Experiments and strategies for the assignment of fully13 C/15N-labelled polypeptides by solid state NMR

Author(s):
Straus, Suzana K.; Bremi, Tobias; Ernst, Richard R.

Publication date:
1998

Permanent link:
https://doi.org/10.3929/ethz-b-000422785

Originally published in:
Journal of Biomolecular NMR 12(1), https://doi.org/10.1023/A:1008280716360

Journal of Biomolecular NMR, **12:** 39–50, 1998.
KLUWER/ESCOM

Experiments and strategies for the assignment of fully $^{13}C/^{15}N$-labelled polypeptides by solid state NMR

Suzana K. Straus, Tobias Bremi and Richard R. Ernst*
Laboratorium für Physikalische Chemie, ETH Zentrum, CH-8092 Zürich, Switzerland

Received 9 December 1997; Accepted 26 January 1998

Key words: assignment, fully $^{13}C/^{15}N$-labelled biomolecules, heteronuclear correlation, homonuclear decoupling, resolution enhancement, solid state NMR

Abstract

High-resolution heteronuclear NMR correlation experiments and strategies are proposed for the assignment of fully $^{13}C/^{15}N$-labelled polypeptides in the solid state. By the combination of intra-residue and inter-residue ^{13}C-^{15}N correlation experiments with ^{13}C-^{13}C spin-diffusion studies, it becomes feasible to partially assign backbone and side-chain resonances in solid proteins. The performance of sequences using ^{15}N instead of ^{13}C detection is evaluated regarding sensitivity and resolution for a labelled dipeptide (L-Val-L-Phe). The techniques are used for a partial assignment of the ^{15}N and ^{13}C resonances in human ubiquitin.

Introduction

In liquid state NMR spectroscopy, multidimensional homonuclear and heteronuclear chemical shift correlation methods provide a basis for the resonance assignment of proteins (Wüthrich, 1986; Ikura et al., 1990; Kay et al., 1991; Grzesiek and Bax, 1992; Clore and Gronenborn, 1993; Kay, 1995; Cavanagh et al., 1996; Dotsch et al., 1996). A full resonance assignment of at least the backbone atoms is a prerequisite for structural or dynamical investigations. In the liquid state, assignment and structure determination of fully $^{13}C/^{15}N$-labelled proteins of up to 30 kDa is today possible (Bax et al., 1990; Clore et al., 1990). Recent developments (Tjandra and Bax, 1997; Pervushin et al., 1997) might allow access to structures of even larger proteins.

Also for investigations of proteins in the solid state, an assignment of resonances is mandatory. Proton resonance alone is unsuited for resonance assignments in the solid state because even the most sophisticated resolution enhancement techniques, combining magic angle sample spinning (MAS) and multiple pulse dipolar decoupling (CRAMPS), do not reduce the line width below 100 Hz in amino acids (Naito et al., 1991). Since low γ nuclei, such as ^{13}C and ^{15}N, have smaller dipolar couplings and a larger chemical shift range than protons, they are usually the nuclei of choice in solid state studies of peptides and proteins. To circumvent the need of a full assignment of all resonances and to avoid heavily overlapping lines, recourse is often taken to selective labelling allowing for investigations at specific sites (Griffiths et al., 1995; Lansbury et al., 1995; Heller et al., 1996; Hirsh et al., 1996). For a structural or dynamical study of an entire protein in solid state, a large number of selectively labelled samples would have to be prepared. Such a procedure is time-consuming, expensive, and hardly practical.

To ease the chemical requirements of sample preparation, it would be desirable to use fully ^{13}C- and/or ^{15}N-labelled molecules for assignment and structure determination purposes in the solid state. Taking advantage of recent developments, significant improvements in the resolution of heteronuclear MAS spectra are possible. First, heteronuclear proton decoupling by the TPPM sequence (Bennett et al., 1995) efficiently removes the heteronuclear broadening, and, secondly, by homonuclear decoupling (Straus et al., 1996) it is possible to also suppress the homonuclear

*To whom correspondence should be addressed.

J splittings. This opens possibilities for a resonance assignment of larger peptides and proteins and allows the investigation of structural and dynamical features of polypeptides in the solid state (Straus et al., 1997).

In the liquid phase, the backbone resonances of fully labelled proteins are usually assigned by methods like HCACO, HNCO, HNCA, H(CACO)N, etc. Some of them establish intra-residue connectivities (HNCA, HCACO, etc.), while others connect resonances of neighbouring amino acids (HNCO, H(CACO)N, etc.). Knowing the primary structure of a protein, normally an unambiguous heteronuclear assignment is possible with these methods.

The strategy proposed here for solid state applications is closely related to the concepts used in the liquid phase. It starts with an identification of the ^{13}C resonances belonging to the same amino acid residue based on homonuclear dipolar ^{13}C correlation spectra in analogy to Straus et al. (1997). For the intra-residue assignment of ^{15}N resonance, a $^{13}C^{\alpha}$-^{15}N (CAN) correlation is proposed with high resolution in the indirectly observed ^{13}C dimension (Straus et al., 1996) and with detection on ^{15}N. For the inter-residue assignment, $N_{i+1}(CO)CA_i$ and $CA_i(CO)N_{i+1}$ experiments are suggested to establish, in a directed manner, a connectivity between the $^{13}C^{\alpha}$ resonance of residue i and the ^{15}N resonance of residue i+1. The advantages and disadvantages of ^{15}N detection are analysed for the proposed experiments. The possible increase in resolution when applying homonuclear ^{13}C decoupling during the indirectly observed time periods may be decisive for using ^{15}N detection.

In the 'Assignment strategy and pulse sequences' section, the strategy and the proposed pulse sequences are described. This is followed in the 'Experiments on the dipeptide' section by the application to the fully $^{13}C/^{15}N$-labelled dipeptide L-Val-L-Phe to explore the achievable signal-to-noise (S/N) ratio and the possible gain in resolution by homonuclear decoupling when using ^{15}N instead of ^{13}C detection. In the 'Partial assignment of ubiquitin' section, the methods are applied to fully $^{13}C/^{15}N$-labelled ubiquitin, a 76-residue protein. This example demonstrates that the methods provide a promising basis for the assignment of proteins in the solid state.

Assignment strategy and pulse sequences

The backbone assignment is based exclusively on the $^{13}C^{\alpha}$ and ^{15}N resonances which have a sufficiently large chemical shift dispersion. ^{13}CO resonances are only used as relay stations for two-step polarization transfers as their dispersion turned out to be insufficient for an adequate resolution. The intra-residue $^{13}C_i^{\alpha}$-$^{15}N_i$ correlation can easily be achieved by a direct cross-polarization step while the sequential inter-residue $^{13}C_i^{\alpha}$-$^{15}N_{i+1}$ correlation requires a two-step transfer: (i) from $^{13}C_i^{\alpha}$ to $^{13}CO_i$ by spin diffusion and (ii) by cross polarization from $^{13}CO_i$ to $^{15}N_{i+1}$. In addition, ^{13}C spin diffusion is employed for the identification of the types of amino acid residues based on their side-chain resonances.

The proposed heteronuclear correlation experiments are given in Figure 1. The pulse sequence in Figure 1a correlates $^{13}C^{\alpha}$ with ^{15}N within the same residue by the heteronuclear cross-polarization step CP2 (Hediger et al., 1994,1995; Metz et al., 1994). This 'CAN' pulse sequence is related to the 'NCA' sequence proposed by Sun et al. (1995) and Baldus et al. (1996). The reversed order is chosen such that resolution enhancement by homonuclear ^{13}C-^{13}C scalar decoupling during t_1 becomes possible (Straus et al., 1996). The resolution in ω_2 is additionally improved by applying GARP-1 decoupling (Shaka et al., 1985) on the ^{13}C channel during ^{15}N detection. GARP-1 was found to be more efficient than other decoupling sequences (Croasmun and Carlson, 1994), such as CW decoupling, WALTZ-16, and DIPSI-2, in this case.

The $N_{i+1}(CO)CA_i$ and $CA_i(CO)N_{i+1}$ sequences, given in Figures 1b and c, respectively, are designed to establish sequential inter-residue connectivities. They both yield a $^{13}C_i^{\alpha}$ -$^{15}N_{i+1}$ correlation spectrum with the difference that the first experiment uses ^{13}C detection while the second employs ^{15}N detection. In the ^{13}C-detected experiment (Figure 1b), magnetization is transferred from protons to $^{15}N_{i+1}$ where it evolves under heteronuclear decoupling in t_1. The magnetization is then transferred via cross polarization to the neighbouring carbon spins $^{13}CO_i$ and $^{13}C_i^{\alpha}$. Subsequently, a Gaussian $\pi/2$ pulse, applied to the carbonyl region, selectively flips the $^{13}CO_i$ magnetization to the z-axis, and from there a proton-driven spin-diffusion process to the neighbouring $^{13}C_i^{\alpha}$ is allowed to take place. The $^{13}C_i^{\alpha}$ magnetization is then brought into the transverse plane by a non-selective $\pi/2$ pulse and detected under high-power proton decoupling. By selectively using the ^{13}CO as a relay nucleus, the transfer becomes directed and only correlations between $^{13}C_i^{\alpha}$ and $^{15}N_{i+1}$ appear in the 2D spectrum, thus reducing the number of peaks. Note

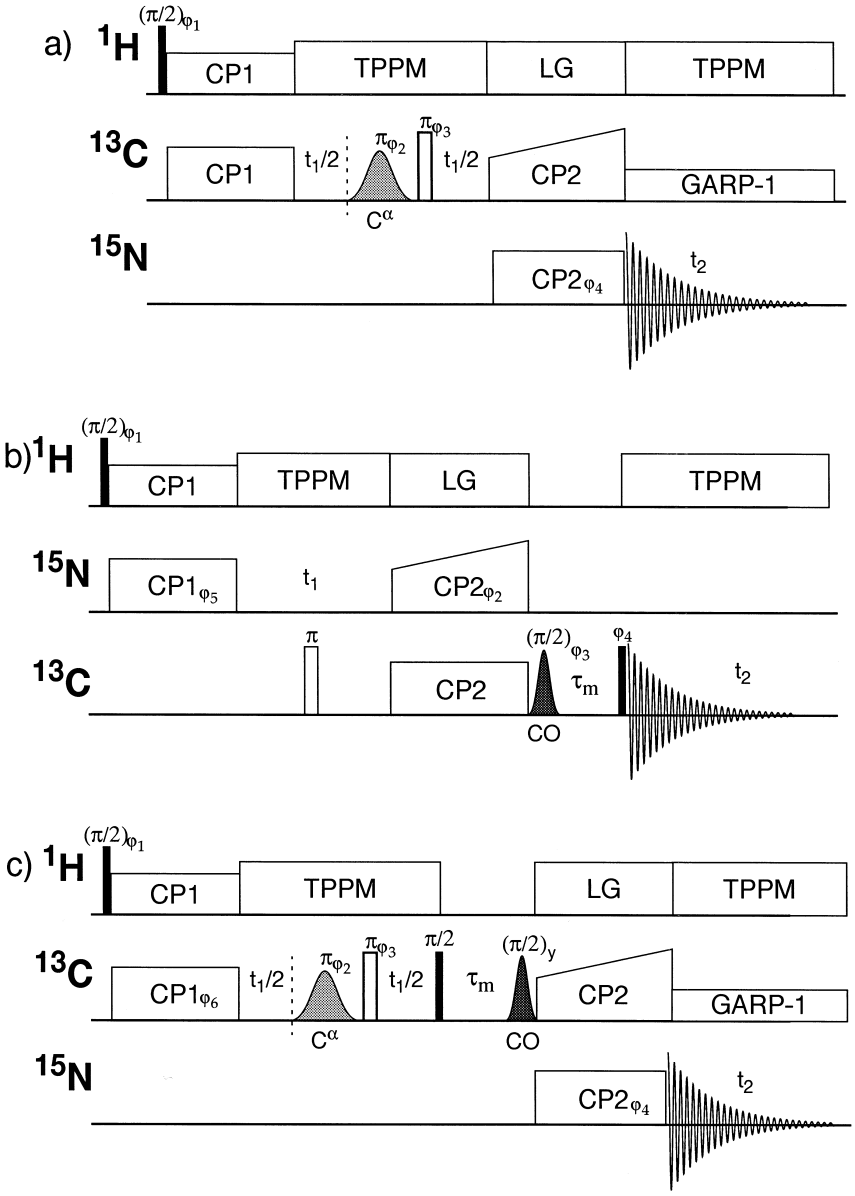

Figure 1. Pulse sequences to correlate ^{13}C chemical shifts with ^{15}N shifts: (a) CAN experiment, correlating ^{13}C–^{15}N pairs within a residue; (b) $N_{i+1}(CO)CA_i$ and (c) $CA_i(CO)N_{i+1}$ experiments, used to correlate $^{13}C_i^{\alpha}$-$^{15}N_{i+1}$ pairs. The minimal phase cycle in (a), resulting in artefact-free spectra with subtraction of all out-of-band resonances, is (φ_1 = y,−y; φ_2 = x,x,x,x,y,y,y,y; φ_3 = x; φ_4 = x,x,−x,−x; and receiver phase: x,−x,−x,x,−x,x,x,−x. Quadrature detection in t_1 is achieved by TPPI, incrementing φ_4. In (b), the minimal phase cycle is: φ_1 = y,−y; φ_2 = x,x,−x,−x; φ_3 = y; φ_4 = y,y,y,y,−y,−y,−y,−y; and receiver phase: x,−x,−x,x,−x,x,x,−x. Quadrature detection in t_1 is achieved by TPPI, incrementing φ_5. In (c), the phases are as for (a). The phase φ_6 is incremented for TPPI. The first cross-polarization step (CP1) consists of Hartmann–Hahn cross-polarization, whereas for the second step (CP2) a linear amplitude sweep was applied on one of the channels (see text). Lee–Goldburg decoupling (LG) (Lee and Goldburg, 1965) was applied on the ^1H channel during CP2 (Baldus et al., 1996). All pulses except the shaped ones are high power ($\omega_H/2\pi$ = 80.0 kHz; $\omega_C/2\pi$ = 62.5 kHz; $\omega_N/2\pi$ = 38.5 kHz). The low-power Gaussian pulses are described in the text. Heteronuclear proton decoupling uses TPPM (Bennett et al., 1995). During ^{15}N detection, GARP-1 ^{13}C decoupling (Shaka et al., 1985) is applied.

that the magnetization transferred to $^{13}C^{\alpha}_{i+1}$ during the cross-polarization process is dephasing during τ_m.

In the ^{15}N-detected version of the above experiment (Figure 1c), magnetization is first transferred from protons to $^{13}C^{\alpha}_{i}$ where it evolves under hetero- and homonuclear decoupling in t_1. The magnetization is then flipped along z and undergoes a proton-driven spin-diffusion process (as described above) to the neighbouring carbon atom, $^{13}CO_i$. With a selective Gaussian $\pi/2$ pulse, the $^{13}CO_i$ magnetization is then brought into the transverse plane and subsequently transferred by cross polarization to the neighbouring $^{15}N_{i+1}$ atom of the next residue, the evolution of which is then detected under high-power proton decoupling. Again, the ^{15}N-detected experiment has the advantage that homonuclear decoupling is possible leading to increased resolution in the ^{13}C dimension. Note that in both experiments the chemical shift information of ^{13}CO could also be detected in an additional evolution period after the spin-diffusion step, allowing the assignment of these resonances as well.

Under the present experimental conditions (spinning frequency $\nu_r = 15$ kHz and available rf field strength $\omega_1(^{13}C)/2\pi \approx 60$ kHz), it was found that proton-driven $^{13}C^{\alpha}$-^{13}CO transfer is not only the easiest to implement but also equally efficient as rf-driven schemes. At a magnetic field strength of 9.4 T, the chemical shift frequency difference is $\nu_{^{13}C^{\alpha}}$-$\nu_{^{13}CO} \approx$ 12 kHz. It is therefore not far from the first spinning sideband matching condition such that the rotor-driven transfer mechanism (Andrew et al., 1963; Colombo et al., 1987) enhances the spin-diffusion rate. Simple schemes, such as biselective spin-locking of the $^{13}C^{\alpha}$ and ^{13}CO resonances with Hartmann–Hahn matching of the fields, turned out to be less efficient than the proton-driven transfer. More advanced techniques, such as RIL (Baldus and Meier, 1997), proved to be impossible to implement in an efficient form due to the limitations of the short rotor-cycle time in comparison with the time requirements of the pulses at the low available rf power. In situations where stronger rf fields are available, rf-driven schemes (Meier, 1994), e.g. RIL (Baldus and Meier, 1997), DRAMA (Tycko and Smith,1993), SEDRA (Gullion and Vega, 1992), may prove to be more efficient than proton-driven spin diffusion.

Experiments on the dipeptide

For a sensitivity and resolution test of the pulse sequences illustrated in Figure 1, experiments were performed on fully $^{13}C/^{15}N$-labelled L-Val-L-Phe, purchased from Cambridge Isotope Laboratories. Fifty milligrams of the sample were ground and packed into a 4 mm rotor.

The experiments were performed on a Bruker DMX-400 NMR spectrometer with a Bruker triple resonance MAS probehead. The spinning speed was set to 14.5 kHz for all experiments and was controlled by a Bruker MAS control unit with a stability of ± 2 Hz. The ramp used for the ^{13}C-^{15}N cross-polarization step was generated using xShape, which is part of the Bruker XWIN-NMR package, using 64 points, with a depth of ± 7 kHz centred at the first sideband Hartmann–Hahn matching condition (Baldus et al., 1996). The proton $\pi/2$ pulse length was 3.1 μs ($\omega_H/2\pi = 80.6$ kHz), the carbon $\pi/2$ pulse length was 4.0 μs ($\omega_C/2\pi = 62.5$ kHz), and the nitrogen $\pi/2$ pulse length was 6.5 μs ($\omega_N/2\pi = 38.5$ kHz). For the chosen spinning speed of 14.5 kHz, it was found that the optimal pulse length for TPPM proton decoupling (Bennett et al., 1995) was 6.4 ms and the optimal phases were $\pm 13°$. The durations of the cross-polarization steps, CP1 and CP2, were optimized using a 1D version of the pulse sequence illustrated in Figure 1a, without the π(sel)-π(non-sel.) pulse pair. The optimal durations were found to be $\tau_{CP1} = 3$ ms and $\tau_{CP2} = 10$ ms. The Gaussian shape for the selective pulse was generated with the program xShape using 256 points with a truncation value of 3%. The amplitude of the pulse was calibrated as previously described in Straus et al. (1996).

For the triple resonance experiments described here, careful filtering of the transmitter rf channels proved to be necessary to avoid losses in S/N. Cross-talk between the different channels could effectively be removed by the following rf filters: a ^1H band pass (-60 dB at ± 150 MHz) in the proton channel, a ^{15}N-stop high-pass filter (-100 dB at ^{15}N frequency) in the carbon channel, and a ^{13}C-stop low-pass filter (-100 dB at ^{13}C frequency) in the nitrogen channel.

To assess the sensitivity of the proposed experiments, it is necessary to know the S/N performance of the ^{13}C and the ^{15}N channels in single-pulse exper-

iments. Theoretically, one expects a ratio (Hoult and Richards, 1976)

$$\frac{(S/N)_{^{13}C}}{(S/N)_{^{15}N}} = \frac{\gamma_{^{13}C}^{5/2} T_{2,^{13}C}}{\gamma_{^{15}N}^{5/2} T_{2,^{15}N}} \quad (1)$$

when no apodization functions (which would allow an optimization of the sensitivity on account of resolution) are applied to the FIDs. From the line widths in the experimental data of Figure 2, one finds for $^{13}C^{\alpha}$(Phe) and ^{15}N(Phe) a ratio $\Delta\nu(^{15}N)/\Delta\nu(^{13}C) = T_{2,^{13}C}/T_{2,^{15}N} = 44$ Hz/134 Hz ≈ 0.328 and for the expected ratio $(S/N)_{^{13}C}/(S/N)_{^{15}N} \approx 3.18$. This compares with an experimental sensitivity ratio of 3.5 (Figure 2), which shows that the ^{13}C channel of the spectrometer is 10% more sensitive than the ^{15}N channel.

When comparing the sensitivities of experiments starting with a Hartmann–Hahn cross-polarization step from ^{1}H to either ^{13}C or ^{15}N, the ratio of S/N ratios is modified to

$$\left[\frac{(S/N)_{^{13}C}}{(S/N)_{^{15}N}}\right]_{CP} = \frac{\gamma_{^{13}C}^{3/2} T_{2,^{13}C}}{\gamma_{^{15}N}^{3/2} T_{2,^{15}N}} f \quad (2)$$

where the factor f takes into account the measured ratio of channel sensitivities of $f = 1.10$. This leads to an expected ratio

$$\left[\frac{(S/N)_{^{13}C}}{(S/N)_{^{15}N}}\right]_{CP} = 1.41 \quad (3)$$

The performance of the pulse sequences in Figure 1 regarding sensitivity and resolution is evaluated based on the following six 2D experiments:
(i) $^{15}N(\omega_1)$-$^{13}C(\omega_2)$ correlation experiment (NCA);
(ii) $^{13}C(\omega_1)$-$^{15}N(\omega_2)$ correlation experiment (CAN);
(iii) $^{13}C^{\alpha}(\omega_1)$-$^{15}N(\omega_2)$ correlation experiment, with homonuclear decoupling on $^{13}C^{\alpha}$ (CAN – Figure 1a);
(iv) $^{15}N(\omega_1)$-^{13}CO(sel)-$^{13}C(\omega_2)$ correlation experiment (N_{i+1}(CO)CA_i – Figure 1b);
(v) $^{13}C(\omega_1)$-^{13}CO(sel)-$^{15}N(\omega_2)$ correlation experiment (CA_i(CO)N_{i+1});
(vi) $^{13}C^{\alpha}(\omega_1)$-^{13}CO(sel)-$^{15}N(\omega_2)$ correlation experiment, with homonuclear decoupling on $^{13}C^{\alpha}$ (CA_i(CO)N_{i+1} – Figure 1c).

All data sets were acquired with the same experiment time of 10 h and 50 min. The sweep widths and number of t_1 increments were chosen such that the observed line widths are not broadened by truncation effects. The t_1 and t_2 times for ^{15}N evolution were 36.6 ms. For ^{13}C without homonuclear decoupling, an evolution time of 12.8 ms was selected, and with homonuclear decoupling 24.4 ms was selected. All spectra were processed without apodization functions to ensure that the S/N ratios were comparable. The signal amplitude, S, was determined from the intensity of the cross peak in the slice through the 2D spectrum with the maximum intensity. The noise amplitude, N, was measured as the rmsd from a slice of the spectrum without cross peaks. The S/N values (Ernst et al., 1987) are summarized in Table 1.

As can be seen from the values in Table 1, detection on ^{15}N yields an S/N ratio worse by a factor of ca. 1.5 ((i)/(iii) and (iv)/(vi) in Table 1) in comparison to ^{13}C detection for all experiments. This agrees well with the expected value of Eq. 3. The line widths achieved with the ^{15}N-detected experiments, however, are narrower by a factor of 2. Often the slight loss in sensitivity is more than counterbalanced by the gain in resolution. Comparing the pulse schemes in Figures 1b and c, one notices that during ^{13}C detection in Figure 1b no ^{15}N decoupling is applied while GARP-1 ^{13}C decoupling is used in Figure 1c. In fact, ^{15}N decoupling hardly would affect the line width due to the relatively broad $^{13}C^{\alpha}$ resonance lines and the weak coupling $J_{C^{\alpha}-N} \approx 5$ Hz. On the other hand, ^{13}C GARP-1 decoupling during ^{15}N detection has a significant line narrowing effect because the ^{15}N lines are inherently narrower and a larger coupling ($J_{CO-N} \approx$ 10–15 Hz) is involved as well. This leads to a line narrowing from 40 Hz to 30 Hz by ^{13}C GARP-1 decoupling.

The intensity of cross peaks arising from polarization transfers between non-bonded heteronuclei was found to be at least 10 times smaller than the one-bond cross peaks for all experiments and mixing times presented here.

Partial assignment of ubiquitin

To explore the feasibility of assigning proteins in the solid state, the pulse sequences described in the 'Assignment strategy and pulse sequences' section were applied to a sample of fully ^{13}C/^{15}N-labelled human ubiquitin, M1-Q2-I3-F4-V5-K6-T7-L8-T9-G10-K11-T12-I13-T14-L15-E16-V17-E18-P19-S20-D21-T22-I23-E24-N25-V26-K27-A28-K29-I30-Q31-D32-K33-E34-G35-I36-P37-P38-D39-Q40-Q41-R42-L43-I44-F45-A46-G47-K48-Q49-L50-E51-D52-G53-R54-T55-L56-S57-D58-Y59-N60-I61-Q62-K63-E64-S65-T66-L67-H68-L69-V70-L71-R72-L73-R74-G75-G76,

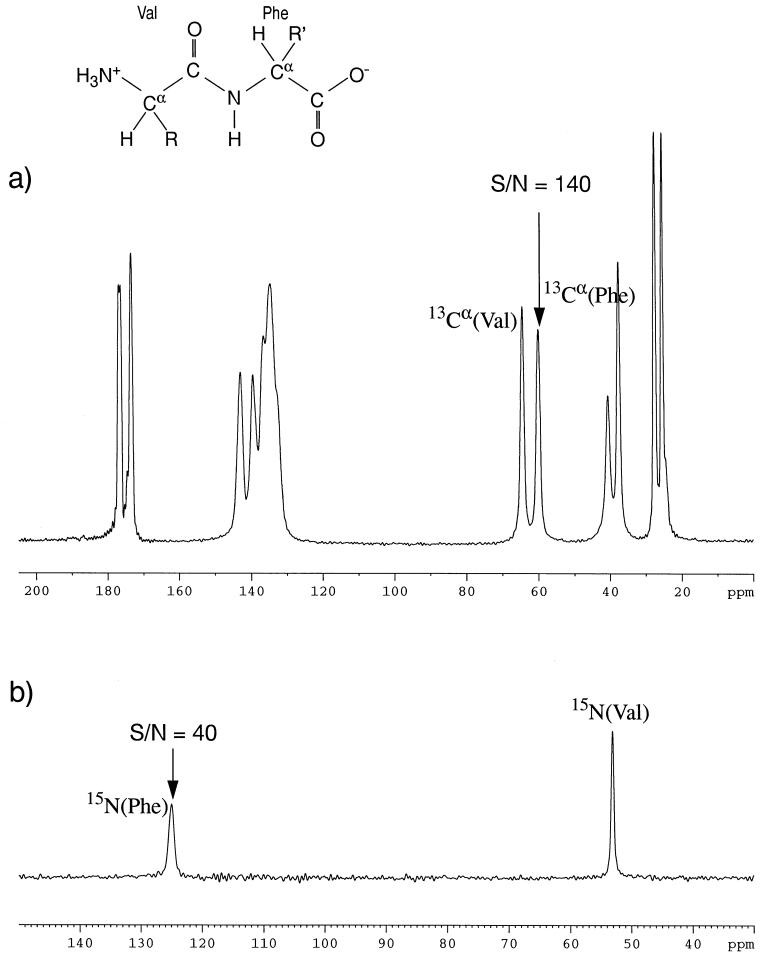

Figure 2. 1D single-pulse spectra (no cross polarization from protons) of the fully $^{13}C/^{15}N$-labelled dipeptide L-Val-L-Phe: (a) ^{13}C spectrum; (b) ^{15}N spectrum. Both spectra were acquired with 32 scans and a recycle delay of 2 min.

which was kindly provided by A.J. Wand (State University of New York at Buffalo, U.S.A.). Twenty milligrams of the sample (prepared and purified by lyophilization as described by Ecker et al. (1987)) were ground and packed into a 4 mm rotor. The sample was equilibrated in a water-saturated atmosphere and the rotor was stored in a moist container. The line widths obtained for the sample were $^{13}C^{\alpha}$ FWHH = 120–135 Hz (with TPPM proton decoupling but without homonuclear decoupling) and ^{15}N FWHH = 100–105 Hz (with TPPM proton decoupling but without ^{13}C decoupling).

The intra-residue assignment of the backbone atoms is based on NCA (Figure 3a) and CAN spectra (Figure 3b) which relate the ^{15}N backbone resonances to the adjacent $^{13}C^{\alpha}$ spin. As expected, the CAN experiment which allows for homonuclear ^{13}C-^{13}C decoupling shows better resolution than the NCA experiment. For the identification of the amino acid residues, a ^{13}C-^{13}C correlation spectrum by proton-driven spin diffusion at an extended mixing time of $\tau_m = 20$ ms, which favours multiple-step transfers (Figure 4), is useful. The inter-residue assignment takes advantage of $N_{i+1}(CO)CA_i$ (Figure 5a) and $CA_i(CO)N_{i+1}$ spectra (Figure 5b). Note that no cross peaks arising from polarization transfers between non-directly bonded heteronuclei were found.

Note that the ^{15}N-detected $CA_i(CO)N_{i+1}$ spectrum in Figure 5b also shows increased resolution compared to Figure 5a, visible in particular in the $^{13}C_i^{\alpha}(Gly) \rightarrow {}^{15}N_{i+1}(Xxx)$ region marked with a box. All four expected cross peaks (Gly^{10}-Lys^{11}, Gly^{35}-Ile^{36}, Gly^{47}-Lys^{48}, Gly^{53}-Arg^{54}) are resolved, in contrast to the same region in the $N_{i+1}(CO)CA_i$ spectrum

Table 1. Experimental S/N ratios and line widths of the $^{13}C^{\alpha}$ resonance taken from (a) the $^{13}C_i^{\alpha}$(Phe)-$^{15}N_i$(Phe) cross peaks in 2D heteronuclear correlation spectra using NCA and CAN (illustrated in Figure 1a) and (b) the $^{13}C_i^{\alpha}$(Val)-$^{15}N_{i+1}$(Phe) cross peaks in spectra using the sequences of 2D heteronuclear correlation spectra using the sequences of Figures 1b and c for fully $^{13}C/^{15}N$-Iabelled L-Val-L-Phe. All values are given for identical experiment times of 10 h and 50 min.

(a)	Experiment	Detected nucleaus	S/N[a]	Phe-$^{13}C^{\alpha}$ FWHH[b] (Hz)
(i)	NCA	^{13}C	265[c]	128.1
(ii)	CAN	^{15}N	150	128.2
(iii)	CAN (with ^{13}C homonuclear decoupling)	^{15}N	180	64.9
(b)	Experiment	Detected nucleus	S/N[a]	Val-$^{13}C^{\alpha}$ FWHH[b] (Hz)
(iv)	$N_{i+1}(CO)CA_i$	^{13}C	60	110.0
(v)	$CA_i(CO)N_{i+1}$	^{15}N	40	110.1
(vi)	$CA_i(CO)N_{i+1}$ (with ^{13}C homonuclear decoupling)	^{15}N	43	59.8

[a]See text for description on how it was determined.
[b]Line width measured as the full width at half height (FWHH).
[c]Error in all S/N ratios has a standard deviation of 5%.

in Figure 5a. However, while for smaller systems the $CA_i(CO)N_{i+1}$ experiment is advantageous as has been shown in the 'Experiments on the dipeptide' section of this paper, for proteins like ubiquitin, this is not the case since the S/N is low when using ^{15}N detection. Some of the cross peaks visible in Figure 5a, which will be used for assignment, are hidden in the noise in Figure 5b, indicating some of the limitations of the $CA_i(CO)N_{i+1}$ experiment. Therefore, only the spectrum in Figure 5a will be used in the following.

The limited resolution of the spectra in Figures 3–5 suggests that a full assignment of ubiquitin based on these spectra alone is not yet feasible. However, there are a sizeable number of well-resolved peaks that allow for a partial assignment. Because only fragments can be addressed, it is understandable that much of the amino acid identification has to be based on characteristic chemical shift ranges known from the study of other proteins (Wishart et al., 1991). We demonstrate in the following the assignment strategy for Pro[19], Pro[37], and Pro[38], followed by the assignment of Ala[28] and Ala[46] and of a few further residues.

Proline assignment

The most revealing data for the identification of proline residues are contained in the CAN spectrum of Figure 3b. Proline residues differ from other amino acids by the fact that the backbone nitrogen is adjacent to C^{α} and C^{δ}. Because the $^{13}C^{\delta}$ chemical shifts are not far from the $^{13}C^{\alpha}$ shifts, both are excited by the selective pulse in the sequence of Figure 1a, and each ^{15}N(Pro) must be correlated with two ^{13}C resonances. This leads to the identification of the three proline residues in Figure 3b. Cross checks are possible by the expected chemical shift ranges for $^{13}C^{\delta}$(Pro) ($47.5 < \delta_{^{13}C^{\delta}(Pro)} < 52$ ppm) and $^{13}C^{\alpha}$(Pro) $58.7 < \delta_{^{13}C^{\alpha}(Pro)} < 67$ ppm) as well as by the appearance of spin-diffusion cross peaks $^{13}C^{\alpha} \rightarrow {}^{13}C^{\delta}$ in the spectrum of Figure 4.

The assignment of the three proline residues is based on the $N_{i+1}(CO)CA_i$ spectrum of Figure 5a. From the primary sequence, one finds that Pro[19] is a neighbour to Ser[20] and that Pro[37] is flanked by another proline, Pro[38], which is itself a neighbour to Asp[39]. Having identified the $^{15}N(Pro^{38})$-$^{13}C^{\alpha}(Pro^{37})$ cross peak in Figure 5a, also the assignment of Pro[19] is firm. The assignment of ^{15}N and $^{13}C^{\alpha}$ of all three prolines allows then also to find the inter-residue cross peaks $^{15}N(Pro^{19})$-$^{13}C^{\alpha}(Glu^{18})$, $^{15}N(Ser^{20})$-$^{13}C^{\alpha}(Pro^{19})$, $^{15}N(Pro^{37})$-$^{13}C^{\alpha}(Ile^{36})$, and $^{15}N(Asp^{39})$-$^{13}C^{\alpha}(Pro^{38})$ in Figure 5a. This leads then to the assignments of the cross peak $^{13}C^{\alpha}(Ser^{20})$-$^{15}N(Ser^{20})$ in Figure 3b and of $^{13}C^{\alpha}(Ser^{20})$-$^{13}C^{\beta}(Ser^{20})$ in Figure 4.

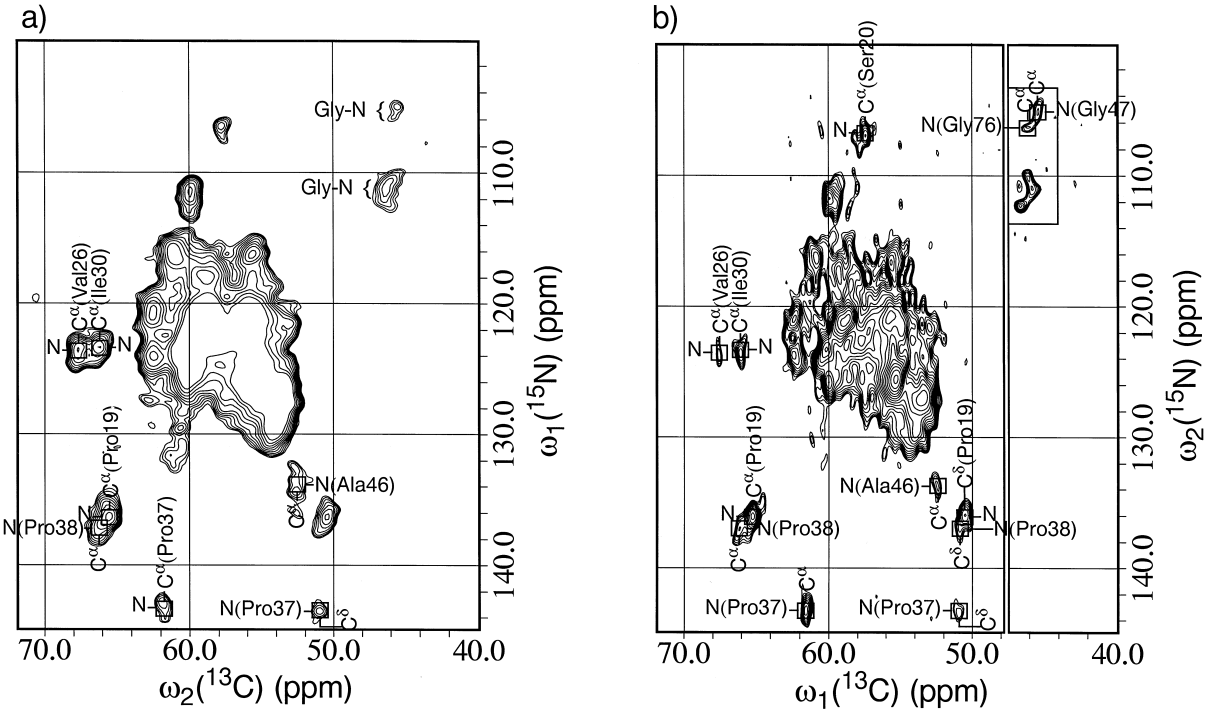

Figure 3. $^{13}C^{\alpha}$-^{15}N cross-peak regions from heteronuclear 2D correlation spectra of $^{13}C/^{15}N$-labelled ubiquitin using (a) the NCA pulse sequence (according to Sun et al. (1995) and Baldus et al. (1996)) (experiment time = 2.0 days; number of scans = 156) and (b) the CAN pulse sequence as presented in Figure 1a, with homonuclear decoupling in ω_1, (experiment time = 4.3 days; number of scans = 384). Note that the region to the left in the spectrum in (b) was obtained from an experiment with the carrier of the selective pulse placed at 58 ppm while for the right region it was moves to 45 ppm. The spectrum in (b) was mirrored so that the displays in (a) and (b) match. The low-power Gaussian pulse length was set to 500 μs. The cross-polarization times were set to $\tau_{CP1} = 3$ ms and $\tau_{CP2} = 10$ ms, respectively. The recycle delay was 3.5 s. The spectra, referenced with respect to TMS and $^{15}NH_4Cl$, are plotted at the same amplitude levels, with the contour levels increasing by the factor 1.1.

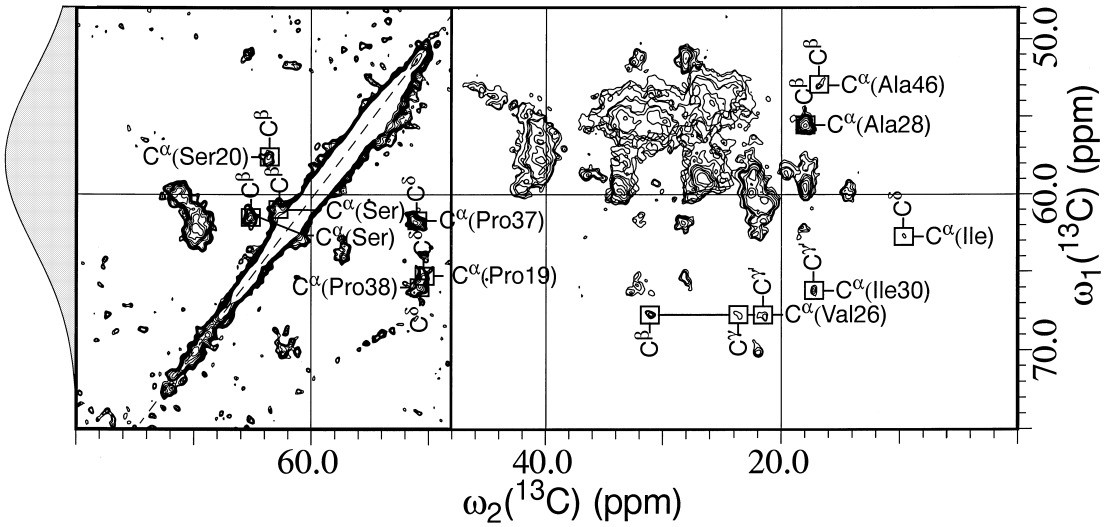

Figure 4. 2D ^{13}C-^{13}C correlation spectra of $^{13}C/^{15}N$-labelled ubiquitin, using a proton-driven spin-diffusion sequence for mixing and homonuclear decoupling in ω_1, (Straus et al., 1996,1997). The mixing time, τ_m, was set to 20 ms and the length of the low-power Gaussian pulse was 1 ms (the inversion profile is schematically drawn on the left side). Thirty-two transients were added, resulting in an experiment time of 19 h (recycle delay = 3.1 s). The spectrum, referenced with respect to TMS, is plotted with contour levels increasing by the factor 1.1. The region to the left of the spectrum was plotted at a lower level such that the cross peaks on both sides of the diagonal are visible.

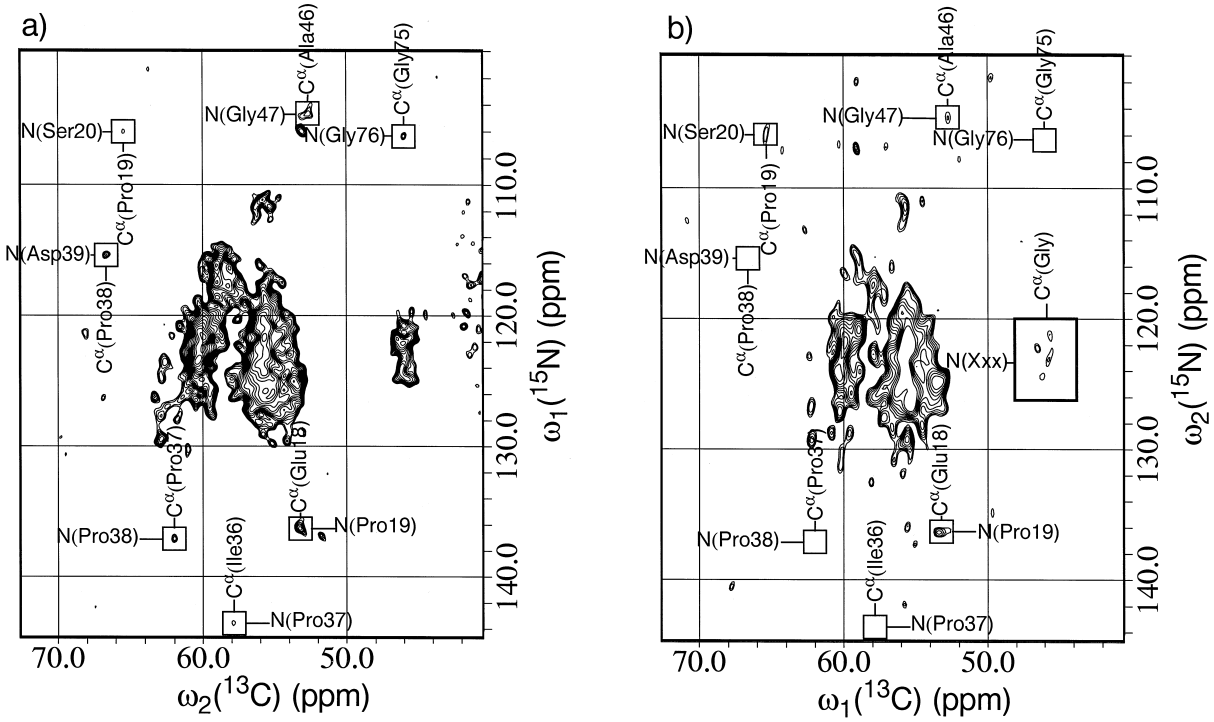

Figure 5. Sections of the heteronuclear 2D correlation spectra of $^{13}C/^{15}N$-Iabelled ubiquitin using (a) the $N_{i+1}(CO)CA_i$ pulse sequence of Figure 1b and (b) the $CA_i(CO)N_{i+1}$ pulse sequence of Figure 1c. The experiment time was 5.3 days in (a) (number of scans = 640; recycle delay = 3.5 s) and 7.1 days in (b) (number of scans = 1200; recycle delay = 3.0 s). The cross-polarization times were set to $\tau_{CP1} = 3$ ms and $\tau_{CP2} = 10$ ms, respectively, and the mixing time to $\tau_m = 12$ ms. The spectra, referenced with respect to TMS and $^{15}NH_4Cl$, are plotted at the same amplitude levels, with the contour levels increasing by the factor 1.1.

Alanine assignment

The assignment of the two alanine residues in ubiquitin can start from a carbon spin-diffusion spectrum recorded with a sufficiently short mixing time ($\tau_m = 6$ ms) to allow only one-step transfers. In the cross-peak region characteristic for $^{13}C^{\alpha}$ shifts $50 < \delta_{^{13}C^{\alpha}} < 70$ ppm and for methyl groups $10 < \delta_{^{13}CH_3} < 20$ ppm, one finds only two cross peaks at (55.5 ppm, 18.0 ppm) and at (52.9 ppm, 16.8 ppm). Based on the chemical shift, one can inspect the CAN spectrum of Figure 3b and identify a $^{13}C^{\alpha}$-^{15}N peak at (52.9 ppm, 133.8 ppm). The other alanine $^{13}C^{\alpha}$-^{15}N resonance is not resolved.

Using the $N_{i+1}(CO)CA_i$ spectrum, one can then proceed to the sequential assignment of the two alanines. One of the alanines, Ala^{46}, is a neighbour to a glycine, Gly^{47}. Since glycines have characteristic ^{15}N chemical shifts which are typically around 110 ppm (see Figure 3a or b), the corresponding cross peak $^{15}N(Gly^{47})$-$^{13}C^{\alpha}(Ala^{46})$ can easily be identified in the $N_{i+1}(CO)CA_i$ spectrum using the alanine chemical shifts determined from Figure 4 and the glycine shifts of Figure 3b. A full assignment of the residues Ala^{46} and Gly^{47}, and by exclusion also of Ala^{28}, is then possible (see Table 2). Using the chemical shift–secondary structure relationship (Wishart et al., 1991), one may even conclude that Ala^{28} must be in a helix due to its high $^{13}C^{\alpha}$ chemical shift, while Ala^{46} must be in a β-sheet or in a loop region. This is in agreement with the tertiary structure of human ubiquitin as determined by X-ray crystallography (Vijay-Kumar et al., 1987).

Further assignment

Using the above strategy and the correlation experiments as given in Figure 1, further assignments are possible and are also listed in Table 2. Three glycine residues in ubiquitin (Gly^{47}, Gly^{75}, Gly^{76}) can be assigned, and also one of the glutamic acids (Glu^{18}), a valine (Val^{26}), a lysine (Lys^{27}), two of the seven isoleucines (Ile^{30} and Ile^{36}) and an asparagine (Asp^{39}) can be identified. The lower half of Table 2 gives the ^{13}C chemical shifts of the remaining three valine residues in ubiquitin (Val^{5}, Val^{17}, and Val^{70}), two

Table 2. ^{15}N and ^{13}C assignments[a] of some amino acid residues in ubiquitin. Unless explicitly indicated in brackets, the error is estimated to be ±2 ppm[b]

Residue	^{15}N	$^{13}C^{\alpha}$	^{13}CO	Side-chain ^{13}C's
Glu^{18}	n.r.[c]	53.1	n.r.	n.r.
Pro^{19}	136.0	65.6	175.7	C^{δ} 50.8
Ser^{20}	106.0	57.7	174.8 (0.6)	C^{β} 63.6
Val^{26}	123.6	67.8	178.4 (0.6)	C^{β} 31.4
				C^{γ} 23.9; C^{γ} 22.0
Lys^{27}	121.2	n.r.	n.r.	n.r.
Ala^{28}	n.r.	55.6	179.8	C^{β} 18.0
Ile^{30}	123.4	66.2	178.7 (0.6)	C^{β} 37.4
				C^{γ} 31.5; $C^{\gamma'}$ 17.2
				C^{γ} 15.5
Gln^{31}	126.0	n.r.	n.r.	n.r.
Ile^{36}	n.r.	58.1	n.r.	n.r.
Pro^{37}	143.4	61.8	173.4	C^{δ} 51.2
Pro^{38}	137.1	66.3	175.9	C^{δ} 51.1
Asp^{39}	115.4	n.r.	n.r.	n.r.
Ala^{46}	133.8	52.9	177.7 (0.4)	C^{β} 16.8
Gly^{47}	105.0	45.6	n.r.	–
Gly^{75}	111.3	46.3	n.r.	–
Gly^{76}	106.2	45.8	n.r.	–
Val$^{5/17/70}$	n.r.	58.8	n.r.	C^{β} 36.7
				C^{γ} 22.6; C^{γ} 20.0
Val$^{5/17/70}$	n.r.	60.7 (0.6)	n.r.	C^{β} 35.3
				C^{γ} 21.8; C^{γ} 21.1 (0.4)
Val$^{5/17/70}$	n.r.	61.0 (0.6)	n.r.	C^{β} 34.1
				C^{γ} 21.4; C^{γ} n.r.
Ile$^{3/13/23/44/61}$	n.r.	62.6	n.r.	C^{β} 37.1
				C^{γ} 28.1; $C^{\gamma'}$ 18.4
				C^{δ} 9.6
Ile$^{3/13/23/44/61}$	n.r.	59.7	n.r	C^{β} 42.0
				C^{γ} 25.4; $C^{\gamma'}$ 18.4
				C^{δ} 14.8
Ser$^{57/65}$	n.r.	61.5	172.6 (0.4)	C^{β} 65.1
Ser$^{57/65}$	n.r.	61.0	n.r.	C^{β} 62.8

[a]Shifts in ppm, with respect to external adamantane (δ_{CH} = 29.5 ppm with respect to TMS) for the ^{13}C resonances and external $^{15}NH_4Cl$ (δ_N = 24.9 ppm) for the ^{15}N resonances.
[b]See Straus et al. (1997) for the error estimation.
[c]Not resolved.

isoleucines (Ile^3, Ile^{13}, Ile^{23}, Ile^{44} or Ile^{61}) and two serines (Ser^{57} and Ser^{65}) which however could not be positioned uniquely in the amino acid sequence. In total, 16 residues of ubiquitin (∼20%) can be sequentially assigned without recourse to liquid state NMR. Comparing the chemical shifts obtained in this study to those obtained in the liquid state (Wand et al., 1996), one finds that the differences ($\delta_{liquid} - \delta_{solid}$) lie within a range of ±1ppm, except for the terminal glycines Gly^{75} and Gly^{76}.

Conclusions

This study demonstrates that by a combination of advanced NMR techniques it is possible to at least par-

tially assign proteins in the solid state. It is sufficient to have fully $^{13}C/^{15}N$-labelled proteins available. No selective labelling is required. This considerably facilitates the practical applicability and widens the range of problems treated in this manner.

It is found that the most useful experiment is a heteronuclear inter-residue correlation using an $N_{i+1}(CO)CA_i$ sequence or possibly a $CA_i(CO)N_{i+1}$ experiment. The latter one as well as the intra-residue correlation experiment CAN and the ^{13}C-^{13}C spin-diffusion experiment allow for high resolution by homonuclear ^{13}C-^{13}C decoupling during the evolution period. A combination of all these experiments is found to be a powerful tool for assignments in the solid state.

Obviously the goal of a complete assignment of a medium-sized protein in the solid state is still remote. But the present work gives us the hope that by further methodological advances a full assignment is foreseeable. In particular, the usage of (unfortunately time-consuming) 3D (Tycko, 1996; Sun et al., 1997) and possibly 4D correlation methods should improve the spectral resolution to such an extent that a sufficient number of peaks can be resolved to complete the assignment. Work in this direction is in progress.

Acknowledgements

The $^{13}C/^{15}N$-labelled sample of ubiquitin was kindly provided by Prof. A.J. Wand (State University of New York, Buffalo, U.S.A.). Helpful discussions with Prof. Dieter Suter and Dr. Sabine Hediger are gratefully acknowledged. S.K.S. would like to thank NSERC for their financial support. This work has been supported by the Swiss National Science Foundation.

References

Andrew, E.R., Bradbury, A., Eades, R.G. and Wynn, T. (1963) *Phys. Lett.*, **4**, 99–100.

Baldus, M., Geurts, D.G., Hediger, S. and Meier, B.H. (1996) *J. Magn. Reson.*, **A118**, 140–144.

Baldus, M. and Meier, B.H. (1997) *J. Chem. Phys.*, submitted.

Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) *J. Magn. Reson.*, **88**, 425–431.

Bennett, A.E., Rienstra, C.E., Auger, M., Lakshmi, K.V. and Griffin, R.G. (1995) *J. Chem. Phys.*, **103**, 6951–6958.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, CA.

Clore, G.M., Bax, A., Driscoll, P.C., Wingfield, P.T. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 8172–8184.

Clore, G.M. and Gronenborn, A.M. (Eds.) (1993) *NMR of Proteins*, Macmillan, London.

Colombo, M.G., Meier, B.H. and Ernst, R.R. (1987) *Chem. Phys. Lett.*, **146**, 189–196.

Croasmun, W.R. and Carlson, R.M.K. (Eds.) (1994) *Two-dimensional NMR Spectroscopy, Applications for Chemists and Biochemists*, VCH, Weinheim.

Dotsch, V., Matsuo, H. and Wagner, G. (1996) *J. Magn. Reson.*, **B112**, 95–100.

Ecker, D.J., Muhammad, I.K., Marsk, J., Butt, T.R. and Crooke, S.T. (1987) *J. Biol. Chem.*, **262**, 3524–3527.

Ernst, R.R., Bodenhausen, G. and Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon, Oxford.

Griffiths, J.M., Ashburn, T.T., Auger, M., Costa, P.R., Griffin, R.G. and Lansbury Jr., P.T. (1995) *J. Am. Chem. Soc.*, **117**, 3539–3546.

Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **99**, 201–207.

Gullion, T. and Vega, S. (1992) *Chem. Phys. Lett.*, **194**, 423–428.

Hediger, S., Meier, B.H., Kurur, N.D., Bodenhausen, G. and Ernst, R.R. (1994) *Chem. Phys. Lett.*, **223**, 283–288.

Hediger, S., Meier, B.H. and Ernst, R.R. (1995) *Chem. Phys. Lett.*, **240**, 449–456.

Heller, J., Kolbert, A.C., Larsen, R., Ernst, M., Bekker, T., Baldwin, M., Prusiner, S.B., Pines, A. and Wemmer, D.E. (1996) *Protein Sci.*, **5**, 1655–1661.

Hirsh, D.J., Hammer, J., Maloy, W.L., Blazyk, J. and Schaefer, J. (1996) *Biochemistry*, **35**, 12733–12741.

Hoult, D.I. and Richards, R.E. (1976) *J. Magn. Reson.*, **24**, 71–85.

Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.

Kay, L.E., Ikura, M., Guang, Z. and Bax, A. (1991) *J. Magn. Reson.*, **91**, 422–428.

Kay, L.E. (1995) *Prog. Biophys. Mol. Biol.*, **63**, 277–299.

Lansbury Jr., P.T., Costa, P.R., Griffiths, J.M., Simon, E.J., Auger, M., Halverson, K.J., Kociska, D.A., Hendsch, Z.S., Ashburn, T.T., Spencer, R.G.S., Tidor, B. and Griffin, R.G. (1995) *Nat. Struct. Biol.*, **2**, 990–998.

Lee, M. and Goldburg, W.I. (1965) *Phys. Rev.*, **A140**, 1261–1271.

Meier, B.H. (1994) *Adv. Magn. Opt. Reson.*, **18**, 1–116.

Metz, G., Xiaoling, W. and Smith, S.O. (1994) *J. Magn. Reson.*, **A110**, 219–227.

Naito, A., Root, A. and McDowell, C.A. (1991) *J. Phys. Chem.*, **95**, 3578–3581.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.

Straus, S.K., Bremi, T. and Ernst, R.R. (1996) *Chem. Phys. Lett.*, **262**, 709–715.

Straus, S.K., Bremi, T. and Ernst, R.R. (1997) *J. Biomol. NMR*, **10**, 119–128.

Sun, B.Q., Costa, P.R. and Griffin, R.G. (1995) *J. Magn. Reson.*, **A112**, 191–198.

Sun, B.Q., Rienstra, C.M., Costa, P.R., Williamson, J.R. and Griffin, R.G. (1997) *J. Am. Chem. Soc.*, **119**, 8540–8546.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

Tycko, R. and Smith, S.O. (1993) *J. Chem. Phys.*, **98**, 932–943.

Tycko, R. (1996) *J. Biomol. NMR*, **8**, 239–251.

Vijay-Kumar, S., Bugg, C.E. and Cook, W.J. (1987) *J. Mol. Biol.*, **194**, 531–544.
Wand, A.J., Urbauer, J.L., McEvoy, R.P. and Bieber, R.J. (1996) *Biochemistry*, **35**, 6116–6125.
Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.*, **222**, 311–333.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.