Journal of Biomolecular NMR, **15:** 1–14, 1999.
KLUWER/ESCOM

Resonance assignment of $^{13}C/^{15}N$ labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR

Mei Hong*
Department of Polymer Science & Engineering, University of Massachusetts, Amherst, MA 01003, U.S.A.

Received 4 May 1999; Accepted 23 June 1999

Key words: isotopic labeling, multidimensional correlation, proteins, resonance assignment, sequential connectivity, solid-state NMR

Abstract

The comprehensive structure determination of isotopically labeled proteins by solid-state NMR requires sequence-specific assignment of ^{13}C and ^{15}N spectra. We describe several 2D and 3D MAS correlation techniques for resonance assignment and apply them, at 7.0 Tesla, to ^{13}C and ^{15}N labeled ubiquitin to examine the extent of resonance assignments in the solid state. Both interresidue and intraresidue assignments of the ^{13}C and ^{15}N resonances are addressed. The interresidue assignment was carried out by an N(CO)CA technique, which yields N_i-$C\alpha_{i-1}$ connectivities in protein backbones via two steps of dipolar-mediated coherence transfer. The intraresidue connectivities were obtained from a new 3D NCACB technique, which utilizes the well resolved Cβ chemical shift to distinguish the different amino acids. Additional amino acid type assignment was provided by a ^{13}C spin diffusion experiment, which exhibits ^{13}C spin pairs as off-diagonal intensities in the 2D spectrum. To better resolve carbons with similar chemical shifts, we also performed a dipolar-mediated INADEQUATE experiment. By cross-referencing these spectra and exploiting the selective and extensive ^{13}C labeling approach, we assigned 25% of the amino acids in ubiquitin sequence-specifically and 47% of the residues to the amino acid types. The sensitivity and resolution of these experiments are evaluated, especially in the context of the selective and extensive ^{13}C labeling approach.

Introduction

The determination of the three-dimensional structure of noncrystalline biological solids such as membrane proteins and insoluble protein aggregates by NMR remains an important and challenging problem today. To achieve this goal, it is desirable to have a repertoire of solid-state NMR techniques that are capable of yielding a large number of structural constraints from each experiment and sample. A necessary element in such a methodology is isotopic labeling, used to enhance the signal-to-noise ratios of the NMR spectra. Although site-specific ^{13}C and ^{15}N labeling has been shown to allow the measurement of structural constraints with high precision (Creuzet et al., 1991; Shon et al., 1991; Ketchem et al., 1993; McDowell and Schaefer, 1996; Weliky and Tycko, 1996; Feng et al., 1997; Long et al., 1998), it is an inefficient process, requiring major synthetic efforts and yielding only a limited amount of information per experiment. Therefore, increasing attention has recently been directed to uniform $^{13}C/^{15}N$ labeling of proteins and to the design of experiments that yield multiple distances or torsion-angle constraints. However, before structural measurements on these uniformly labeled proteins can be evaluated, their NMR spectra need to be assigned to the amino acid sequences of the proteins.

Resonance assignment of proteins in solution is currently accomplished by a number of multidimensional correlation techniques that encode the chemical shift frequencies of various spins during different time periods and utilize scalar couplings to transfer the coherence between these spins (Wagner and Wüthrich,

*To whom correspondence should be addressed at: Department of Chemistry, Iowa State University, Ames, IA 50011, U.S.A. E-mail: mhong@iastate.edu

1982; Ernst et al., 1987; Clore and Gronenborn, 1991; Grzesiek and Bax, 1992a; Bax, 1994; Cavanagh et al., 1996). In principle, analogous correlation schemes can also be employed in designing multidimensional magic-angle spinning (MAS) experiments to achieve resonance assignment of solid proteins (Tycko, 1996). In the solid state, one can use not only scalar couplings but also dipolar couplings for coherence transfer (Baldus and Meier, 1996; Lesage et al., 1997; Hong and Griffin, 1998). In fact, due to its sharp distance dependence ($1/r^3$), the dipolar interaction can distinguish one-bond connectivities from multiple-bond connectivities quite effectively. Under MAS, heteronuclear and homonuclear dipolar couplings can be reintroduced by various radio frequency (rf) sequences (Baldus et al., 1994; Bennett et al., 1994; Nielsen et al., 1994; Gregory et al., 1995; Sun et al., 1995; Griffin et al., 1998; Hohwy et al., 1998; Rienstra et al., 1998).

In this paper, we demonstrate four 2D and 3D MAS correlation techniques that together yield the resonance assignment of the 76-residue protein, ubiquitin, to a significant extent. We first describe a 2D N(CO)CA correlation technique that manifests sequential $C\alpha_{i-1}$-N_i connectivities along the protein backbone. The technique is formally complementary to, but simpler than, the ^{15}N-^{13}C correlation experiment introduced recently (Hong and Griffin, 1998). A similar N(CO)CA experiment was recently published by Straus et al. (1998); however, the detailed implementations of the two experiments differ, with substantially higher sensitivities achieved by the current experimental design. Since more complete assignment requires better site resolution than is feasible by 2D spectroscopy alone, we introduce a 3D NCACB technique that allows amino acid type assignment by utilizing the Cβ chemical shifts. The ^{15}N-^{13}Cα cross sections of the 3D spectrum display distinct peaks that were unresolved in the 2D ^{15}N-^{13}C spectra. On the other hand, the ^{13}Cα-^{13}Cβ cross sections of the 3D NCACB spectrum can be compared to a 2D ^{13}C homonuclear correlation spectrum, obtained via ^{1}H-driven spin diffusion, which reveals ^{13}C connectivities as off-diagonal intensities. We show that additional amino acid type assignment from this experiment can be extracted based on the chemical shifts of carbons further down the side chains. To better resolve and assign carbons with similar chemical shifts, which appear close to the diagonal of the spin diffusion spectrum, we also employ a double-quantum INADEQUATE experiment.

These heteronuclear and homonuclear correlation experiments are demonstrated on two ubiquitin samples. Both are uniformly labeled in ^{15}N, but while one is uniformly labeled in ^{13}C, the other is selectively and extensively labeled in ^{13}C (Hong, 1999a; Hong and Jakes, 1999). The selective ^{13}C labeling was achieved by using a specifically ^{13}C-labeled glycerol, [2-^{13}C] glycerol, as the sole carbon source in the defined media for protein expression. By breaking the ^{13}C spin network with unlabeled carbon sites, selective and extensive ^{13}C labeling reduces the effects of one-bond ^{13}C-^{13}C scalar couplings (∼55 Hz) and the multi-spin dipolar couplings, thereby enhancing the resolution of the NMR spectra. As analyzed previously, the spectral linewidth is the critical factor determining the extent and uniqueness of resonance assignment (Tycko, 1996). We recently showed that due to the selective and extensive ^{13}C labeling, relatively well resolved spectra of proteins can be obtained, even when the samples were unoriented and uncrystallized, and the experiments were carried out under moderate spinning speeds (∼7 kHz) and at a moderate field strength (7.0 Tesla) (Hong, 1999a; Hong and Jakes, 1999). We also determined the labeling levels of each labeled carbon semi-quantitatively, using solution ^{13}C-^{1}H correlation NMR. Most of these labeling levels agree well with the theoretical prediction based on the biosynthetic pathways (Tables 1 and 2 in (Hong, 1999a)). In this paper, we address some of the unique aspects of resonance assignment on such a selectively and extensively ^{13}C-labeled protein. We show that knowledge of the labeling patterns and labeling levels affords additional information on resonance assignment that would be difficult to extract from the spectra of fully ^{13}C-labeled proteins.

Pulse sequences for resonance assignment

N(CO)CA 2D correlation

We showed recently that sequential connectivities between an amide nitrogen of residuei and the Cα site of residue $i-1$ can be obtained using a time-dependent, dipolar-mediated ^{15}N-^{13}C correlation experiment (Hong and Griffin, 1998). At long dipolar coherence transfer times, weak two-bond N_i-$C\alpha_{i-1}$ cross peaks between two adjacent amino acid residues can be observed in conjunction with strong one-bond N_i-$C\alpha_i$ cross peaks. But the dominant intraresidue cross peaks are of little utility for sequence-specific assignment, thus an experiment that selectively removes

these N_i- $C\alpha_i$ resonances will simplify the spectra and provide better site resolution. The N(CO)CA experiment is designed to select sequential N_i-$C\alpha_{i-1}$ cross peaks via two steps of coherence transfer, each involving strong one-bond dipolar couplings. The first step relays the magnetization of amide ^{15}N of residue i to the directly bonded carbonyl carbon of residue $i-1$, while the second transfer connects the CO_{i-1} with $C\alpha_{i-1}$.

The pulse sequence is displayed in Figure 1a and the corresponding phase cycles are listed in Table 1. After cross polarization from ^{1}H to ^{13}C, a train of rotor-synchronized 180° pulses is applied on the ^{13}C and ^{15}N channels to reintroduce the ^{13}C-^{15}N dipolar coupling. This REDOR mixing period (Gullion and Schaefer, 1989), together with a pair of simultaneous 90° pulses, converts the ^{13}C magnetization (C_x) to ^{15}N antiphase magnetization (C_zN_y), which then evolves during the t_1 period. Next, the chemical-shift-modulated ^{15}N transverse magnetization is reconverted to ^{13}C single-quantum coherence by an identical ^{13}C-^{15}N mixing period. Before the end of the second REDOR period, a short C-H dipolar dephasing period of about 30 μs is introduced by turning off the ^{1}H decoupler. This dephasing period destroys the magnetization of the protonated Cα sites while retaining most of the carbonyl intensities. As a result, the intraresidue N_i-$C\alpha_i$ cross peaks are removed from the final spectrum. The selected carbonyl polarization is subsequently transferred to the neighboring Cα via ^{13}C-^{13}C dipolar coupling. We chose the CMR7 sequence (Rienstra et al., 1998) to reintroduce the homonuclear dipolar interaction because of its relative insensitivity to resonance offset and rf field inhomogeneity, but other homonuclear recoupling sequences with comparable performance can also be employed (Hohwy et al., 1998, 1999). By choosing a short C-C mixing time, we can selectively recouple the one-bond ^{13}CO-^{13}Cα spin pairs. During the t_2 period, the Cα signals are detected along with the untransferred carbonyl magnetization. Due to the double-quantum nature of the CMR7 Hamiltonian, the Cα resonances exhibit opposite sign of intensities from the CO signals.

The application of the CMR7 sequence is made possible by the moderate spinning speed used, since the required sevenfold increase in the ^{13}C rf field strength over the spinning speed imposes strong demands on the ^{13}C as well as the ^{1}H rf field strengths. If higher spinning speeds are used, other recoupling sequences with smaller ratios of the rf field to the spinning speed (Sun et al., 1995; Hohwy et al., 1999) or with small rf duty cycles (Bennett et al., 1992) will be more suitable.

During the experiment, the proton carrier frequency is set on resonance for the aliphatic protons, except during the ^{15}N evolution period, when it is shifted 1.3 kHz downfield to be on resonance for the amide proton. A moderate resolution enhancement in the ^{15}N dimension is observed due to this on-resonance ^{1}H decoupling.

NCACB 3D correlation

In the 2D ^{15}N- ^{13}C correlation experiment described previously (Hong, 1999a), the spectral resolution is limited by overlapping signals in the region (^{15}N,^{13}C) = (110–130 ppm, 52–62 ppm). With the exception of Pro, Gly, Ser, Thr, Val, and Ala, which exhibit relatively distinct ^{15}N or ^{13}Cα chemical shifts, most amino acids have their $^{15}N_i$-$^{13}C\alpha_i$ resonances within this region. To reduce resonance overlap, we incorporated a third dimension into the correlation scheme. Sidechain carbons, whose chemical shifts have larger dispersions than the Cα and the carbonyl carbons, are good candidates for the additional dimension. The simplest experiment is an NCACB technique, where, after the ^{15}N and ^{13}Cα chemical shifts are encoded in two indirect dimensions, the magnetization is transferred to Cβ by a homonuclear dipolar recoupling sequence and detected during the t_3 period. The 3D pulse sequence, shown in Figure 1b, resembles that of the N(CO)CA experiment to a large extent. The only differences are the removal of the C-H dephasing period and the reduction of the ^{1}H-^{13}C contact time at the beginning of the sequence. The latter serves to minimize the CO signals while retaining most of the Cα intensities. In this way, we suppress the N_i-CO_{i-1}-$C\alpha_{i-1}$ pathway, which simplifies the spectrum, reduces the necessary ω_2 spectral width, and minimizes the size of the 3D data set.

The ^{13}C carrier frequency is set in the middle of the Cα and Cβ chemical shift region during most of the experiment, except for the ^{13}Cα evolution (t_2) period, when it is shifted by 3.1 kHz to be downfield from the Cα region. This enables off-resonance detection of the aliphatic carbons in the ω_2 dimension, which avoids potential zero-frequency artifacts that would interfere with the real signals.

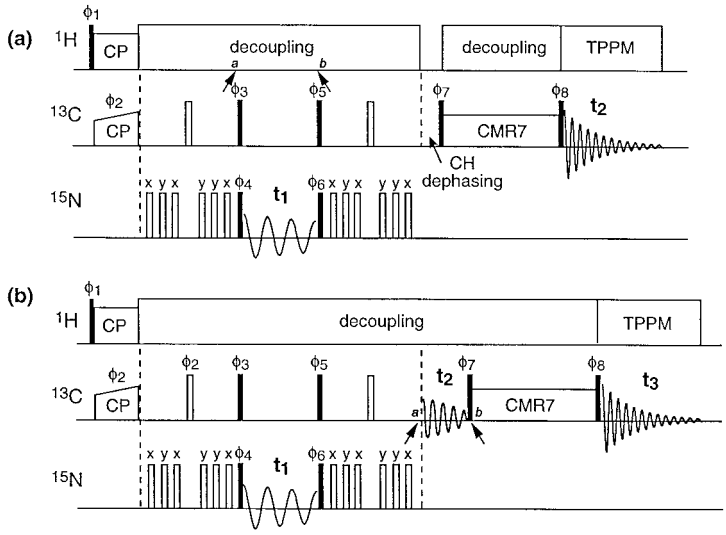

Figure 1. Pulse sequences for resonance assignments. (a) 2D N(CO)CA correlation. (b) 3D NCACB correlation. Open and filled rectangles indicate 180° and 90° pulses, respectively. ^{13}C-^{13}C double-quantum polarization transfer is achieved by the CMR7 sequence (Rienstra et al., 1998). Points *a* and *b* correspond to the frequency switching times on the ^{1}H (a) and ^{13}C (b) channel. The phase cycles are listed in Table 1.

Table 1. Phase cycles for the N(CO)CA and NCACB experiments. The phase numbering is indicated in Figure 1

ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	Receiver
+x	+x	+x	+y	−x	+x −x	+y −y	+y −x −y +x	2 3 0 1
−x	+x	+x	+x	−x	−x +x	+y −y	−x −y +x +y	3 0 1 2
	+y	+y	+y	−y	+y −y	−x +x	−y +x +y −x	0 1 2 3
	+y	+y	+y	−y	−y +y	−x +x	+x +y −x −y	1 2 3 0
	−x	−x	−x	+x	−x +x	+y −y	+y −x −y +x	0 1 2 3
	−x	−x	−x	+x	+x −x	+y −y	−x −y +x +y	1 2 3 0
	−y	−y	−y	+y	−y +y	−x +x	−y +x +y −x	2 3 0 1
	−y	−y	−y	+y	+y −y	−x +x	+x +y −x −y	3 0 1 2

Materials and methods

Sample preparation

The ubiquitin samples were obtained from VLI-research (Malvern, PA). The human ubiquitin gene was expressed in *E. coli*. Cells were grown in M9 minimal media, whose sole nitrogen source was $(^{15}NH_4)_2SO_4$ and the sole carbon source was either [2-^{13}C] glycerol (Cambridge Isotope Labs) or [U-$^{13}C_6$] glucose. Below we shall refer to these two samples as ^{13}C2-ubiquitin and U-^{13}C,^{15}N-ubiquitin, respectively. For a typical ubiquitin expression experiment, a 5 L culture was induced and the solid paste was purified by column chromatography to homogeneity. The purity of the final products was analyzed by gel electrophoresis and amino acid composition analysis to be >99%. The pools of purified protein solution were extensively dialysed against distilled water and lyophilized (http://www.vli-research.com/). The lyophilized protein was packed into a 4 mm MAS rotor and hydrated by direct addition of distilled and deionized H_2O. The water content was determined gravimetrically to be about 30% (w/w). It is well known that a hydrated powder yields sharper NMR lines than a lyophilized sample due to the reduction of structural heterogeneity (Poole and Finney, 1983; Kennedy and Bryant, 1990; Gregory et al., 1993). About 10 mg of U-^{13}C,^{15}N-ubiquitin was used for the N(CO)CA experiment, and about 30 mg of ^{13}C2-ubiquitin was used for the 3D NCACB experiment and the ^{13}C 2D experiments.

NMR spectroscopy

All NMR experiments were performed on a Bruker (Billerica, MA) DSX-300 spectrometer operating at 75.5 MHz for ^{13}C and 30.4 MHz for ^{15}N. A 4-mm ^{1}H/^{13}C/^{15}N triple-resonance MAS probehead was used. The spinning speeds ranged from 6000 Hz to 7000 Hz, regulated by a Bruker MAS control unit. All experiments were conducted at room temperature, 293 ± 2 K. Proton decoupling fields of 105–125 kHz were used. Typical carbon and nitrogen 90° pulse lengths were 4 μs and 6 μs, respectively. A linearly ramped rf field was applied on the ^{13}C channel during ^{1}H-^{13}C cross polarization to minimize dependence on the Hartmann–Hahn condition and maximize signal intensity (Metz et al., 1994). The CP contact time was typically 0.5 ms. During the acquisition period, TPPM decoupling (Bennett et al., 1995) was applied on the ^{1}H channel with a π pulse phase angle of ±10° to ± 14°.

To set up the triple-resonance experiments, the efficiencies of homonuclear and heteronuclear polarization transfer were separately optimized before the two parts are combined in the final sequence. The efficiency of the REDOR-based (Gullion and Schaefer, 1989; Hing et al., 1992; Hong and Griffin, 1998) ^{15}N-^{13}C transfer was typically 20–25%, measured as the intensity of the ^{13}C signal that passed the ^{15}N-^{13}C double- and zero-quantum filter relative to the intensity of the CP signal. The efficiency of the ^{13}C-^{13}C transfer under the CMR7 sequence was typically 40–45% after 1.1–1.2 ms of mixing. Higher efficiencies could not be attained, most likely due to the presence of 1.5–2 μs windows between the 2π pulses, which were required to allow time for phase switching. For the N(CO)CA experiment, the C-H dipolar dephasing period, designed to destroy the signals of protonated carbons, reduced the intensity of the carbonyl resonance by another 20%. Thus, the overall efficiency of the N(CO)CA experiment, judged by the intensity of the negative Cα signals relative to their positive CP signals, varied from −5% to −10%. The NCACB experiment had a slightly higher overall efficiency, since no C-H dephasing period was necessary.

The signal-averaging time of the 3D NCACB experiment was minimized by carefully choosing the ω_1 and ω_2 spectral widths and the quadrature detection scheme. Both ^{15}N and ^{13}Cα chemical shifts were measured off-resonance so that only cosine data sets needed to be recorded. The dwell time for the ^{15}N (ω_1) dimension was chosen to be one rotor period, while the ^{13}Cα (t_2) evolution period was incremented at a step of half a rotor period. At the spinning speed of 6000 kHz and the ^{15}N Larmor frequency of 30.4 MHz, the dwell time for ^{15}N evolution translates to an effective spectral width of 98.5 ppm. This is sufficient for covering most of the ^{15}N chemical shift range, except for the amine peaks, which were aliased into an empty region downfield from the amides. The recycle delay for the 3D experiment was optimized to 1.2 s. The relatively rapid ^{1}H T_1 relaxation is a common feature of proteins due to the presence of various modes of motions.

The proton-driven ^{13}C spin diffusion experiment was carried out using a 2D exchange sequence in which the ^{1}H decoupler was turned off during the mixing time (Szeverenyi et al., 1982). The cross peaks of one-bond and two-bond ^{13}C-^{13}C spin pairs were detected after a mixing time of 100 ms.

The ^{13}C dipolar INADEQUATE experiment was performed using CMR7 for homonuclear recoupling. The evolution period was incremented at a step of 1/7 of the CMR7 cycle time in order to generate a sufficiently large double-quantum spectral width (Hong, 1999b). At the spinning speed used (7 kHz), this corresponds to a ω_1 spectral width of 325 ppm.

Data processing

The spectra were processed using the XWINNMR software. A typical 2D spectrum was processed with 1024 and 256 complex points in the ω_2 and ω_1 dimensions, respectively. Gaussian broadening was employed in each dimension. For the 3D NCACB spectrum, 512, 128, and 128 points were used in the ω_3, ω_2 and ω_1 dimensions, respectively.

Results and discussion

2D N(CO)CA spectrum of U-^{13}C,^{15}N- ubiquitin

The N(CO)CA spectrum of uniformly ^{13}C, ^{15}N-labeled ubiquitin, shown in Figure 2a, demonstrates sequential assignment based on the N_i-Cα_{i-1} cross peaks, in the absence of the intraresidue N_i-Cα_i peaks. Effectively, it represents the difference spectrum between the one-bond ^{15}N-^{13}C spectrum, acquired at short mixing times, and the multiple-bond ^{15}N-^{13}C spectrum, acquired at long mixing times (Hong, 1999a). To illustrate this, we show the aliphatic region of the multiple-bond ^{15}N-^{13}C spectrum in Figure 2b for comparison. The assignment proceeds by identifying resonances in the N(CO)CA spectrum whose (ω_1, ω_2) frequencies uniquely equal the ω_1 frequency and

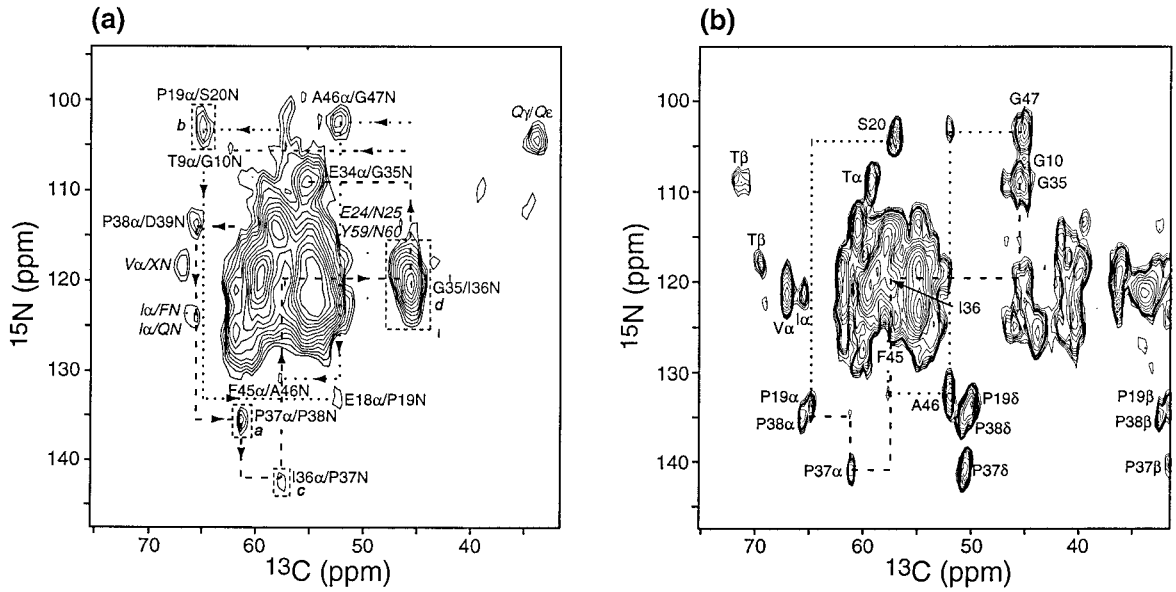

Figure 2. (a) 2D N(CO)CA spectrum of U-^{13}C,^{15}N-ubiquitin. (b) Multiple-bond ^{15}N-^{13}C correlation spectrum of ^{13}C2-ubiquitin for comparison (Hong, 1999a). Four groups of sequential connectivities involving 14 resonances are assigned in the N(CO)CA spectrum. Linetypes differentiate the various connectivity paths while arrows guide the eye for these paths. A few resonances are assigned to the amino acid type or to one of a few possible sequence-specific residues. The acquisition time for the N(CO)CA spectrum was 13 h (number of scans per t_1 point = 96; number of t_1 points = 78; recycle delay = 3 s). Spinning speed: 6.5 kHz. ^{1}H decoupling field during acquisition: 109 kHz.

ω_2 frequency of two separate peaks in the multiple-bond ^{15}N-^{13}C spectrum. For example, the resonance labeled *a* at (61.3 ppm, 136.5 ppm) in the N(CO)CA spectrum (Figure 2a) corresponds to the 'junction' of resonances (61.1 ppm, 142.1 ppm) and (65.8 ppm, 136.3 ppm) in the multiple-bond ^{15}N-^{13}C spectrum (Figure 2b). These two 'parent' peaks have been previously assigned to prolines due to their uniquely downfield ^{15}N chemical shifts. The presence of the cross peak in the N(CO)CA spectrum indicates that these two proline residues must be direct neighbors in the amino acid sequence. Thus the two parent proline peaks are unambiguously assigned to Pro37 and Pro38. Note that a weak cross peak between Pro37 and Pro38 is also observed in Figure 2b, due to direct through-space transfer from Pro38 ^{15}N to Pro37 ^{13}Cα. However, it is far less distinct than the corresponding peak in the N(CO)CA spectrum (Figure 2a). This demonstrates the efficiency of the two-step dipolar coherence transfer and the spectral simplification resulting from the removal of intraresidue resonances.

By superimposing the N(CO)CA spectrum with the multiple-bond ^{15}N-^{13}C correlation spectrum, we can identify several other sequential connectivities. For example, resonance *b* at (65.0 ppm, 103.7 ppm) in Figure 2a links a Pro resonance (65.0 ppm, 134.5 ppm) with a Ser resonance (57.0 ppm, 103.6 ppm) in Figure 2b. Since a unique Pro–Ser pair exists in the amino acid sequence of ubiquitin as Pro19–Ser20, this cross peak allows the assignment of the two intraresidue resonances in the multiple-bond ^{15}N-^{13}C spectrum. In total, we assigned four groups of connectivities, indicated by different linetypes in Figure 2: Glu34–Gly35–Ile36–Pro37–Pro38–Asp39, Phe45–Ala46–Gly47, Glu18–Pro19–Ser20, and Thr9–Gly10. By design of the experiment, the consecutive turning points of each connectivity path alternate between finite and zero intensities. Those turning points with zero intensities correspond to the positions of the intraresidue $^{15}N_i$-$^{13}C\alpha_i$ peaks that are manifested in the multiple-bond ^{15}N-^{13}C spectrum (Figure 2b) but have been removed by the N(CO)CA technique.

These resonances are assigned with slightly different degrees of confidence. For peaks that are not well resolved, the assignment particularly relies on the knowledge of the amino acid sequence of the protein. For example, the ^{13}Cα chemical shift of Ile36 (57.4 ppm) is unmistakably defined by its cross peak with Pro37, resonance *c* (Figure 2a). Its ^{15}N chemical shift, on the other hand, is not clearly known from the N(CO)CA spectrum due to resonance overlap.

Since Ile36 neighbors Gly35, which has a characteristic $^{13}C\alpha$ chemical shift well separated from all other amino acids, and since only one large resonance centered at (45.5 ppm, 120 ppm) (peak *d*) could result from glycine Cα, we expect that the intraresidue ^{15}N-$^{13}C\alpha$ resonance of Ile36 must appear at a ^{15}N frequency of about 120 ppm. This assignment is corroborated by the clear presence of the N-CO cross peak of Ile36 in the multiple-bond ^{15}N-^{13}C spectrum (area not shown) (Hong, 1999a).

In addition to sequence-specific assignment, several more peaks in the N(CO)CA spectrum can be either assigned to amino acid types or to a limited number of sequence-specific possibilities, based on the unique chemical shift ranges of these amino acids. These inconclusive assignments, shown in italics in Figure 2a, are nevertheless useful in combination with other correlation experiments.

The resonance assignments were made independent of the liquid-state NMR chemical shifts known for ubiquitin, utilizing only the characteristic chemical shift ranges of amino acids and peptides and the connectivity patterns in the spectra. After the assignments had been made in the solid state, they were compared with the liquid-state chemical shifts of ubiquitin (Wang et al., 1995). For ambiguous assignments, the various possibilities are indicated in order not to exaggerate the extent of independent assignment by solid-state NMR.

Straus and co-workers recently demonstrated a similar N(CO)CA correlation experiment with a sequence that differs from the current version in three main aspects (Straus et al., 1998). First, the previous sequence employed CP instead of REDOR to produce polarization transfer from ^{15}N to ^{13}C. Second, it used a Gaussian 90° pulse instead of a transverse C-H dephasing period to select the carbonyl signals over Cα. Third, the previous sequence relied on proton-driven spin diffusion to transfer the polarization from CO_{i-1} to $C\alpha_{i-1}$. We find that the coherence transfer schemes used in the current experiment yielded overall higher sensitivity than the previous sequence. This is judged from the signal-averaging times for the spectra, given similar sample amounts ($\sim$10 mg) and similar experimental conditions. Cross peaks such as Pro37α–Pro38N, Ile36α–Pro37N, Glu18α–Pro19N, and Pro19α–Ser20N are visible in the current spectrum with comparable signal-to-noise ratios as the previous spectrum. However, the current 2D N(CO)CA spectrum was acquired in 13 h, which is significantly shorter than the 5.3 days reported in the previous paper. This suggests that one or several of our sequence elements have higher coherence transfer efficiency. Specifically, the efficiency of double-REDOR ^{15}N-^{13}C mixing was typically 20–25%. Although this efficiency is far below unity and lower than several cross polarization schemes (Hediger et al., 1994; Baldus et al., 1996), the transfer was achieved in a relatively short time (about 1.2 ms), thus minimizing signal loss due to transverse relaxation. The CO-Cα polarization transfer was also more rapid and better defined by homonuclear dipolar recoupling (about 1 ms) than by proton-driven spin diffusion (about 100 ms).

More complete sequential assignment of the ^{13}C and ^{15}N resonances in the protein backbone requires improved resolution compared to that of the 2D N(CO)CA spectrum. Although the distinct chemical shifts of several special amino acids such as proline, serine, and glycine result in partial site resolution, the spectrum in the (110–130 ppm, 52–62 ppm) region still contains many unresolved resonances. This can be partially attributed to the severe line broadening caused by uniform ^{13}C labeling of the protein. Using the glycine signals as a benchmark, the linewidths of the CH_2 resonances of the uniformly labeled ubiquitin are about 2.2 ppm, while the linewidths in the ^{15}N-^{13}C spectra of selectively labeled ubiquitin are about 1.3 ppm.

3D NCACB spectrum of $^{13}C2$- ubiquitin

Enhanced site resolution can be achieved either by reducing the spectral linewidths or by increasing the dimensionality of the NMR spectrum. Narrower lines may be obtained by crystallizing the protein, which is possible for the well-behaved ubiquitin (Cook et al., 1979). However, we consider such a strategy not particularly meaningful, since it cannot be applied to noncrystalline and disordered proteins, which are the ultimate goal of the current technique development. Higher-resolution spectra may also be achieved by increasing the spinning speed (e.g., to 30 kHz) and using very high magnetic field strengths (e.g., 17.6 Tesla). However, since these conditions are not yet routine, we have pursued the second route, three-dimensional spectroscopy, to enhance the site resolution of solid protein spectra. Specifically, Cβ chemical shifts are utilized to disperse the ^{15}N-$^{13}C\alpha$ resonances in a third dimension (Grzesiek and Bax, 1992b).

The [2-^{13}C] glycerol labeling scheme generates $^{13}C\alpha$-$^{13}C\beta$ spin pairs for 9 out of the 20 amino acids, thus it is reasonably suited for demonstrating the 3D NCACB experiment. The 3D spectrum

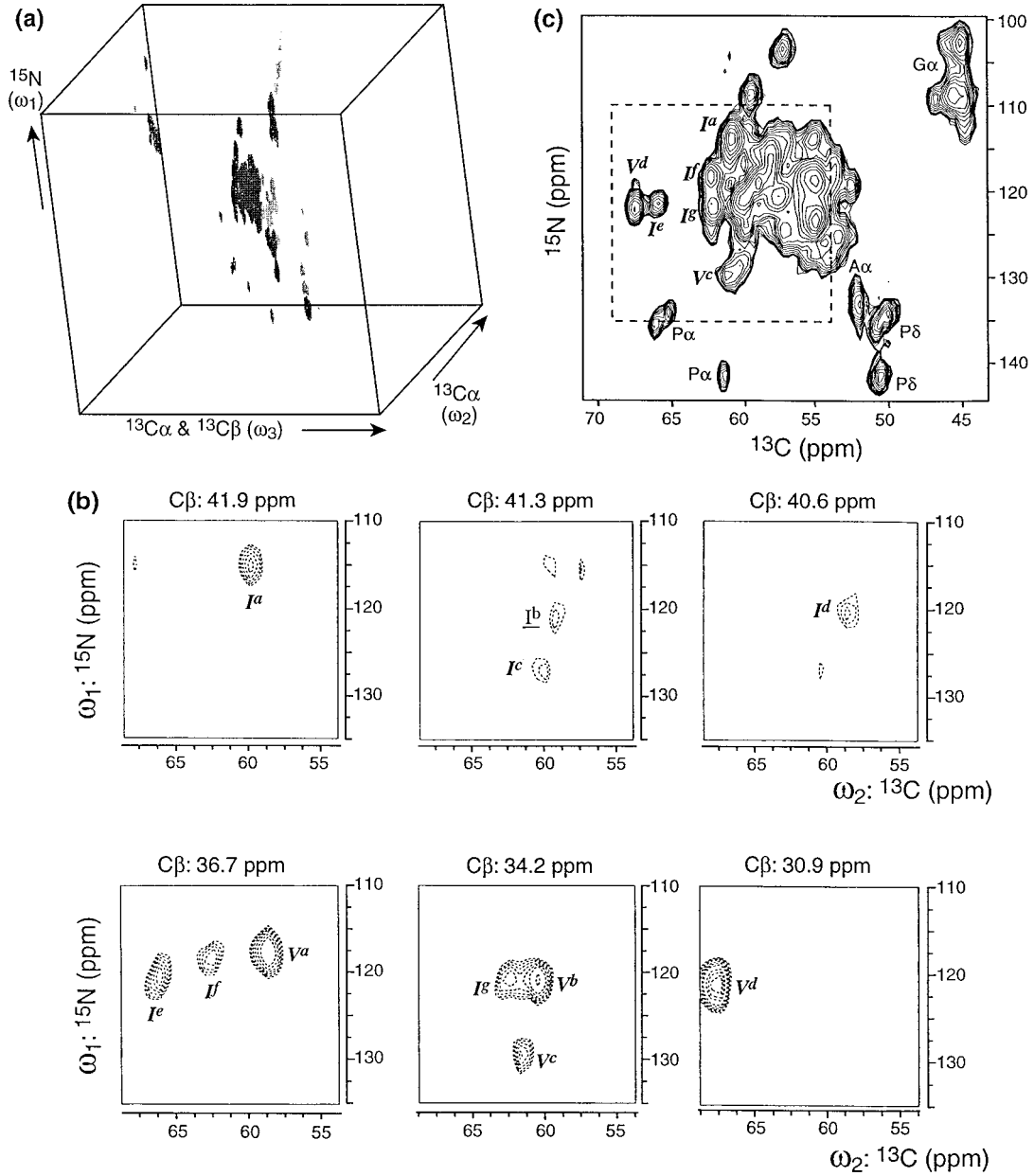

Figure 3. 3D NCACB spectrum of ^{13}C2-ubiquitin for amino acid type assignment. (a) 3D spectrum. Dark and light contours represent positive and negative intensities, respectively. (b) Selected $^{15}N_i$-$^{13}C\alpha_i$ (ω_1-ω_2) cross sections from various $^{13}C\beta$ frequencies. Dashed contours indicate negative intensities. Bold italics indicate assignment confirmed by other experiments. Underlines indicate tentative assignment. (c) 2D one-bond ^{15}N-^{13}C correlation spectrum for comparison (Hong, 1999a). The area enclosed by a dashed square corresponds to the spectral region shown in the cross sections. Note the higher resolution of the NCACB spectrum due to dispersion along the third dimension. The acquisition time for the 3D experiment was 52 h (number of scans per t_1 point = 32; number of t_1 points = 40; number of t_2 points = 116; recycle delay = 1.2 s). Spinning speed: 6 kHz. ^{1}H decoupling field during acquisition: 111 kHz.

of ^{13}C2-ubiquitin is shown in Figure 3a. The positive intensities (dark contours) correspond to the Cα resonances while the negative intensities (light contours) correspond to the Cβ peaks. Figure 3b displays six ^{15}N-^{13}Cα (ω_1-ω_2) cross sections extracted from various Cβ (ω_3) frequencies. The assignment begins with the identification of all four valine residues (labeled as V^a–V^d) in ubiquitin. These manifest as intense peaks with large round lineshapes in the region (^{13}Cα,^{13}Cβ) = (58–68 ppm, 30–37 ppm). The assignment is made on two grounds. First, valine is the only amino acid with 100% labeling for both Cα and Cβ in the [2-^{13}C] glycerol labeling scheme. Thus the ^{13}Cα-^{13}Cβ spin-pair probability of valine is the highest among all 20 amino acids, and the corresponding signal intensities should be the largest. Second, the average valine Cα and Cβ chemical shifts are 61 ppm and 31 ppm, respectively (Wüthrich, 1986). These are consistent with the frequencies of the four resonances when conformation-dependent secondary chemical shifts (Wishart et al., 1992) are taken into account. One of these four valine resonances, V^d, is corroborated by the time-dependent ^{15}N-^{13}C spectra as the resonance at (^{15}N, ^{13}C) = (122.0 ppm, 67.2 ppm) (Hong, 1999a). The other three valine peaks had not been assigned previously in the multiple-bond 2D ^{15}N-^{13}C spectrum (Figure 2b) due to insufficient spectral resolution. This demonstrates that 3D NCACB spectroscopy significantly surpasses 2D ^{15}N-^{13}C spectroscopy in the information content.

The amino acid with the next highest ^{13}Cα-^{13}Cβ spin pair probability is isoleucine: its Cβ is 100% labeled while its Cα is labeled at 50%, thus the Cα-Cβ pairs are doubly labeled at ∼50%. In the remaining resonances in the ^{15}N-^{13}Cα cross sections, the peaks labeled I^a (115.1 ppm, 59.8 ppm, 41.9 ppm), I^c (127.4 ppm, 60.0 ppm, 41.3 ppm), I^d (120.4 ppm, 58.5 ppm, 40.6 ppm), I^e (120.4 ppm, 66.3 ppm, 36.7 ppm), and I^f (118.5 ppm, 62.5 ppm, 36.7 ppm) can be assigned to isoleucine with confidence, since they show relatively strong intensities and fall into the characteristic chemical shift regions of isoleucine, which are centered around (123 ppm, 60 ppm, 37 ppm) (Wüthrich, 1986). In addition, I^a, I^e and I^f have clear correspondences in the one-bond ^{15}N-^{13}C 2D spectrum (Figure 3c) (Hong, 1999a), while I^c and I^d are confirmed by the ^{13}C spin diffusion spectrum (Figure 4) and the INADEQUATE spectrum (Figure 5) described below.

The combined use of signal intensities based on the known ^{13}C labeling levels and the characteristic chemical shift ranges reduces the ambiguity of the amino acid type assignment. For example, the Cβ chemical shifts of valine range from 29 ppm (α-helix) to 38 ppm (β-sheet), while the Cβ chemical shifts of isoleucine range from 35 ppm (α-helix) to 44 ppm (β-sheet) (Spera and Bax, 1991). The partial overlap would have made the assignment of the V^a (^{13}Cβ: 36.7 ppm) peak in Figure 3b uncertain. However, the distinctly stronger intensity of this resonance indicates that it must be valine. Additional chemical shift correlation also aids the assignment. For example, the resonance labeled I^g has an intensity and chemical shift values that could make it a possible valine if the NCACB spectrum alone is considered. However, in the ^{13}C spin diffusion spectrum (Figure 4), the corresponding Cα-Cβ peak at (60.8 ppm, 34.5 ppm) shares a cross peak with an upfield methyl signal at 9.5 ppm. Since valine Cγ is unlabeled and resonates around 19 ppm, while isoleucine Cδ1 is labeled and resonates around 11 ppm, this Cα-Cβ resonance belongs unequivocally to an isoleucine. When all spectra are evaluated, only one resonance in Figure 3b, I^b, cannot be conclusively assigned.

The 3D NCACB spectrum was acquired in a relatively short time: 32 scans were coadded for each slice, resulting in slightly over two days for the entire experiment. With this signal-averaging time, amino acids derived from the citric acid cycle, which have lower levels of doubly labeled Cα-Cβ pairs than valine and isoleucine, could not be observed clearly. According to the labeling pattern for [2-^{13}C] glycerol (Table 2 in (Hong, 1999a)), the Cα and Cβ sites are labeled at 25% and 50%, respectively, for Glu, Gln, Arg, and Pro, which derive from α-ketoglutarate. For Asp, Asn, Met, and Thr, whose common precursor is oxaloacetate, the Cα and Cβ sites are labeled at 50% and 25%, respectively. Therefore, the probability for doubly labeled Cα-Cβ pairs in these amino acids is about 12.5%, which is far below the 100% for valine and the 50% for isoleucine.

Although uniform ^{13}C labeling in principle allows more resonances to be observed in the 3D spectrum, it will also deteriorate the spectral resolution due to the ^{13}C-^{13}C J-coupling and dipolar couplings. The reduced resolution will not only complicate the analysis of the ^{15}N-^{13}Cα cross sections but will also make it difficult to separate the ^{13}Cβ resonances, thus hindering the extraction of the 2D ^{15}N-^{13}Cα cross sections in the first place.

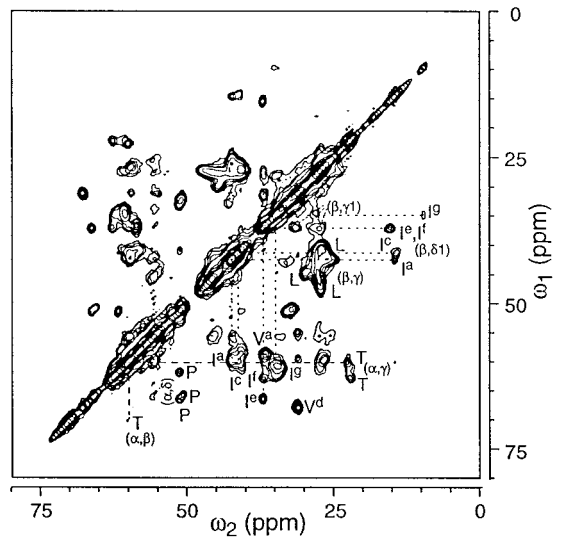

Figure 4. 2D proton-driven ^{13}C spin diffusion spectrum of ^{13}C2-ubiquitin, acquired with a mixing time of 100 ms. Greek letters in brackets specify the type of cross peaks. Superscripts for valine and isoleucine match the assignment in the NCACB spectrum. Dotted lines guide the eye for the connectivity paths. The acquisition time was 15 h (number of scans per t_1 point = 48; t_1 dwell time = 57.1 μs; number of t_1 points = 400; recycle delay = 2.8 s). Spinning speed: 7 kHz. ^{1}H decoupling field during acquisition: 116 kHz.

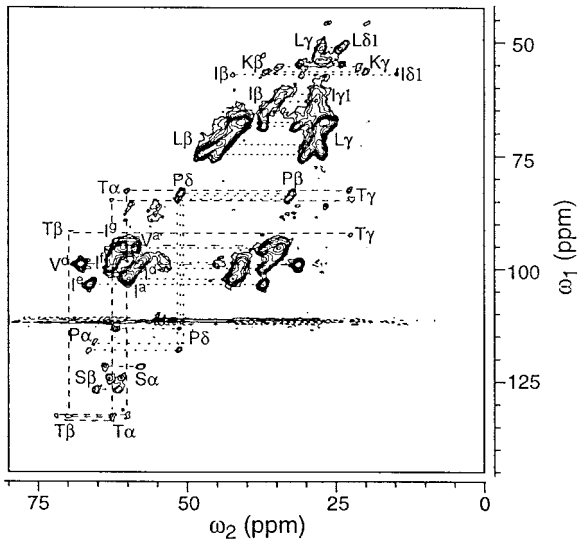

Figure 5. 2D ^{13}C dipolar INADEQUATE spectrum of ^{13}C2- ubiquitin, acquired with a double-quantum excitation time of 571 μs. Superscripts for valine and isoleucine match the assignment in the NCACB spectrum. Dotted lines guide the eye for the connectivity paths. The acquisition time was 22.5 h (number of scans per t_1 point = 64; t_1 dwell time = 40.8 μs; number of t_1 points = 800; recycle delay = 1.5 s). Spinning speed: 7 kHz. ^{1}H decoupling field during acquisition: 125 kHz.

2D ^{13}C spin diffusion spectrum of ^{13}C2-ubiquitin

Amino acid type assignment can also be made by direct correlation of the ^{13}C chemical shifts in a 2D ^{13}C homonuclear correlation spectrum. Such a spectrum may be obtained in a number of ways, depending on the method of polarization transfer and the coherence order involved. We first show the result of a ^{1}H-driven spin diffusion experiment, which involves only single-quantum ^{13}C coherence (Szeverenyi et al., 1982; McDermott et al., 1999). The mixing period for polarization transfer can be adjusted to probe connectivities between carbons separated by varying numbers of bonds.

In the spectrum of ^{13}C2-ubiquitin acquired with 100 ms of mixing time (Figure 4), five types of residues, Ile, Val, Leu, Thr, and Pro, are identified with relative ease. Isoleucines are labeled at the Cα, Cβ, Cγ1 and Cδ1 positions by the [2-^{13}C] glycerol labeling scheme. Among these sites, Cγ1 has the lowest labeling level and Cβ the highest, thus cross peaks between Cγ1 and Cδ1 are not observed while cross peaks between Cβ and Cδ1 can be easily identified. The distinct upfield shift of Cδ1 (∼11 ppm) allows the assignment of four peaks in the region (34–43 ppm, 10–16 ppm) to the Cβ-Cδ1 couplings of isoleucine. Once the Cβ chemical shifts are identified, Cβ-Cα cross peaks are found by tracing the connectivity paths. Five Cα-Cβ cross peaks at (60.8 ppm, 34.5 ppm), (62.8 ppm, 37.1 ppm), (66.3 ppm, 37.1 ppm), (59.5 ppm, 42.2 ppm), and (59.5 ppm, 42.2 ppm), agree with the I^g, I^f, I^e , I^c, and I^a resonances in the NCACB spectrum, respectively. The Cβ-Cδ1 peak of I^f at (36.7 ppm, 15.2 ppm) overlaps with that of I^e, which explains the particularly strong intensity of this resonance. Finally, three isoleucine Cβ-Cγ1 cross peaks are identified near the spectral diagonal in the region (34–37 ppm, 29–33 ppm).

Valine Cα-Cβ peaks can be identified based on their strong intensities due to 100% labeling and the distinctly downfield shift of the Cα site. Thus two resonances at (67.8 ppm, 31.0 ppm) and (58.5 ppm, 36.7 ppm) are assigned to valine V^d and V^a, respectively. Leucines have 100% labeled Cβ-Cγ pairs, whose unique chemical shifts around (40 ppm, 25 ppm) allow them to be assigned straightforwardly. There are nine leucine residues in ubiquitin, thus resonance overlap in this region is relatively severe, and only three major peaks can be resolved. These Cβ-Cγ peaks are missing in the NCACB spectrum due to the long distance between the amide ^{15}N and the Cβ.

Threonines are characterized by strong Cα-Cγ cross peaks (doubly labeled at ∼25%) around (60 ppm, 20 ppm) and weak Cβ-Cα cross peaks (doubly labeled at ∼12.5%) near the spectral diagonal around (70 ppm, 60 ppm). The intensity distribution of these peaks results from the low labeling level of Cβ (∼ 25%), which is also reflected by the lack of Cβ-Cγ cross peaks in the spectrum. Based on the strong Cα-Cγ resonances, two threonine resonances are identified, out of a total of seven threonine residues in ubiquitin. Finally, the prolines are labeled at the Cα (∼25%), Cβ (50%) and Cδ (100%) sites of the ring. The directly bonded and ∼25% labeled Cα-Cδ pairs give rise to three strong resonances in the region (60–66 ppm, 50–52 ppm). These assignments corroborate the result of the ^{15}N-^{13}C 2D spectra.

The proton-driven spin diffusion experiment contains two 90° pulses before and after the mixing period. Thus it essentially constitutes a stimulated-echo experiment, which refocuses the magnetization at $t_2 = t_1$. According to Fourier transform theory, this indicates that the widths of the diagonal peaks correspond to the intrinsic homogeneous linewidths of the spectrum (Schmidt-Rohr and Spiess, 1994). The presence of sharp diagonal peaks with truncation wiggles indicates that the homogeneous linewidth of the spectrum is smaller than the typical linewidths of the off-diagonal peaks. This suggests possibilities for further improvement of the spectral resolution. Based on the non-diagonal intensities, the linewidths of methyl (e.g. Ile Cδ1), methylene (e.g., Leu Cβ), and methine (e.g., Val Cα) resonances are approximately 0.8 ppm, 1.5 ppm, and 1.2 ppm, respectively.

The initial rate of spin diffusion depends on the square of the dipolar coupling between the two exchanging spins and on an integral that describes the overlap between two ^{13}C resonances in the presence of proton coupling (Linder et al., 1985). As a result, it is necessary to use hundreds of milliseconds or longer for polarization transfer over several angstroms. In this sense spin diffusion is an inefficient method for coherence transfer. Due to the complexity of the overlap integral, cross peak intensities in a spin diffusion spectrum do not correlate with internuclear distances quantitatively. Moreover, the spin diffusion spectrum suffers from a dominant diagonal, which obscures cross peaks close to it. To probe the ^{13}C-^{13}C bond connectivities more efficiently, with more emphasis on directly bonded carbons, and without neglecting spin pairs with small chemical shift differences, we employed a ^{13}C double-quantum technique.

2D ^{13}C dipolar-INADEQUATE spectrum of ^{13}C2-ubiquitin

The dipolar-mediated INADEQUATE experiment correlates the double-quantum signals of two coupled spins with the single-quantum chemical shift. The double-quantum coherence can be excited by the ^{13}C-^{13}C dipolar coupling, which is reintroduced by the CMR7 sequence in our experiment (Hong, 1999b). A double-quantum excitation time of 571 μs, corresponding to one cycle of the sequence under 7 kHz of sample spinning, was employed. Figure 5 displays the aliphatic region of the INADEQUATE spectrum of ^{13}C2-ubiquitin. Two groups of intense signals are observed. First, in the region $\omega_1 = 93$–105 ppm, the Cα-Cβ couplings of two valines (V^a and V^d) and five isoleucine residues (I^a, I^d, I^e, I^f, and I^g) are observed. These corroborate the assignment of the spin diffusion spectrum and the 3D NCACB spectrum. In particular, the pair of signals at (98.7 ppm, 58.0 ppm) and (98.7 ppm, 40.7 ppm) corresponds to the I^d Cα-Cβ resonance in the NCACB spectrum but was unresolved in the spin diffusion spectrum. This attests to the higher resolution of the INADEQUATE spectrum compared to the spin diffusion experiment. Second, in the region $\omega_1 = 60$–80 ppm, the strong and broad peaks can be assigned to the 100% labeled Cβ-Cγ pairs of leucine residues. The weaker and more closely spaced resonances are assigned to the Cβ-Cγ1 couplings of isoleucine based on the chemical shifts and the ∼25% labeling level of the ^{13}Cβ-^{13}Cγ1 pairs (Hong, 1999a). These isoleucine resonances were easy to miss in the spin diffusion spectrum due to their proximity to the strong diagonal. The double-quantum filtration of uncoupled spins by the INADEQUATE technique simplifies the detection of these coupled spins with similar chemical shifts.

In addition to these intense resonances, three clusters of weak cross peaks are observed in the spectrum. From $\omega_1 = 113$ ppm to 135 ppm, we identified proline Cα-Cδ, serine Cβ-Cα, and threonine Cβ-Cα resonances. The observation of the serine peaks is at first surprising, since only Cα should be labeled in this amino acid according to the [2-^{13}C] glycerol labeling protocol. However, since the Cα site is labeled at 100%, even a low scrambling level of 15% at Cβ would have increased the ^{13}Cβ-^{13}Cα spin-pair probability to 15%. This would come close to the spin-pair probability of many amino acids produced from the citric acid cycle, whose cross peaks are detectable in the spectrum. For example, the threonine Cβ-Cα sites, which resonate slightly downfield from the ser-

ine peaks in the ω_1 dimension, are doubly labeled at about 12.5%. Therefore, the manifestation of the serine cross peaks indicates label scrambling at Cβ, which had not been detected in previous heteronuclear and homonuclear spectra. It demonstrates the capability of double-quantum filtration in revealing low percentages of spin pairs.

From ω_1 = 80 ppm to 93 ppm, proline Cδ-Cβ, threonine Cα-Cγ and Cβ-Cγ cross peaks are identified. The presence of a large splitting between (92.5 ppm, 70.0 ppm) and (92.5 ppm, 22.5 ppm), which is assigned to a threonine Cβ-Cγ pair, attests to the broadband nature of the technique. Again, a strong correlation is found between the signal intensities and the labeling levels. For example, the threonine Cβ-Cγ sites (∼12.5% labeling) exhibit much weaker signals than the Cα-Cγ pairs (∼25% labeling), although they are better recoupled due to the shorter internuclear distance.

In the most upfield region of the ω_1 dimension (45–57 ppm), the cluster of resonances is tentatively assigned to isoleucine Cβ-Cδ1, lysine Cβ-Cγ, and leucine Cγ-Cδ1 based on the chemical shifts. The assignment of the lysine and leucine peaks assumes scrambling at the Cβ site and the Cδ1 site, respectively.

Assignment from complementary [1-^{13}C] glucose and [2-^{13}C] glycerol labeling

Due to the partial complementarity of the [1-^{13}C] glucose and [2-^{13}C] glycerol labeling schemes, additional resonance assignment can be extracted by comparing the spectra of the differently labeled samples. For example, in the ^{15}N-^{13}C spectrum of ^{13}C2-ubiquitin, two adjacent resonances at (121.7 ppm, 65.7 ppm) and (122.0 ppm, 67.2 ppm) were originally assigned to valine Cα and isoleucine Cα without distinguishing between the two (Hong, 1999a). In the ^{15}N-^{13}C spectrum of ^{13}C1-ubiquitin (not shown), however, only one peak is observed at (122.0 ppm, 65.5 ppm). Since valine Cα is unlabeled in the [1-^{13}C] glucose scheme, while isoleucine Cα is labeled, the peak in the ^{13}C1-ubiquitin spectrum is definitively assigned to isoleucine N-Cα. Thus the remaining peak in the ^{13}C2-ubiquitin spectrum is attributed to valine.

Summary of ^{13}C and ^{15}N assignment of ubiquitin by solid-state NMR

We have developed and demonstrated several 2D and 3D MAS correlation techniques to test the extent of ^{13}C and ^{15}N assignment of proteins in the solid state. These include three heteronuclear experiments: N(CO)CA, NCACB, and the time-dependent ^{15}N-^{13}C correlation described previously (Hong, 1999a), and two homonuclear techniques: proton-driven ^{13}C spin diffusion and ^{13}C dipolar INADEQUATE (Hong, 1999b). Combining all the spectra, we are able to assign the backbone ^{15}N and ^{13}C resonances and some of the side-chain carbons of 19 out of 76 residues in ubiquitin in a sequence-specific fashion. These conclusively assigned residues are underlined in the amino acid sequence of ubiquitin in Figure 6. In addition, a larger number of resonances has been assigned to the amino acid types, as compiled in Table 2. These assignments were made in the solid state based on the spectral connectivities, the characteristic chemical shifts, and the selective and extensive ^{13}C labeling patterns. Since they were made independent of the solution-state chemical shifts, they represent the true extent of assignment currently feasible in our hands for a protein of this size.

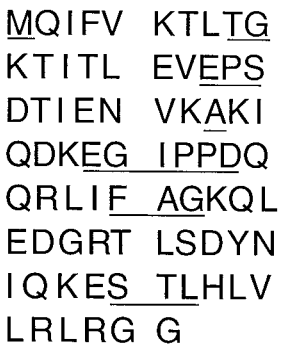

Figure 6. Amino acid sequence of ubiquitin. Underlined residues have been assigned by the techniques described here and in two recent publications (Hong, 1999a, b).

The amino acid type assignment involved various side-chain ^{13}C chemical shifts, including Cβ, Cγ, and Cδ. It can be seen that the most completely and easily assigned residues are either those with distinct chemical shifts, or those that are labeled at high levels and with high selectivity. Thus, all proline and valine resonances were type-assigned. Interestingly, although all three prolines in ubiquitin could be identified sequence-specifically, no valine resonances could be assigned sequentially, due to the lack of unique chemical shifts or unique labeling pattern of their neighboring residues. In addition to these two amino acids, several other amino acids lend themselves conveniently to type-assignment by virtue of their distinct chemical shifts: glycine (Cα), alanine (Cβ), leucine (Cβ and Cγ), isoleucine (Cδ1), serine (Cβ), threonine (Cβ), lysine (Cϵ), and arginine (Cδ) (Tycko, 1996).

Table 2. Summary of type-assigned amino acid residues in ubiquitin, using ^{15}N-^{13}C and ^{13}C-^{13}C MAS multidimensional correlation spectroscopy

Amino acid	Total no.	No. assigned	Amino acid	Total no.	No. assigned
Gly	6	3[a,c]	Gln	6	0
Ser	3	3[a−c,f]	Pro	3	3[a−c,e−g]
Ala	2	2[a−c]	Arg	4	1[a,b,g]
Leu	9	3[b,e−g]	Asp	5	0
Val	4	4[a−e,f]	Asn	2	1[b,g]
His	1	1[h]	Met	1	1[a,b]
Phe	2	2[a,c,h]	Thr	7	3[a−c,e−g]
Tyr	1	1[e,h]	Lys	7	1[a,b,f,g]
Glu	6	1[b,c]	Ile	7	6[a−g]

[a] 2D ^{15}N-^{13}C experiment on ^{13}C2-ubiquitin.
[b] 2D ^{15}N-^{13}C experiment on ^{13}C1-ubiquitin.
[c] 2D N(CO)CA experiment on U-^{13}C,^{15}N-ubiquitin.
[d] 3D NCACB experiment on ^{13}C2-ubiquitin.
[e] 2D ^{13}C spin diffusion experiment on ^{13}C2-ubiquitin.
[f] 2D ^{13}C dipolar INADEQUATE experiment on ^{13}C2-ubiquitin.
[g] 2D ^{13}C dipolar INADEQUATE experiment on ^{13}C1-ubiquitin.
[h] 1D ^{13}C CP.

In comparison, the amino acids produced from the citric acid cycle were more difficult to assign because of their similar labeling patterns, low labeling levels, and chemical shift overlap. Further experimental schemes, yielding larger ratios of chemical shift dispersion over linewidths, are necessary for tackling the assignment of these difficult amino acids.

Conclusions and outlook

The ability to assign the NMR resonances to the protein sequence provides the basis for efficient structure elucidation of extensively labeled proteins. We have shown here that multidimensional MAS correlation spectroscopy in conjunction with selective and extensive ^{13}C labeling makes it possible to assign a significant fraction of the ^{13}C and ^{15}N resonances of a model protein in the solid state. The N(CO)CA technique yields sequential backbone connectivities, while 3D NCACB correlation substantially enhances the site resolution and permits amino acid type assignment. The proton-driven ^{13}C spin diffusion experiment extends the type assignment by utilizing not only Cα and Cβ chemical shifts but also the chemical shifts of carbons further down the side chain. Finally, the ^{13}C dipolar INADEQUATE experiment supplements the type-assignment by removing the dominant spectral diagonal, thus allowing the detection of low levels of spin pairs such as the Cα and Cβ of serine and the Cβ and Cγ of threonine. Combining all four experiments presented here, and the previous ^{15}N-^{13}C 2D correlation technique, we have achieved sequence-specific assignment (mainly in the backbone) for 25% of the amino acids in ubiquitin and type-assignment for 47% of the residues.

All correlation techniques demonstrated here utilize dipolar-mediated coherence transfer schemes, which allow more rapid transfer than scalar-coupling-based pulse sequences without undue loss of transfer specificity at short mixing times. In fact, one can take advantage of the mixing-time dependence of dipolar transfer to probe one- to multiple-bond connectivities. This has been shown to be useful in 2D ^{15}N-^{13}C correlation spectroscopy.

The assignment techniques shown here are ultimately limited by two factors: the spectral linewidths and the incomplete labeling of the selective and extensive ^{13}C labeling protocol. Given the significant line narrowing already achieved by the selective and extensive ^{13}C labeling approach, the resolution may be further enhanced by higher power ^{1}H decoupling, better shift dispersion under higher magnetic field strengths, more sophisticated sample preparation methods, and faster spinning. For the few directly bonded and highly labeled ^{13}C spin pairs, such as Val Cα and Cβ, techniques that remove ^{13}C-^{13}C one-bond J-couplings will be useful for reducing the linewidths, although such experiments may be costly in sensitivity (Straus et al., 1996). Finally, more labeled samples with variations

of the selective and extensive ^{13}C labeling approach will also be valuable for obtaining additional assignment.

Even with the incomplete resonance assignment presented here, much structural information should now become accessible. For example, multiple backbone and side-chain torsion angles can be measured from a single experiment to provide information on the secondary structure as well as the local tertiary structure of proteins (Hong, 1999a). Further work along these lines is currently pursued in our laboratory.

Acknowledgements

This work is supported by the National Science Foundation through a POWRE grant (MCB-9870373) and a MRSEC grant at UMass.

References

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

Baldus, M. and Meier, B.H. (1996) *J. Magn. Reson.*, **A121**, 65–69.

Baldus, M., Tomaselli, M., Meier, B.H. and Ernst, R.R. (1994) *Chem. Phys. Lett.*, **230**, 329–336.

Bax, A. (1994) *Curr. Opin. Struct. Biol.*, **4**, 738–744.

Bennett, A.E., Ok, J.H., Griffin, R.G. and Vega, S. (1992) *J. Chem. Phys.*, **96**, 8624–8627.

Bennett, A.E., Griffin, R.G. and Vega, S. (1994) In *Recoupling of Homo- and Heteronuclear Dipolar Interactions in Rotating Solids*, Vol. 33 (Eds., P. Diehl, E. Fluck and E. Kosfeld), Springer, Berlin, pp. 1–77.

Bennett, A.E., Rienstra, C.M., 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. and Gronenborn, A.M. (1991) *Annu. Rev. Biophys. Chem.*, **20**, 29–63.

Cook, W.J., Suddath, F.L., Bugg, C.E. and Goldstein, G. (1979) *J. Mol. Biol.*, **130**, 353–355.

Creuzet, F., McDermott, A., Gebhard, R., van der Hoef, K., Spijker-Assink, M.B., Herzfeld, J., Lugtenburg, J., Levitt, M.H. and Griffin, R.G. (1991) *Science*, **251**, 783–786.

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

Feng, X., Verdegem, P.J.E., Lee, Y.K., Sandstrom, D., Eden, M., Bovee-Geurts, P., de Grip, W.J., Lugtenburg, J., de Groot, H.J.M. and Levitt, M.H. (1997) *J. Am. Chem. Soc.*, **119**, 6853–6857.

Gregory, D., Mitchell, D.J., Stringer, J.A., Kiihne, S., Shiels, J.C., Callahan, J., Mehta, M.A. and Drobny, G.P. (1995) *Chem. Phys. Lett.*, **246**, 654–663.

Gregory, R.B., Gangoda, M., Gilpin, R.K. and Su, W. (1993) *Biopolymers*, **33**, 513–519.

Griffin, R.G., Costa, P.R., Gross, J.D., Hatcher, M., Hong, M., Hu, J., Mueller, L., Rienstra, C.M., Fesik, S. and Herzfeld, J. (1998) *Experimental NMR Conference*, Asilomar, CA.

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

Grzesiek, S. and Bax, A. (1992b) *J. Magn. Reson.*, **96**, 432–440.

Gullion, T. and Schaefer, J. (1989) *J. Magn. Reson.*, **81**, 196–200.

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

Hing, A.W., Vega, S. and Schaefer, J. (1992) *J. Magn. Reson.*, **96**, 205–209.

Hohwy, M., Jakobsen, H.J., Eden, M., Levitt, M.H. and Nielsen, N.C. (1998) *J. Chem. Phys.*, **108**, 2686–2694.

Hohwy, M., Rienstra, C.M., Jaroniec, C.P. and Griffin, R.G. (1999) *J. Chem. Phys.*, **110**, 7983–7992.

Hong, M. (1999a) *J. Magn. Reson.*, **139**, 389–401.

Hong, M. (1999b) *J. Magn. Reson.*, **136**, 86–91.

Hong, M. and Griffin, R.G. (1998) *J. Am. Chem. Soc.*, **120**, 7113–7114.

Hong, M. and Jakes, K. (1999) *J. Biomol. NMR*, **14**, 71–74.

Kennedy, S.D. and Bryant, R.G. (1990) *Biopolymers*, **29**, 1801–1806.

Ketchem, R.R., Hu, W. and Cross, T.A. (1993) *Science*, **261**, 1457–1460.

Lesage, A., Auger, C., Caldarelli, S. and Emsley, L. (1997) *J. Am. Chem. Soc.*, **119**, 7867–7868.

Linder, M., Henrichs, M., Hewitt, J.M. and Massa, D.J. (1985) *J. Chem. Phys.*, **82**, 1585–1598.

Long, J.R., Dindot, J.L., Zebroski, H., Kiihne, S., Clark, R.H., Campbell, A.A., Stayton, P.S. and Drobny, G.P. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 12083–12087.

McDermott, A.E., Polenova, T., Montelione, G., Zilm, K., Paulsen, E., Martin, R. and Coker, G. (1999) *40th Experimental NMR Conference*, Orlando, FL.

McDowell, L.M. and Schaefer, J. (1996) *Curr. Opin. Struct. Biol.*, **6**, 624–629.

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

Nielsen, N.C., Bildsoe, H., Jakobsen, H.J. and Levitt, M.H. (1994) *J. Chem. Phys.*, **101**, 1805–1812.

Poole, P.L. and Finney, J.L. (1983) *Int. J. Biol. Macromol.*, **5**, 308–310.

Rienstra, C.M., Hatcher, M.E., Mueller, L.J., Sun, B.Q., Fesik, S.W. and Griffin, R.G. (1998) *J. Am. Chem. Soc.*, **120**, 10602–10612.

Schmidt-Rohr, K. and Spiess, H.W. (1994) *Multidimensional Solid-state NMR and Polymers*, Academic Press, San Diego, CA.

Shon, K.-J., Kim, Y., Colnago, L.A. and Opella, S.J. (1991) *Science*, **252**, 1303–1305.

Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.

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. (1998) *J. Biomol. NMR*, **12**, 39–50.

Sun, B.-Q., Costa, P.R., Kocisko, D., Lansbury, P.T.J. and Griffin, R.G. (1995) *J. Chem. Phys.*, **102**, 702–707.

Szeverenyi, N.M., Sullivan, M.J. and Maciel, G.E. (1982) *J. Magn. Reson.*, **47**, 462–475.

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

Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.

Wang, A.C., Grzesiek, S., Tschudin, R., Lodi, P.J. and Bax, A. (1995) *J. Biomol. NMR*, **5**, 376–382.

Weliky, D. and Tycko, R. (1996) *J. Am. Chem. Soc.*, **118**, 8487–8488.

Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) *Biochemistry*, **31**, 1647–1651.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.