, as a function of mixing time. V88 diagonal peak ●, A139 cross peak □, V141 cross peak ■, F109 cross peak ♦, L90 cross peak ◊, non-selective inversion recovery o. The solid lines are simulations to the experimental data with $r_{NH} = 0.102$ nm, $B_0 = 17.6$ T, $\tau_r = 6.6$ ns, $\Delta\sigma(H) = 10$ ppm. (b) Dependence of the diagonal and some cross peak intensities for A89, using the data from Figure 9, as a function of mixing time. A89 diagonal peak ●, G106 cross peak ■, E107 cross peak □, Y91 cross peak ◊, V88 cross peak ♦, non-selective inversion recovery o. The solid lines are simulations to the experimental data which used similar parameters to those for Figure 10a.

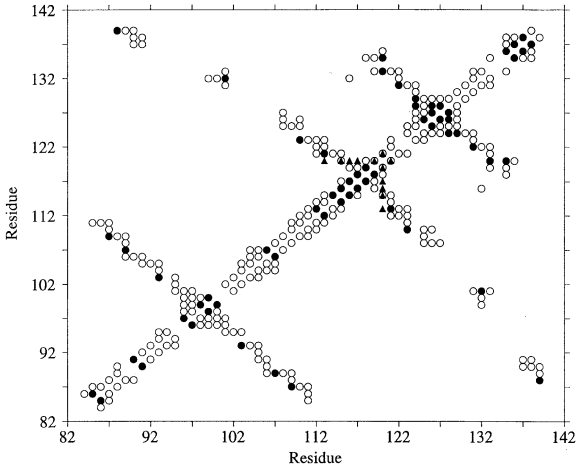

Figure 11. A diagonal plot representing the 162 inter-amide NOEs identified from an analysis of 3D HSQC-NOESY-HSQC datasets. The inter-amide NOEs established in the previous NMR study (Morton et al., 1996) are shown as black circles. The filled triangles represent the NOEs involving the tryptophan side chain indole protons of residues W119 and W120.

rmsd between these two sets was 6.31±3.30 Å. We associate the difference between these two sets of structures with the correct fold and the mirror image fold. A superposition of the mean structure of the smaller family, after inversion of one of the coordinate sets, onto the mean structure of the larger family gave an rmsd of 1.01 Å; this is consistent with the smaller family being a mirror image of the larger family and was confirmed by direct visualisation. Analysis of these two families of NMR structures using PROCHECK_nmr (Laskowski et al., 1993) shows that the larger set has a higher percentage of residues falling in the most favourable region of the Ramachandran plot (24% compared to 17%). From the larger set all 20 structures were selected, and these are displayed in Figure 12 [1]. The backbone atoms from the X-ray structure are also displayed in Figure 12 [1] and this comparison shows that the larger set of NMR structures is associated with the correct fold. These 20 structures were used for the comparisons reported in Table 1. The global fold of the compact β barrel, which consists of two orthogonal β-sheets (Noble et al. 1993 and Morton et al. 1996), is readily discernable in Figure 12 [1]. However, the DSSP algorithm of Kabsch and Sander (1983) does not completely identify the known secondary structure elements for any of the 20 NMR structures. We have calculated much larger sets of structures than those reported above, and in all cases it was found that the resulting family of structures could be separated into two classes, with the larger set always being identified with the correct fold. Venters et al. (1995) have also pointed out that, if a resonable number of inter-amide NOEs could be observed between protons up to about 7 Å apart, then this NOE data set could be used to calculate the global fold of a protein. The global fold has been established here using, on average, 2.84 NOE restraints per residue.

From the results of calculation [2] ten structures were found to have significantly higher NOE energies than the remainder and each of these was also found to have a large NOE violation (> 0.3 Å) with a significantly higher total energy. These ten structures were excluded from any further analysis. Inspection of the pairwise backbone rmsd values between each of the remaining 40 structures indicated that they could be separated into two sets with 23 and 17 members, respectively. The larger and smaller set were found to have a mean global pairwise backbone rmsd value of 2.56 ± 0.37 Å and 2.97 ± 0.37 Å, respectively. However, the mean global pairwise backbone rmsd between these two sets was 6.02 ± 3.33 Å. As be-

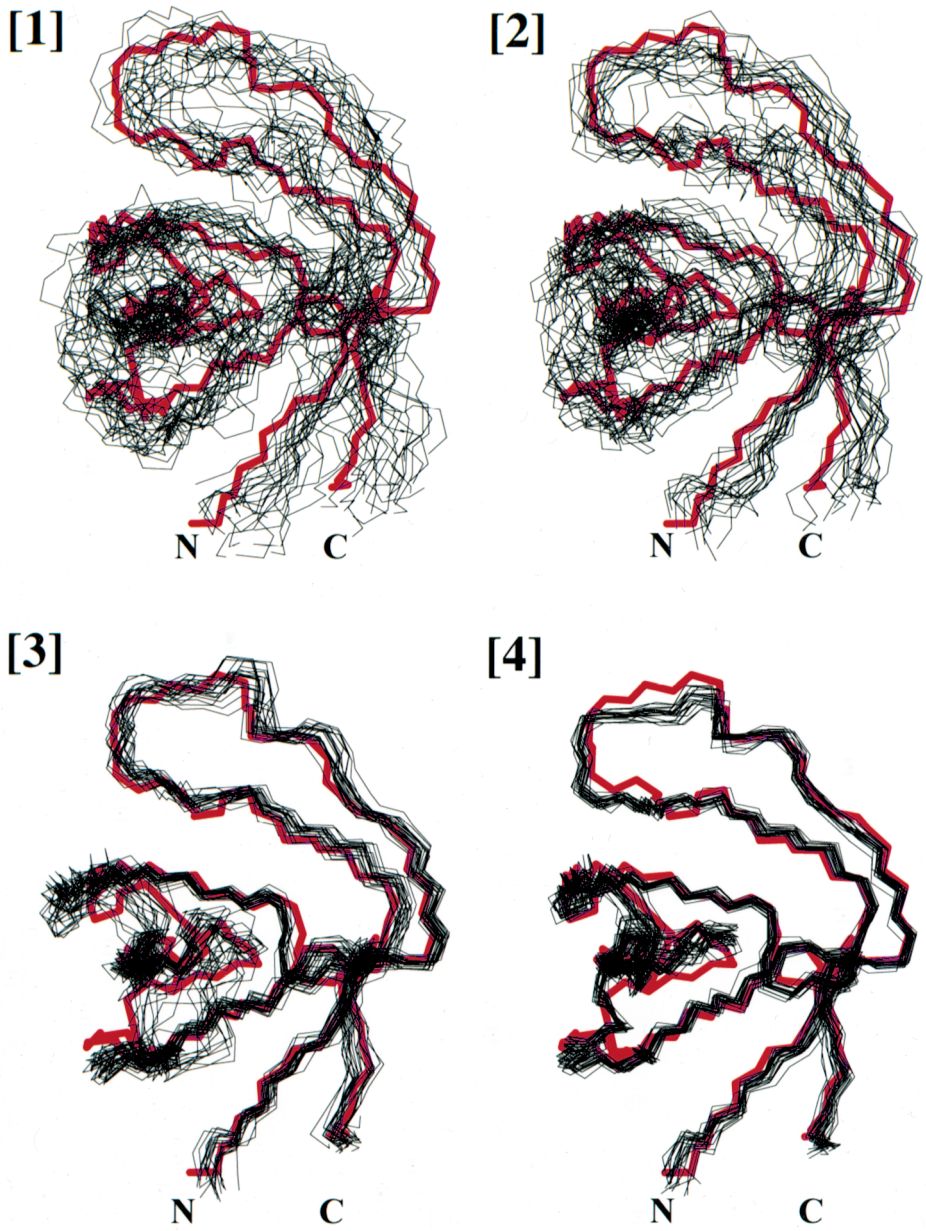

Figure 12. The family of 20 structures resulting from calculations [1], [2], [3] and [4]. The backbone atoms for residues 85–141 of the X-ray structure are superimposed upon each of the families of NMR structures and is displayed in red.

fore, we associate the difference between these two sets of structures with the correct fold and the mirror image fold. A superposition of the mean structure of the smaller family, after inversion of one of the coordinate sets, onto the mean structure of the larger family gave an rmsd of 1.38 Å; this is consistent with the smaller family being a mirror image of the larger family and was confirmed by direct visualisation. Analysis of these two families of NMR structures using PROCHECK_nmr also shows that the larger set has a higher percentage of residues falling in the most favourable region of the Ramachandran plot (47% compared to 31%). From the larger set a family of 20 structures was selected, associated with the lowest total energy, and these structures are displayed in Figure 12 [2]. These 20 structures were used for the comparisons reported in Table 1. The DSSP algorithm recognizes secondary structure elements for

residues 87–89, 109–111, 119–124, 131–133 and 138–140. Thus, the additional backbone ϕ dihedral angle restraints based on the previously determined experimental ^3J(HN–Hα) values, have improved the precision of the structure determination which is also evident from Table 1.

From the results of calculation [3] 20 structures, with the lowest total and NOE energies, were selected and these are shown in Figure 12 [3]. These 20 structures were used for the comparisons reported in Table 1. These structures are similar to those reported previously (Morton et al. 1996).

From the results of calculation [4] 20 structures were selected with the lowest total energy and these are shown in Figure 12 [4]. Analysis of the 20 final structures from calculations [3] and [4], using PROCHECK_nmr, shows very little improvement in the quality of stuctures after addition of extra interamide NOEs. This occurs because the NMR structures from calculation [3] already show very few violations of the Ramachandran allowed regions. However, the extra NOEs have reduced the rmsd compared to the NMR structures from calculation [3] (see Table 1). Furthermore, within each family of 20 structures from calculations [3] and [4] the average pairwise backbone rmsd for all residues was found to be 1.08 Å and 0.65 Å, respectively. This decrease to the pairwise backbone rmsd value indicates that the additional NOEs have improved the precision of the NMR structures.

Conclusion

We have shown that studies of a highly enriched {^{15}N,^{2}H} protein can yield interesting new information. The accepted model for the spin diffusion limit in normal ^{1}H-rich proteins, where $\omega\tau_r \gg 1$ is that the ^{1}H relaxation of all the spins becomes uniform (Kalk and Berendsen, 1976). This is not observed in a sample with dilute ^{1}H spins and has allowed the detection of local dipolar interactions between amide protons up to 6 to 7 Å apart. This additional long-range NOE information can be used to provide additional structural restraints in the NMR structure refinement process. A second advantage is the significantly improved spectral resolution in the ^{1}H dimension of a perdeuterated sample; this may be particularly important in those systems, such as partially folded proteins, where the chemical shift dispersion is poor. We have also shown that perdeuterated samples allow analysis of the ^{1}H self-, cross- and transverse-relaxation characteristics. The results shown here suggest that perdeuterated samples might also prove useful for iterative back-calculation structure refinements.

Acknowledgements

This work is a contribution from the Oxford Centre for Molecular Sciences (OCMS), which is supported by the BBSRC, EPSRC and MRC. The 750 MHz NMR instrument was partly funded by the LINK protein engineering project with industrial partners Zeneca and Oxford Instruments. I.D.C. also thanks the Wellcome Trust for financial support. T.K.M. thanks the INLAKS foundation for financial support. H.K. acknowledges funding from OCMS. We thank Dr. C. Redfield for providing the program to calculate intrinsic amide exchange rates.

References

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford, U.K.

Anet, F.A.L. and O'Leary, D.J. (1990) *J. Magn. Reson.*, **86**, 358–370.

Bai, Y., Milne, J.S., Mayne, L. and Englander, S.W. (1993) *Proteins*, **17**, 75–86.

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Letts.*, **69**, 185–189.

Brünger, A.T. (1992) *X-PLOR Manual v.3.1*, Yale University, New Haven, CT, USA.

Campbell, I.D. and Freeman, R. (1973) *J. Magn. Reson.*, **11**,143–162.

Campbell, I.D., Dobson, C.M., Ratcliffe, R.G and Williams, R.J.P. (1978) *J. Magn. Reson.*, **29**, 397–417.

Chothia, C. (1979) *Nature*, **254**, 304–308.

Cooke, M.P. and Perlmutter, R.M. (1989) *New Biol.*, **1**, 66–74.

Crespi, H.L., Rosenberg, R.M. and Katz, J.J. (1968) *Science*, **161**, 795–796.

Devore, J.L. (1991) *Probability and Statistics for Engineering and Sciences*, Duxbury Press, Belmont, California.

Englander, S.W. and Kallenbach, N.R. (1983) *Q. Rev. Biophys.*, **16**, 521–655.

Farmer, B.T. and Venters, R.A. (1995) *J. Am. Chem. Soc.*, **117**, 4187–4188.

Forsen, S.H. and Hoffman, R.A. (1963) *J. Chem. Phys.*, **39**, 2892–2901.

Freeman, R. and Hill, H.D.W. (1971) *J. Chem. Phys.*, **54**, 3367–3377.

Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.*, **90**, 420–425.

Garcia de la Torre, J. and Bloomfield, V.A. (1981) *Quart. Rev. Biophys.*, **14**, 81–139

Garcia de la Torre, J., Navarro, S., Lopez Martinez, M.C., Diaz, F.G. and Lopez Cascales, J.J. (1994) *Biophys. J.*, **67**, 530–531.

Gardner, K.H., Rosen, M.K., and Kay, L.E. (1997) *Biochemistry*, **36**, 1389–1401.
Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.
Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 627–638.
Grzesiek, S., Wingfield, P., Stahl, S., Kaufman, J.D. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 9594–9595.
Hoult, D.J. and Richards, R.E. (1975) *Proc. R. Soc. London A*, **344**, 311–340.
Huang, T. and Shaka, A.J. (1995) *J. Magn. Reson. Ser. A*, **112**, 275–279.
James, T.L. (1994) *Methods of Enzymology*, **Volume 239** C, p. 416.
Jameson, C.J. (1996) *Encyclopaedia of Nuclear Magnetic Resonance*, (Ed.) Grant, D.M. and Harris, R.R. *Isotope Effects on Chemical Shifts and Coupling constants*, pp. 2638–2655.
Kabsch W. and Sander C. (1983) *Biopolymers*, **22**, 2577–2637.
Kalk, A. and Berendson, H.J.C. (1976) *J. Magn. Reson.*, **24**, 343–365.
Kay, L.E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.
Koradi, R., Billeter, M. and Wuthrich, K. (1996) *J. Mol. Graphics.*, **14**, 51–55.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.
Le Master, D.M. and Richards, F.M. (1988) *Biochemistry*, **27**, 142–150.
Liepinsh, E., Otting, G. and Wuthrich, K. (1992) *J. Biomol NMR*, **2**, 447–465.
Markley, J.L., Putter, I. and Jardetzky, O. (1968) *Science*, **161**, 1249–1251.
Markus, M.A., Dayie, K.T., Mastsudaira, P. and Wagner, G. (1994) *J. Magn. Reson. Ser. B*, **105**, 192–195.
McConnell, J. (Ed.) (1987) *The Theory of Nuclear Magnetic Relaxation in Liquids*, Cambridge University Press, Cambridge.
Morton, C.J., Pugh, D.J.R., Brown, E.L.J., Kahmann, J.D., Renzoni, D.A.C., and Campbell, I.D. (1996) *Structure*, **4**, 705–714.
Muchmore, S.W., Sattler, M., Liang, H., Meadows, R. P., Harlan, J.E., Yoon, H.S., Nettesheim, D., Chang, B.S., Thompson, C.B., Wong, S.L., Ng, S.C., and Fesik, S.W. (1996) *Nature*, **381**, 355–341.
Nietlispach, D., Clowes, R.T., Broadhurst, W., Ito, Y., Keeler, J., Kelly, M., Ashurst, J., Oschkinat, H., Doamille, P.J. and Laue, E.D. (1996) *J. Am. Chem. Soc.*, **118**, 407–415.
Noble, M.E.M., Musacchio, A., Saraste, M., Courtneidge, S.A. and Wieranga, R.K. (1993) *EMBO J.*, **12**, 2617–2624.
Ottiger, M. and Bax, A. (1997) *J. Am. Chem. Soc.*, **119**, 8070–8075.
Otting, G., Liepnish, E., Farmer II, B.T., and Wuthrich, K. (1991) *J. BioMol. NMR*, **1**, 209–215.
Sattler, M., Liang, H., Nettesheim, D., Meadows, R.P., Harlan, J.E., Eberstadt, M., Yoon, H.S., Shuker, S.B., Chang, B.S., Minn, A.J., Thompson, C.B., and Fesik, S.W. (1997) *Science*, **275**, 983–986.
Shaka, A.J., Keeler, J. and Freeman, R. (1983) *J. Magn. Reson.*, **53**, 313–340.
Sklenar, V. (1995) *J. Magn. Reson. Ser. A*, **114**, 132–135.
Smith, B.O., Ito, Y., Raine, A., Teichmann, S., Ben-Tovim, L., Nietlispach, D. Broadhurst, R.W., Terada, T., Kelly, M., Oschkinat, H., Shibata, T., Yokoyama, S., and Laue, E.D. (1996) *J. BioMol. NMR*, **8**, 360–368.
Soffe, N., Boyd, J., and Leonard, M. (1996) *J. Magn. Reson. Ser. A*, **116**, 117–121.
Solomon, I. (1955) *Phys. Rev.*, **99**, 559–565.
Spera, S., Ikura, M., and Bax, A. (1991) *J. Biomol. NMR*, **1**, 155–165.
Stott, K., Keeler, J., Van, N.Q., and Shaka, A.J. (1997) *J. Magn. Reson.*, **125**, 302–324.
Tjandra, N., Grzesiek, S., and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6264–6272.
Tjandra, N., Szabo, A., and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6986–6991.
Torchia, D.A., Sparks, S.W., and Bax, A. (1988) *J. Am. Chem. Soc.*, **110**, 2320–2321.
Tsang, P., Wright, P.E., and Rance, M. (1990) *J. Am. Chem. Soc.*, **112**, 8183–8185.
Venters, R.A., Metzler, W.J., Spicer, L.D., Mueller, L., and Farmer, B.T. (1995) *J. Am. Chem. Soc.*, **117**, 9592–9593.
Venters, R.A., Farmer, B.T., Fierke, C.A., and Spicer, L.D. (1996) *J. Mol. Biol.*, **264**, 1101–1116.
Wang, Y., Darøon, F.I., Grzesiek, S., Torchia, D. A., Wingfield, P.T., Kaufman, J.D., Stahl, S.J., Chang, C. and Hodge, C.N. (1996) *Biochemistry*, **35**, 12694–12704.
Werbelow, L. (1996) *Encyclopaedia of Nuclear Magnetic Resonance*, (Ed.) Grant, D.M. and Harris, R.R. *Dynamic Frequency Shift*, pp. 1776–1783.
Wüthrich, K (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R., and Kay, L.E. (1994) *J. Am. Chem. Soc.*, **116**, 11655–11666.
Yamazaki, T., Tochio, H., Furui, J., Aimoto, S., and Kyogoku, Y. (1997) *J. Am. Chem. Soc.*, **119**, 872–880.
Yu, L., Petros, A.M., Schnuchel, A., Zhong, P., Severin, J.M., Walter, K., Holzman, T.F. and Fesik, S.W. (1997) *Nat. Struct. Biol.*, **4**, 483–489.
Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C. and Jardetzky, O. (1994) *J. Mol. Biol.*, **238**, 592–612.
Zhou, M.M, Ravichandran, K.S., Olejniczak, E.T., Petros, A.M., Meadows, R.P., Sattler, M., Harlan, J. E., Wade, W.S., Burakoff, S.J. and Fesik, S.W. (1995) *Nature*, **378**, 584–592.