Assignments of Carbon NMR Resonances for Microcrystalline Ubiquitin

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · JULY 2004

Impact Factor: 12.11 · DOI: 10.1021/ja030547o · Source: PubMed

6 AUTHORS, INCLUDING:

Ann Mcdermott
Columbia University
100 PUBLICATIONS **4,001** CITATIONS
SEE PROFILE

Rachel W Martin
University of California, Irvine
36 PUBLICATIONS **858** CITATIONS
SEE PROFILE

Eric K Paulson
Yale University
11 PUBLICATIONS **461** CITATIONS
SEE PROFILE

Andrew Joshua Wand
University of Pennsylvania
193 PUBLICATIONS **8,640** CITATIONS
SEE PROFILE

J|A|C|S
ARTICLES
Published on Web 05/11/2004

Assignments of Carbon NMR Resonances for Microcrystalline Ubiquitin

Tatyana I. Igumenova,[†,#] Ann E. McDermott,*[,†] Kurt W. Zilm,*[,‡] Rachel W. Martin,[‡] Eric K. Paulson,[‡] and A. Joshua Wand[§]

Contribution from the Department of Chemistry, Columbia University, 3000 Broadway MC 3113, New York, New York 10027, Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06511, and Department of Biochemistry and Biophysics, University of Pennsylvania, The Johnson Research Foundation, Philadelphia, Pennsylvania 19104

Received September 22, 2003; Revised Manuscript Received February 4, 2004; E-mail: aem5@columbia.edu; kurt.zilm@yale.edu

Abstract: Solid-state NMR 2D spectroscopy was used to correlate carbon backbone and side-chain chemical shifts for uniformly ^{13}C,^{15}N-enriched microcrystalline ubiquitin. High applied field strengths, 800 MHz for protons, moderate proton decoupling fields, 80−100 kHz, and high magic angle sample spinning frequencies, 20 kHz, were used to narrow the most of the carbon line widths to 0.5−0.8 ppm. Homonuclear magnetization transfer was effected by matching the proton RF field to the spinning frequency, the so-called dipolar-assisted rotational resonance (DARR) (Takegoshi, K.; Nakamura, S.; Terao, T. *Chem. Phys. Lett.* **2001**, *344*, 631−637), and a mixing time of 20 ms was used to maximize the intensity of one-bond transfers between carbon atoms. This polarization transfer sequence resulted in roughly 14% transfer efficiencies for directly bonded carbon pairs and 4% transfer efficiencies for carbons separated by a third carbon. With this simple procedure, the majority of the one-bond correlations was observed with moderate transfer efficiencies, and many two-bond correlations were also observed with weaker intensities. Spin systems could be identified for more than half of the amino acid side chains, and site-specific assignments were readily possible via comparison with 400 MHz $^{15}N-^{13}C-^{13}C$ correlation spectroscopy (described separately).

Introduction

Assignment of solid-state NMR spectra, i.e., correlation of the isotropic backbone and side-chain ^{15}N and ^{13}C resonances using multidimensional high-resolution methods, has had a surge of activity in the recent past.[2−7] Magic angle spinning (MAS)-based methods utilize isotropic shifts in the context of well-resolved dimensions, which allow for ready identification of side-chain types and thereby offer independent confirmation of assignments. On the other hand, strong anisotropic interactions can be utilized during MAS measurements for recoupling and for identifying angular relations among structural units. These interactions provide a basis for structural, dynamical, and functional characterization of proteins. A key issue for these endeavors is the ultimate resolution of the peaks: clean, nearly complete, and reproducible 2D "fingerprint" spectra for backbone and side chains will be very useful, and line widths below 1 ppm are therefore desirable; furthermore, sample stability of the order of days is also needed. As a test system for optimizing line widths and assignment protocols, we have selected human ubiquitin. At 8.6 kDa or 76 residues, ubiquitin is a relatively small globular protein with mixed secondary structure and excellent thermal stability. Extensive information about the structure and dynamics of ubiquitin is available from X-ray crystallography[8] and solution NMR studies.[9−14] We have used this protein to illustrate that high quality spectra at high magnetic fields are readily attained and to explore the principal issues that limit resolution for solid-state NMR of uniformly enriched materials.

To compile a ^{13}C chemical shift list for ubiquitin in the solid state, it was necessary to use two-dimensional (2D) NMR studies. A variety of solid-state NMR pulse sequences have been

[†] Columbia University.
[‡] Yale University.
[§] University of Pennsylvania.
[#] Present address: Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104.

(1) Takegoshi, K.; Nakamura, S.; Terao, T. *Chem. Phys. Lett.* **2001**, *344*, 631−637.
(2) Straus, S. K.; Bremi, T.; Ernst, R. R. *J. Biomol. NMR* **1998**, *12*, 39−50.
(3) Hong, M. *J. Biomol. NMR* **1999**, *15*, 1−14.
(4) McDermott, A.; Polenova, T.; Bockmann, A.; Zilm, K. W.; Paulsen, E. K.; Martin, R. W.; Montelione, G. T. *J. Biomol. NMR* **2000**, *16*, 209−219.
(5) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. *ChemBioChem* **2001**, *2*, 272−281.
(6) van Rossum, B. J.; Castellani, F.; Rehbein, K.; Pauli, J.; Oschkinat, H. *ChemBioChem* **2001**, *2*, 906−914.
(7) Detken, A.; Hardy, E. H.; Ernst, M.; Kainosho, M.; Kawakami, T.; Aimoto, S.; Meier, B. H. *J. Biomol. NMR* **2001**, *20*, 203−221.
(8) Vijaykumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.
(9) DiStefano, D. L.; Wand, A. J. *Biochemistry* **1987**, *26*, 7272−7281.
(10) Weber, P. L.; Brown, S. C.; Mueller, L. *Biochemistry* **1987**, *26*, 7282−7290.
(11) Wand, A. J.; Urbauer, J. L.; McEvoy, R. P.; Bieber, R. J. *Biochemistry* **1996**, *35*, 6116−6125.
(12) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.
(13) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.
(14) Babu, C. R.; Flynn, P. F.; Wand, A. J. *J. Am. Chem. Soc.* **2001**, *123*, 2691−2692.

10.1021/ja030547o CCC: $27.50

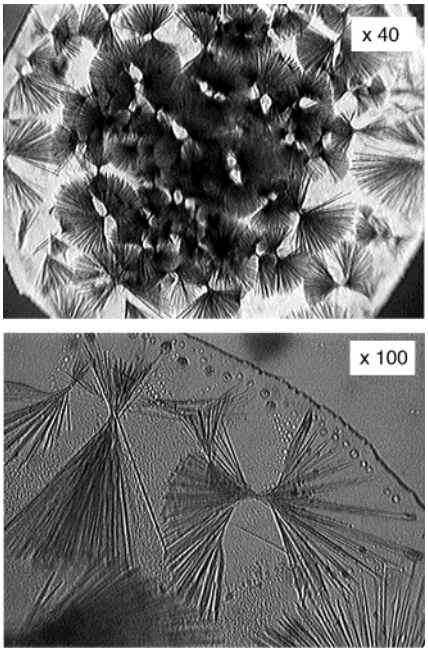

Figure 1. Photographs of the microcrystals of the type used in the NMR experiments. The microcrystals are situated in a hanging drop of diameter 2.3 mm. The crystal dimensions are approximately $10 \times 250\ \mu$m; the small width and the fact that they are twinned make them unsuitable for X-ray diffraction measurements. Crystallization protocol is described in the text.

used to record homonuclear chemical shift correlations in uniformly enriched materials; these range from simple methods such as passive spin diffusion[15] or RFDR,[16,17] to sophisticated pulse sequences such as adiabatic[18,19] or elegant rotor- synchronized methods.[20,21] We elected to use a recently developed scheme involving very low radio frequency (RF) power requirements. In the so-called dipolar-assisted rotational resonance (DARR) diffusion sequence, the proton RF power is adjusted to be an integral multiple of the spinning frequency.[1] At these conditions, the proton homonuclear and ^{1}H$-^{13}$C heteronuclear dipolar couplings, otherwise suppressed through irradiation, are largely restored, having the effect of broadening the carbon lines and assisting carbon magnetization transfer.[1] Importantly, this sequence has very low power requirements and is reasonably experimentally stable, making it a very practical choice for high field protein studies.

Materials and Methods

Uniformly ^{13}C,^{15}N-enriched ubiquitin was prepared by overexpression in *Escherichia coli*; the refolding and purification procedures are described elsewhere.[11] Roughly 7 mg of uniformly ^{13}C ^{15}N-enriched ubiquitin was crystallized by batch methods in 60% 2-methyl-2,4-pentanediol (MPD), 20 mM citrate buffer, pH 4.0−4.2. The crystals formed under these conditions are needles, shown in Figure 1, and appear to correspond to a new space group, relative to those observed in the previous X-ray structural characterizations.[8] These crystals could be prepared in high yield using batch crystallization protocol. In this protocol, the precipitant was slowly added to the protein solution in buffer, and the mixture was left to crystallize for 12 h at 4 °C. The microcrystals formed in batch and hanging drop were found to have identical morphology. Alternative crystallization protocols exist for this sample including use of polyethyleneglycol (PEG) precipitants, which have been discussed in a separate study of protein crystallization protocols for ssNMR spectroscopy.[45] It is noteworthy in this context that at low pH values and high concentrations of alcohols, ubiquitin can adopt a stable partially folded state, the "A form",[22] wherein the N-terminal portion remains nativelike, but the C-terminal part acquires an α-helical character.[23] A mixture of A form and native ubiquitin was observed in the batch crystallization solution, based on the chemical shift of $\epsilon 1/\epsilon 2$ protons of Tyr59.[24,25] The ^{13}C and ^{15}N solid-state NMR spectra of ubiquitin suggest that this solution equilibrium is shifted toward the native state in the process of crystallization. Chemical shifts for some of the upfield methyl and aromatic groups did vary noticeably in comparing the PEG and the MPD crystallized materials (Figure 2B,C), and thus the table of chemical shifts may be considered as a precise description for ubiquitin in this crystalline habit only. It should be noted that in work previously reported in the literature, or unpublished work from our group, lyophilized or rehydrated lyophilized materials can be expected to give line widths of the order 1.5−2.5 ppm, dominated by inhomogeneous dispersion of unknown molecular origin. For this reason, care was taken to maintain the hydration of the sample throughout the MAS experiments. The ^{15}N solid-state NMR spectra were very sensitive to the sample intactness and exhibited significant sharpening with increasing applied magnetic field (Figure 3). It is well-known that sample preparation protocols have a profound effect on the line widths in solid-state NMR spectra of proteins and peptides. NMR lines of freeze-dried proteins and peptides are usually inhomogeneously broadened. Hydration of freeze-dried samples was shown to improve the spectral resolution in ^{13}C cross-polarization magic angle spinning (CP MAS) spectra of natural abundance lysozyme[26] and ubiquitin.[27] In the latter case, $^{13}CH_2$ line widths of >2 ppm and ∼1.3 ppm were reported for uniformly and "selectively and extensively" ^{13}C-labeled samples, respectively. Addition of cryoprotectants such as trehalose and poly(ethylene glycol) can improve line widths for lyophilized proteins.[28] In contrast, solid protein samples with some degree of microscopic order were experimentally shown to give narrow NMR lines for the cases of staphylococcal nuclease,[29] α-lytic protease,[30] BPTI,[4] SH3 domain of α-spectrin,[31] ubiquitin,[32] and triose phosphate isomerase.[33] A viable alternative for noncrystallizable samples is well-controlled precipitation, which resulted in <1 ppm carbon line widths in a variety of intrinsic membrane proteins (A. McDermott, unpublished data).

1D ^{13}C and ^{15}N CP MAS solid-state NMR spectra at 9.4 T were acquired on a Varian/Chemagnetics Infinity Plus 400 spectrometer operating at Larmor frequencies of 396.8 MHz for protons, 99.8 MHz

(15) Szeverenyi, N. M.; Sullivan, M. J.; Maciel, G. E. *J. Magn. Reson.* **1982**, *47*, 462−475.
(16) Bennett, A. E.; Ok, J. H.; Griffin, R. G.; Vega, S. *J. Chem. Phys.* **1992**, *96*, 8624−8627.
(17) Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W. G.; Lansbury, P. T.; Griffin, R. G. *J. Chem. Phys.* **1998**, *108*, 9463−9479.
(18) Verel, R.; Baldus, M.; Ernst, M.; Meier, B. H. *Chem. Phys. Lett.* **1998**, *287*, 421−428.
(19) Verel, R.; Ernst, M.; Meier, B. H. *J. Magn. Reson.* **2001**, *150*, 81−99.
(20) Hohwy, M.; Jakobsen, H. J.; Eden, M.; Levitt, M. H.; Nielsen, N. C. *J. Chem. Phys.* **1998**, *108*, 2686−2694.
(21) Hohwy, M.; Rienstra, C. M.; Jaroniec, C. P.; Griffin, R. G. *J. Chem. Phys.* **1999**, *110*, 7983−7992.
(22) Wilkinson, K. D.; Mayer, A. N. *Arch. Biochem. Biophys.* **1986**, *250*, 390−399.
(23) Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *Biochemistry* **1997**, *36*, 13043−13053.
(24) Harding, M.; Williams, D. H.; Woolfson, D. N. *Biochemistry* **1991**, *30*, 3120−3128.
(25) Pan, Y. Q.; Briggs, M. S. *Biochemistry* **1992**, *31*, 11405−11412.
(26) Gregory, R. B.; Gangoda, M.; Gilpin, R. K.; Su, W. *Biopolymers* **1993**, *33*, 513−519.
(27) Hong, M. *J. Magn. Reson.* **1999**, *139*, 389−401.
(28) Jakeman, D. L.; Mitchell, D. J.; Shuttleworth, W. A.; Evans, J. N. S. *J. Biomol. NMR* **1998**, *12*, 417−421.
(29) Cole, H. B. R.; Torchia, D. A. *Chem. Phys.* **1991**, *158*, 271−281.
(30) Smith, S. O.; Farrjones, S.; Griffin, R. G.; Bachovchin, W. W. *Science* **1989**, *244*, 961−964.
(31) Pauli, J.; van Rossum, B.; Forster, H.; de Groot, H. J. M.; Oschkinat, H. *J. Magn. Reson.* **2000**, *143*, 411−416.
(32) Igumenova, T. I.; Martin, R.; Rienstra, C. M.; Paulson, E.; Wand, A. J.; Zilm, K. W.; McDermott, A. E. Presented at the Experimental NMR Conference, Orlando, FL, 2001; Poster 107.
(33) McDermott, A. Presented at the XXth International Conference on Magnetic Resonance in Biological Systems, Toronto, Canada; 2002.

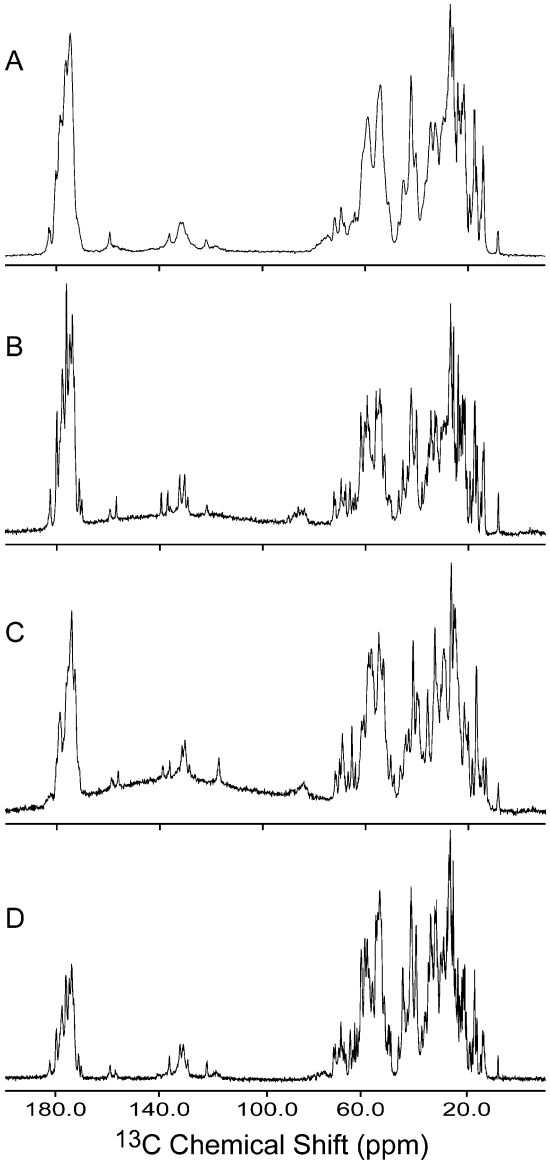

Figure 2. 1D ^{13}C CP MAS spectra of microcrystalline ubiquitin at 400 MHz (A−C) and 800 MHz (D). Spectra A and D were recorded on uniformly ^{13}C,^{15}N-enriched ubiquitin crystallized from MPD. Spectra B and C correspond to the natural abundance protein crystallized from MPD and PEG 8000, respectively. Although resolved splittings are not seen, carbon homonuclear couplings are the major source of apparent line width for our studies of uniformly enriched materials at 400 and 800 MHz, and the line widths depend dramatically on the isotopic composition.

for ^{13}C, and 40.2 MHz for ^{15}N, using a 4-mm triple resonance T3 MAS probe (Varian Instruments), with a spinning frequency of 9 or 10 kHz and a sample temperature of 0 °C. Ramped cross-polarization sequence was used in all cases.[34] TPPM proton decoupling[35] with proton field strength of 80 kHz was applied during the acquisition period.

800 MHz homonuclear ^{13}C spectra of microcrystalline ubiquitin were acquired using a Varian INOVA 800 MHz spectrometer (Yale University) outfitted with a home-built triple resonance probe with 2.5-mm rotors, ca. 19.8 kHz sample spinning frequency, and 89 kHz proton decoupling (TPPM). A sample temperature of 10 ± 5 °C was maintained during the experiment. The carbon excitation pulse was 3.9 μs (90° pulse), and detection windows were 100 kHz in both

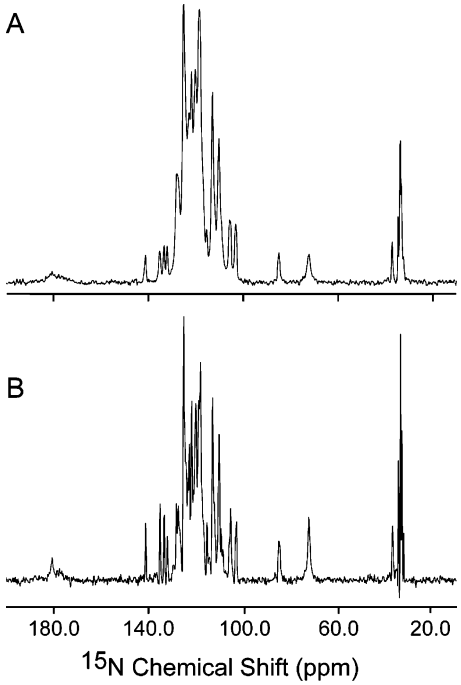

Figure 3. 1D ^{15}N CP MAS spectra of uniformly ^{13}C,^{15}N-enriched microcrystalline ubiquitin at 400 MHz (A) and 800 MHz (B).

dimensions. $^{13}C-^{13}C$ spin diffusion assisted by ^{1}H irradiation[1] was used as a mixing sequence in 2D experiments; a cartoon of this pulse sequence is shown in Figure 4A. The initial carbon excitation was accomplished with a square 500-μs cross-polarization followed by a 20-ms period of spin diffusion mixing, $\tau 1$, to equilibrate the magnetization among the carbons. Eighty scans were acquired per t1 increment, and 1024 increments were used with a recycle delay of 1.5 s. The total experiment time was 35.7 h. Spin diffusion assisted by ^{1}H irradiation was also used for the mixing period of the sequence, with a mixing time $\tau 2$ of 12 ms for the data shown in Figures 4B and 6 and 15 ms for the data shown in Figure 5. Spin diffusion mixing times were chosen to achieve $^{13}C-^{13}C$ transfers throughout the entire side chain. TPPI was used for phase-sensitive detection in the indirect dimension. Data were processed using FELIX (Accelrys, San Diego, CA). Sine-bell window functions with a 45° phase shift were applied in both dimensions. The spectra were referenced externally to TMS using the ^{13}C adamantane methylene peak at 38.56 ppm. The referencing was then readjusted to DSS assuming a 1.7 ppm chemical shift difference between TMS and DSS.[36] A value of 38.48 ppm for adamantane from TMS and an offset of −2.0 for DSS vs TMS have recently been determined; to join this reference scheme a value of 0.22 should be added to each value in the table.[37]

Results

1D ^{13}C spectra of ubiquitin at 800 MHz are shown in Figure 2, along with 400 MHz spectra of uniformly enriched and natural abundance materials. At lower applied field strength, the residual broadening is apparent for the uniformly labeled material (Figure 2A) compared with the natural abundance spectra (Figure 2B,C). The spectra of uniformly enriched materials are comparably better resolved at high magnetic field strength (Figure 2D). The broadenings for the uniformly labeled material at lower applied field strength are primarily due to homonuclear couplings, including J splittings (55 Hz for $^{13}C\alpha-^{13}CO$ and 35 Hz for

(34) Metz, G.; Wu, X. L.; Smith, S. O. *J. Magn. Reson., Ser. A* **1994**, *110*, 219−227.
(35) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951−6958.
(36) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135−140.
(37) Morcombe, C.; Zilm, K. W. *J. Magn. Reson.* **2003**, *162*, 479−486.

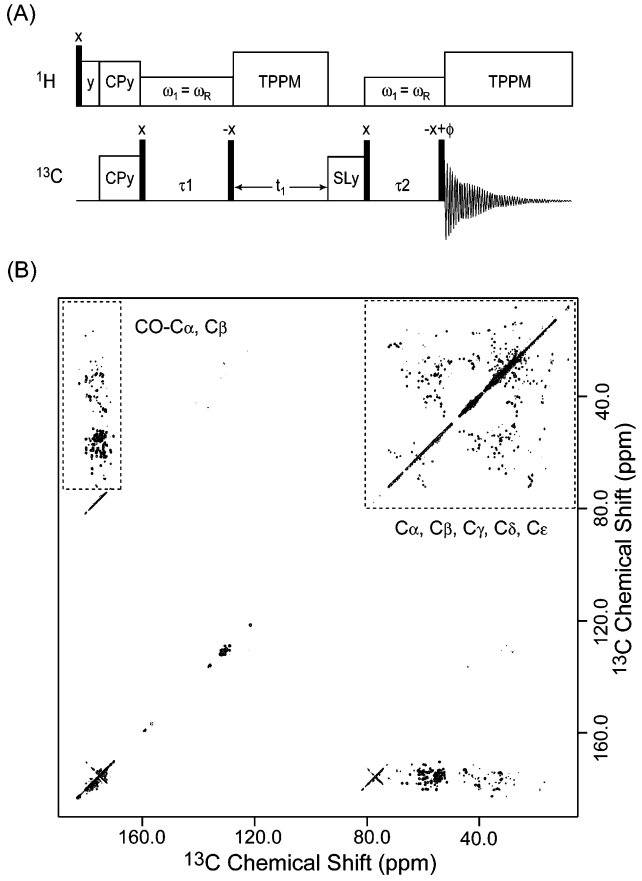

Figure 4. (A) Cartoon of the 2D DARR-based mixing pulse sequence; the experimental parameters are described in the Materials and Methods section. (B) 2D ^{13}C−^{13}C chemical shift correlation spectrum of uniformly ^{13}C,^{15}N-enriched ubiquitin taken at 800 MHz.

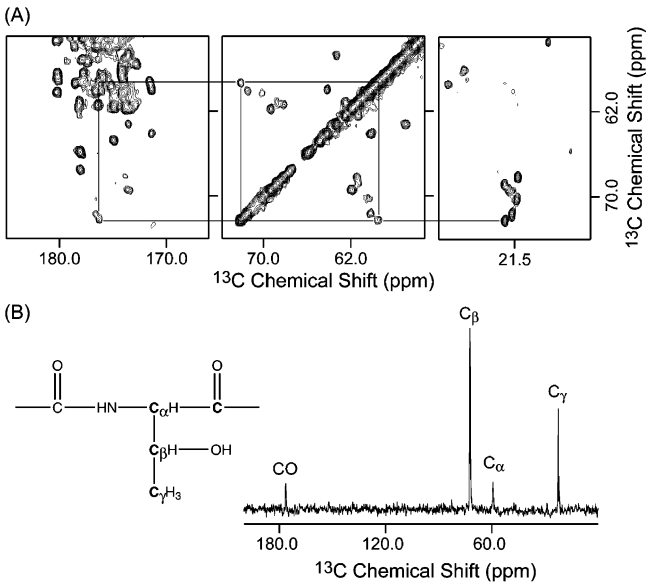

Figure 5. Side-chain "walk" documenting the Thr55 spin system. The expansions of a 2D ^{13}C−^{13}C spectrum corresponding to CO−Cβ, Cα−Cβ, and Cγ−Cβ regions are shown in (A). The spectrum was collected with slightly different conditions compared to those for Figure 4, namely, the number of scans per t1 increment was 16, the mixing time was 20 rather than 12 ms, and the total acquisition time was only 5.6 instead of 35.7 h. (B) Cross section through the Cβ carbon of Thr55.

^{13}Cα−^{13}Cβ), although resolved J splittings cannot be seen principally due to incomplete MAS averaging of the ^{13}C−^{13}C dipolar couplings. The contributions of ^{13}C−^{13}C couplings to the line widths of uniformly enriched materials have been discussed in the literature previously, including protocols for removing ^{13}C−^{13}C J couplings during indirect detection periods[2,38] and "block-labeling" isotope strategies.[39,40] In this study, we utilized free precession carbon detection, with proton decoupling, but no homonuclear decoupling, for both the direct and the indirect dimensions. Under these conditions, the ^{13}C full width half-maximal line widths in uniformly ^{13}C,^{15}N-enriched ubiquitin were of the order of 0.5 ppm at 800 MHz and 1.0−1.5 ppm at 400 MHz. The ^{13}C line widths of natural abundance material crystallized from MPD were of the order of 0.4 ppm at 400 MHz, giving spectral resolution comparable to that achieved at 800 MHz.

The choice of spinning frequency has an important effect upon the apparent line width for uniformly labeled samples due to the incomplete MAS averaging of the ^{13}C−^{13}C dipolar couplings. At high magnetic field strengths and spinning frequencies corresponding to 90−100 ppm at ^{13}C as used here, the spin system is far from any rotational resonance conditions. The spinning frequency far exceeds any dipolar couplings, and the residual dipolar line broadenings and second-order shifts can be expected to be smaller than the other line width contributions, roughly 30 Hz or less. In the following paragraph, we report simply the free precession peak positions, without any correction for dipolar shifts.

In Figure 2, we also compare the ^{13}C natural abundance spectra for the MPD (B) and PEG 8000 (C) crystallized ubiquitin sample. Differences in peak positions, of the order up to 0.5−1.0 ppm, can be seen comparing one formulation to another. Results have been previously reported for lyophilized rehydrated ubiquitin and with "checker-block labeled" lyophilized rehydrated material.[40] The sample crystallized from MPD showed narrow line widths, good dispersion, and good stability during long storage (ca. 6 months). In contrast to the preparation from PEG, no peak doubling was observed. These characteristics made this formulation a natural choice for the 2D experiments.

Figure 3 shows the ^{15}N CP MAS spectra of uniformly enriched ubiquitin, recorded at 400 (A) and 800 (B) MHz. Even at moderate magnetic field strength, spectral resolution is excellent, and several resolved peaks corresponding to individual amino acids can be seen. We have observed experimentally that the difference between ^{15}N CP MAS spectra of ubiquitin in lyophilized and microcrystalline states is much more dramatic than that for the ^{13}C CP MAS spectra. Possibly, ^{15}N spectra are more sensitive to the hydration state and conformational homogeneity of the sample than ^{13}C spectra. Typical ^{15}N line widths are 0.3−0.5 ppm at 800 MHz and 0.75−1.0 ppm at 400 MHz. On the basis of these data, we conclude that ^{15}N-detected 2D experiments, while having lower sensitivity than the ^{13}C-detected experiments, produce better spectral resolution.

The ^{13}C homonuclear correlation spectrum of uniformly enriched ubiquitin collected at 800 MHz is shown in Figure 4B, along with a cartoon for the pulse sequence used (A). A sample spinning frequency of 19.8 kHz was used so as to avoid rotational resonance conditions between the carbonyl carbons

(38) Straus, S. K.; Bremi, T.; Ernst, R. R. *Chem. Phys. Lett.* **1996**, *262*, 709−715.
(39) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255−9264.
(40) Hong, M.; Jakes, K. *J. Biomol. NMR* **1999**, *14*, 71−74.

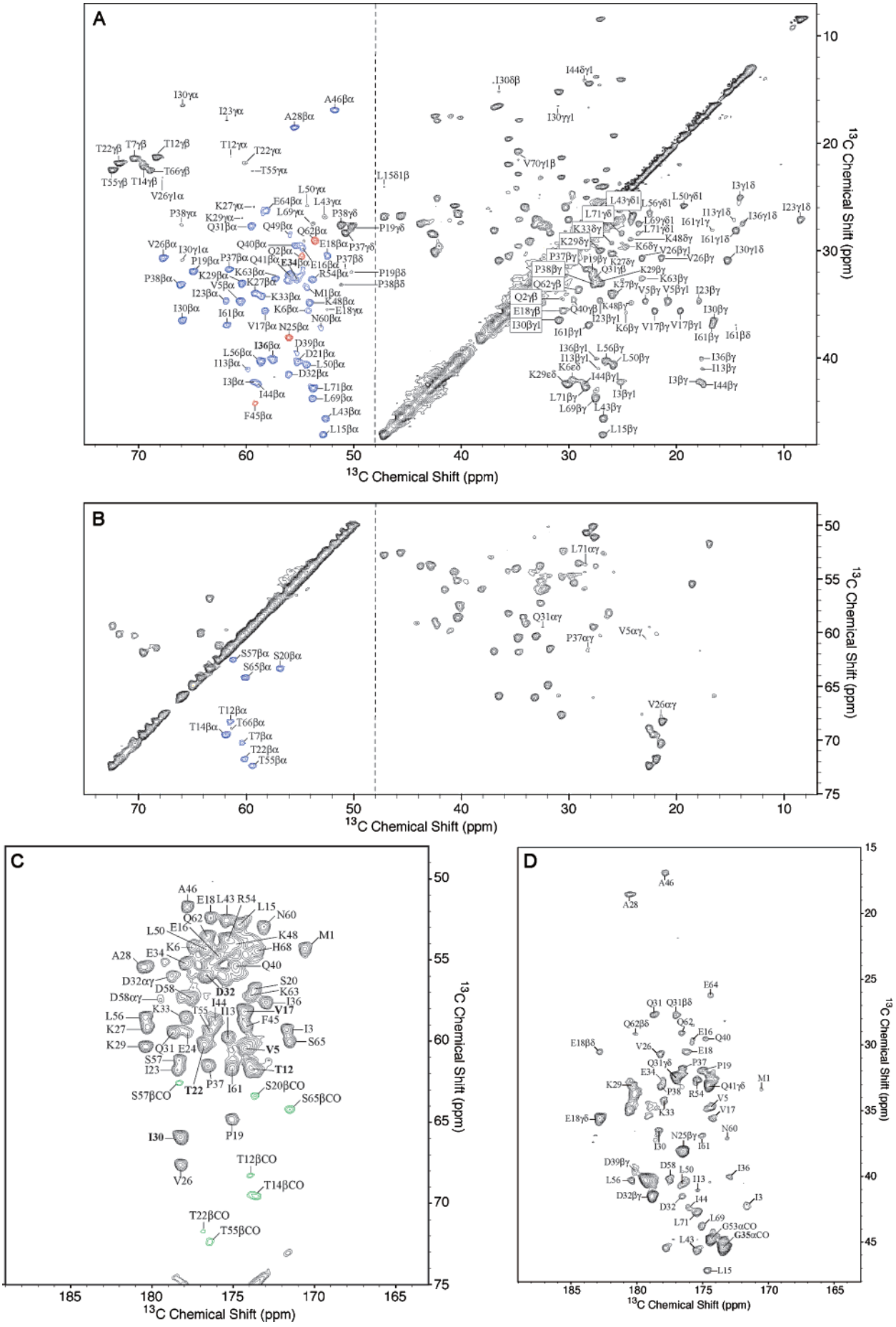
A
B
C
D
^{13}C Chemical Shift (ppm)
^{13}C Chemical Shift (ppm)
^{13}C Chemical Shift (ppm)
^{13}C Chemical Shift (ppm)

Figure 6. Expansion of the aliphatic region of the $^{13}C-^{13}C$ spin diffusion spectrum recorded at 800 MHz. (A) Chemical shift range of 7−75 ppm, 7−48 ppm, which includes all aliphatic cross-peaks, except for Cα−Cβ/Cβ−Cα resonances of threonines and serines. (B) Chemical shift range of 7−75 ppm, 48−75 ppm. Cβ−Cα cross-peaks, also present in the 3D NCACX spectrum, are shown in blue, while correlations found only in the $^{13}C-^{13}C$ diffusion data set are shown in red. Degenerate cross-peaks are labeled in bold and include Ile36 (not resolved from Asp58) and E34 (not resolved from E51). (C) Chemical shift range of 163−189 ppm, 48−75 ppm, including Cα−CO cross-peaks (except for those associated with glycines). Degenerate cross-peaks are labeled in bold and include Thr12 (not resolved from Thr14), Val17 (not resolved from Tyr59 and Glu64), Asp32 (not resolved from Gln41), Ile30 (not resolved from Pro38), Thr22 (not resolved from Thr7), and Val5 (not resolved from Val70). Two-bond Cβ−CO cross-peaks of threonines and serines are shown in green. (D) Chemical shift range of 163−169 ppm, 15−75 ppm, showing Cβ−CO connectivity. The Cα−CO resonance of Gly35 is not resolved from Gly47. Sequential assignments were obtained using a set of 3D $^{15}N-^{13}C-^{13}C$ experiments.

and any of the aliphatic carbons. Many prominent cross-peaks corresponding to directly bonded pairs can be seen in the 800 MHz spectrum, including the CO−Cα region, numerous Cα−Cβ cross-peaks. Many weak two-bond peaks are also observed; these are particularly noticeable in the CO−Cβ region. An example of a 2D "walk" using one of the threonine side chains is shown in Figure 5. Side-chain correlations are illustrated using a threonine Cβ slice as an example. Here the typical signal-to-noise ratio for a 5.6-h experiment can be seen; both one-bond and two-bond transfers are observed.

To estimate Cβ−Cα transfer efficiencies, we selected a set of residues that have resolved diagonal Cβ peaks. Those included five threonines (7, 12, 14, 22, and 55), three leucines (15, 43, and 69), and two serines (20 and 65). For each residue, we took a cross section through the corresponding Cβ chemical shift. The areas of the peaks in a given cross section were then measured, and the transfer efficiencies were calculated as the peak area normalized to the sum of all peaks including the diagonal one for the source spin. Cβ−Cα one-bond correlations typically had average normalized transfer efficiencies of 11% (7−16%), while the average efficiency of the Cβ−Cγ transfer was 17% (8−24%). Two-bond Cβ−CO transfers had an average efficiency of 6% (2−10%). The efficiency of one- and two-bond magnetization transfers shows that the DARR sequence has a far more uniform excitation profile than "passive" spin diffusion and is capable of exciting essentially all of the desired cross-peaks. On the other hand, it is not fully uniform. Pairs near an $n = 1$ condition (where the difference between isotropic chemical shifts, $\Delta\sigma_{iso}$, is equal to the spinning frequency, ω_r) give rise to systematically stronger cross-peaks than those near an $n = 0$ ($\Delta\sigma_{iso} = 0$) condition.

In Figure 6, we show expansions of the aliphatic or Cα−Cβ regions and Cα−CO regions from the 35.7-h experiment (i.e., the same data as in Figure 4), processed with resolution enhancement (shifted sine-bell with a 45° shift), to summarize graphically the spin systems that have been analyzed. In conjunction with 3D NMR data, reported in the accompanying paper, we identified the following 67 spin systems: Gly (3/6), Ala(2/2), Val (4/4), Leu (7/9), Ile (7/7), Thr (6/7), Ser (3/3), Pro (3/3), Phe(2/2), Asn (2/2), Asp (5/5), Gln(6/6), Glu (6/6), Lys (6/7), Arg (2/4), Met (1/1), His (1/1), Tyr (1/1) (no Trp nor Cys). Out of 67 spin systems, 54 could be unambiguously identified through correlations in isotropic shifts, including most of the expected two-bond cross-peaks. All chemical shifts for the spin systems highlighted in Figure 6 are listed in Table 1 as apparent shifts, referenced to DSS. The typical precision in these values was less than 0.1 ppm, as is expected based upon the digital resolution.

Sequence-specific assignment of the 2D $^{13}C-^{13}C$ spectrum was carried out using sequential assignments obtained from a set of 3D experiments carried out at 9.4 T (described separately). This 800 MHz 2D $^{13}C-^{13}C$ experiment provides independent verification of the residue types associated with the backbone chemical shifts and provided the majority of the side-chain carbons. Figure 6A shows an expansion of the aliphatic region of the 2D spectrum, with chemical shift ranges of 7−75 and 7−48 ppm in the direct and indirect dimensions, respectively. This region includes all correlations between the side-chain carbons and all Cβ−Cα correlations with the exception of threonines and serines (which are shown in Figure 6B). Cβ−Cα cross-peaks, which are also present in the 3D NCACX spectrum, are shown in blue; additional Cβ−Cα cross-peaks identified on the basis of Cα and CO chemical shifts are shown in red and include Gln2, Asn25, Phe45, and Gln62. Cβ−Cα cross-peaks that are degenerate are labeled in bold and include Ile36 (not resolved with Asp58) and Glu34 (not resolved with Glu51). Cα−Cβ (or Cβ−Cα) cross-peaks in this spectrum were in excellent agreement with peaks identified from 3D NCACX experiment used for sequential assignment, despite the differences in temperature, applied magnetic field strength, and radio frequency duty cycle. The average standard deviation of carbon chemical shifts within solid-state NMR data sets is 0.11 ppm. It is this agreement that immediately led to the site-specific assignment of almost all resonances in the aliphatic region. In all ubiquitin spectra no evidence for peak "doubling" was observed, as might have been expected for heterogeneous samples; no peaks went unaccounted for.

In contrast to well-resolved aliphatic cross-peaks, the Cα−CO region remains fairly congested even at high applied magnetic field strength, as evident from Figure 6C. All Cα−CO cross-peaks are labeled with just residue numbers. Two-bond Cβ−CO correlations of threonines and serines are shown in green to distinguish them from Cα−CO cross-peaks. Degenerate cross-peaks are labeled in bold and include Thr12 (not resolved with Thr14), Val17 (not resolved with Tyr59 and Glu64), Asp32 (not resolved with Gln41), Ile30 (not resolved with Pro38), Thr22 (not resolved with Thr7), and Val5 (not resolved with Val70). Despite significant spectral overlap, the Cα−CO region proved to be very useful in the verification of sequential assignments, especially when used in conjunction with the Cβ−CO region of the same spectrum, shown in Figure 6D. Clear Cβ−CO cross-peaks are observed for a number of residues; pairs of bidirectional two-bond transfers (β−CO and β−δ) are clearly detected for Gln31, Gln62, and Glu18 side chains. The large cross-peak at (179.3, 40.3) ppm most likely represents one-bond β−γ transfers within Asp21 and Asp58 side chains. Notice that the Cα−CO glycine region is poorly resolved. On the basis of our sequential assignment, we conclude that Gly35 (labeled in bold) is not resolved with Gly47. Three glycine residues belong to either a loop region (Gly10) or C-terminus of the protein (Gly75 and Gly76), with both regions systematically missing from all four data sets.

There are two stretches of amino acids that are systematically missing from our 2D and 3D data sets: the first region includes

Table 1. Isotropic Chemical Shifts for Carbon Resonances in Microcrystalline Ubiquitin, Citrate Buffer, and MPD at pH 4.2[a]

residue	Cα	Cβ	Cγ/Cγ1	Cδ/Cδ1	Cε	CO	residue	Cα	Cβ	Cγ/Cγ1	Cδ/Cδ1	Cε	CO
Met1	54.4	33.4	–		–	170.7	Asp39	55.2	39.6	180.1			176.9
	(54.7)	(33.4)				(171.4)		(56.0)	(40.1)	(nr)			(177.9)
Gln2	54.8	30.6	34.5	–		176.4	Gln40	55.3	29.6	34.7			175.0
	(55.4)	(30.9)	(34.8)			(176.9)		(55.8)	(29.9)	(34.6)			(176.3)
Ile3	59.3	42.3	17.9/25.2	14.1		171.7	Gln41	55.9	32.4	33.1	174.5		176.9
	(59.9)	(42.2)	(18.2/25.2)	(14.6)		(173.2)		(56.9)	(30.8)	(33.6)	(nr)		(177.0)
Phe4	54.1	–	–	–	–	175.4	Arg42	54.5	–	–	–		174.5
	(55.3)					(176.0)		(55.4)					(174.7)
Val5	60.4	34.7	22.9/20.7			174.5	Leu43	52.5	45.7	26.8	24.1		175.5
	(60.7)	(34.3)	(22.7/21.1)			(175.7)		(53.3)	(45.8)	(26.9)	(24.4)		(176.2)
Lys6	54.2	35.7	24.7	29.6	42.0	177.6	Ile44	58.7	42.5	17.5/28.6	14.2		176.3
	(54.8)	(34.6)	(25.2)	(29.5)	(42.4)	(178.0)		(59.2)	(41.2)	(17.8/28.0)	(12.9)		(176.7)
Thr7	60.3	70.3	21.4			176.7	Phe45	59.1	44.2	–	–	–	174.4
	(60.7)	(70.8)	(21.6)			(177.8)		(56.8)	(43.9)				(175.5)
Thr12	61.5	68.3	21.3			174.1	Ala46	51.8	17.0				177.9
	(62.6)	(70.1)	(22.3)			(175.2)		(52.8)	(16.8)				(178.2)
Ile13	59.9	41.1	17.6/27.2	14.6		175.5	Gly47	45.5					173.7
	(60.3)	(41.0)	(17.9/27.2)	(14.6)		(176.1)		(45.5)					(174.6)
Thr14	61.8	69.5	22.3			173.9	Lys48	53.9	34.9	24.2	29.0	–	175.0
	(62.3)	(70.0)	(22.1)			(174.6)		(54.8)	(34.6)	(24.7)	(29.4)		(175.5)
Leu15	52.9	47.2 (47.1)	27.0	24.1		174.7	Gln49	55.8	28.7	–	–		176.1
	(53.0)		(27.2)	(24.4)		(175.4)		(56.2)	(nr)				(176.4
Glu16	54.6	29.8	35.4	–		175.9	Leu50	54.3	40.7	25.9	19.3		176.9
	(55.2)	(29.8)	(35.8)			(176.7)		(54.5)	(41.6)	(26.2)	(19.9)		(177.5)
Val17	58.2	35.6	22.0/19.7			174.3	Glu51	55.5	32.4	37.8	–		175.6
	(58.8)	(36.6)	(22.5/19.8)			(174.9)		(56.1)	(29.4)	(34.8)			(176.3)
Glu18	52.4	30.5	35.6	182.8		176.5	Asp52	55.8	–	–			177.6
	(53.0)	(31.0)	(35.3)	(nr)[b]		(nr)		(56.9)					(178.3)
Pro19	64.8	32.0	27.8	50.2		175.2	Gly53	44.6					174.4
	(65.6)	(31.9)	(28.1)	(50.7)		(176.2)		(45.4)					(175.6)
Ser20	56.8	63.4				173.8	Arg54	53.9	32.7	–	–		175.5
	(57.7)	(63.9)				(175.5)		(54.5)	(33.0)				(176.2)
Asp21	55.1	40.4				176.1	Thr55	59.4	72.4	22.6			176.6
	(56.1)	(41.1)				(177.2)		(60.0)	(72.8)	(22.3)			(177.4)
Thr22	60.0	71.8	21.9			177.0	Leu56	58.6	40.3	26.5	22.5		180.5
	(60.0)	(71.2)	(22.7)			(177.6)		(58.9)	(40.5)	(27.0)	(23.4)		(181.6)
Ile23	61.7	34.7	17.9/27.2	8.5		178.7	Ser57	61.2	62.6				178.4
	(62.7)	(34.8)	(18.4/28.3)	(9.5)		(179.8)		(61.4)	(62.8)				(179.1)
Glu24	59.5	–	–	–		177.9	Asp58	57.3	40.2	179.5			177.7
	(61.0)					(179.9)		(57.7)	(40.7)	(nr)			(178.3)
Asn25	55.9	38.1	176.4			177.0	Tyr59	58.6	–	–	–	–	174.6
	(56.3)	(38.7)	(nr)			(179.2)		(58.6)					(175.5)
Val26	67.8	30.7	21.4/23.2			178.3	Asn60	53.0	37.2	–			173.2
	(67.9)	(31.0)	(21.8/23.9)			(178.8)		(54.3)	(37.7)				(175.1)
Lys27	59.1	34.1	26.0	30.3	–	180.4	Ile61	61.8	37.0	16.7/28.1	14.5		175.1
	(59.5)	(33.7)	(nr)	(nr)		(181.4)		(62.7)	(36.9)	(17.5/28.5)	(14.5)		(175.4)
Ala28	55.4	18.6				180.6	Gln62	53.5	29.1	33.6	180.1		176.6
	(55.6)	(18.0)				(181.1)		(53.9)	(31.8)	(33.7)	(nr)		(176.7)
Lys29	60.3	33.1	27.1	30.1	42.5	180.5	Lys63	57.2	32.7	23.2	29.8	–	174.2
	(60.1)	(33.6)	(24.8)	(26.7)	(42.9)	(181.1)		(58.2)	(32.8)	(24.1)	(29.8)		(176.6)
Ile30	65.8	36.6	16.6/31.0	15.3		178.5	Glu64	58.0	26.4	–	–		174.5
	(66.4)	(36.9)	(17.4/31.3)	(15.5)		(179.1)		(58.7)	(26.3)				(176.1)
Gln31	59.4	27.7	32.5	177.1		178.8	Ser65	60.0	64.3				171.7
	(60.4)	(28.0)	(34.1)	(nr)		(179.7)		(61.2)	(65.7)				(172.9)
Asp32	55.9	41.6	178.8			176.6	Thr66	61.2	68.9	22.6			174.3
	(57.7)	(41.4)	(nr)			(178.2)		(62.7)	(70.6)	(21.8)			(174.6)
Lys33	58.5	34.3	26.0	29.3	–	178.0	Leu67	53.7	–	–	–		175.7
	(58.5)	(34.3)	(25.4)	(29.1)		(178.7)		(54.0)					(176.2)
Glu34	55.1	32.6	–	–		178.0	His68	54.3	28.0 (32.6)	–	–	–	173.7
	(55.6)	(33.5)				(178.8)		(56.3)					(174.6)
Gly35	45.4					173.5	Leu69	53.9	43.8	27.5	23.5		175.2
	(46.2)					(174.8)		(54.0)	(44.5)	(27.7)	(24.1)		(176.2)
Ile36	57.4	40.1	17.7/27.5	13.8		173.2	Val70	60.2	34.7	21.6/20.7			174.3
	(58.0)	(40.7)	(18.0/27.3)	(13.8)		(nr)		(60.9)	(35.0)	(21.6/20.9)			(174.9)
Pro37	61.5	31.8	28.4	50.8		176.7	Leu71	53.8	42.8	28.4	25.1/23.6		175.4
	(61.8)	(32.1)	(28.3)	(51.3)		(nr)		(54.2)	(43.0)	(27.6)	(25.2/24.1)		(178.7)
Pro38	66.0	33.2	27.7	51.2		178.3							
	(66.5)	(33.0)	(27.9)	(51.4)		(179.1)							

[a] Site-specific assignments are proposed based on the set of 3D experiments reported separately. Solution values (phosphate buffer, pH 5.7) from Wand et al.[11] are entered in parenthesis. [b] "nr" stands for not reported.

residues Leu8-Lys11 and corresponds to the β-turn between the first two β-strands of ubiquitin, and the second region includes five residues of ubiquitin C-terminus, Arg72-Gly76. Both "missing" regions show depressed values of generalized order parameters in solution NMR,[12] which is indicative of subnanosecond backbone motions (this is in agreement with other solution NMR studies of ubiquitin[41,42]). Reported main chain crystallographic B factors mirror this trend.[8] For four side chains we consider the side chain assignments to be not secure: Glu51, His68, Gln41, and Thr66. To summarize, we have assigned the following spin systems in microcrystalline ubiquitin: Gly (3/6), Ala (2/2), Val (4/4), Leu (7/9), Ile (7/7), Thr (6/7), Ser (3/3), Pro (3/3), Phe (2/2), Asn (2/2), Asp (5/5), Gln(6/6), Glu (6/6), Arg (2/4), Lys (6/7), Met (1/1), His (1/1), Tyr (1/1) (no Trp nor Cys). In total, 67 out of 76 residues are assigned to various degrees of completion.

To compare the ^{13}C chemical shifts of ubiquitin in solution and solid state, for all protein sites that we consider to be securely assigned in the solid state, we calculated the chemical shift difference between our solid-state values and the solution values reported by Wand et al.[11] The main differences between solution and solid sample conditions are the pH (5.7 versus 4.0−4.2) and buffer (potassium phosphate versus citrate). We then constructed a histogram of chemical shift differences, setting the bin size to 0.5 ppm. The average standard deviation of ^{13}C chemical shifts, taken over all cross-peaks with assignment count larger than two, is 0.11 ppm. Thus, the bin size is approximately 5 times the standard deviation of the mean carbon chemical shift, which ensures >95% probability that a given value of chemical shift falls into the bin. The histogram was fit with a Gaussian function, and the residues falling outside of the 89% confidence interval were considered to be outliers. The histogram with a superimposed Gaussian fit is shown in Figure 7, with outliers for the side-chain Cβ shift mapped onto the crystal structure of ubiquitin. These outliers are largely solvent-exposed ionizable residues. An analogous analysis of CO and Cα chemical shifts has been reported separately and gave a similar standard deviation of the mean σ, ~0.4 ppm. Since there is no crystal structure data for ubiquitin under conditions similar to those of our solid-state NMR experiments, we cannot establish if crystal-packing forces are responsible for the observed perturbation of chemical shifts, but this remains a strong hypothesis.

It is worth noting that it was possible to identify the site-specific assignments correctly for the majority of sites simply by comparison with solution shifts. When the solution[11] and solid-state chemical shifts were similarly referenced, the RMS deviation between corresponding carbon shifts for *random* pairs of a particular side-chain spin system (i.e., two leucines) was typically 5−10 ppm. In many cases, the carbon shifts recorded in the solid state matched those from one of the spin systems reported based on solution NMR data, with an RMS value of less than 1 ppm. Therefore, for proteins with assigned solution NMR spectra, the corresponding solution and solid-state peaks can be correctly identified without de novo assignments in the vast majority of cases.

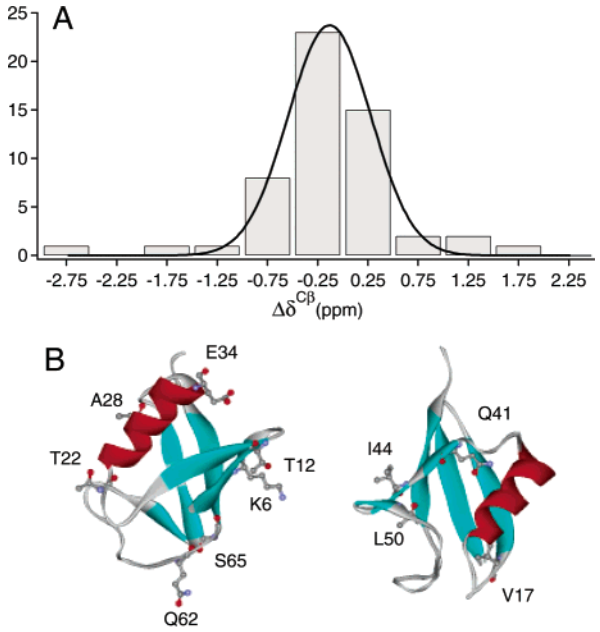

Figure 7. Comparison of solution and solid-state ^{13}Cβ chemical shifts of ubiquitin. (A) Histogram of the observed ^{13}Cβ chemical shift differences between solid- and solution-state NMR. The chemical shift difference is defined as $\Delta\delta^{C\beta} = \delta^{C\beta}_{solid} - \delta^{C\beta}_{solution}$. The solid line represents a fit of the histogram by a Gaussian function $y = A \exp\{-[(x - x0)/width]^2\}$. The best-fit parameters for $x0$ and *width* were found to be -0.145 ± 0.019 and 0.583 ± 0.026, respectively. The standard deviation of the mean σ is trivially calculated from the *width* parameter and is equal to 0.412 ± 0.018. (B) Amino acids with "perturbed" values of Cβ chemical shifts are shown in ball-and-stick representation. Left: Q62, S65, K6, T12, E34, A28, and T22. Right: V17, Q41, I44, and L50.

Conclusions

In this work we have achieved high-quality spectra of a microcrystalline protein, and narrow line widths, good dispersion, and long-term stability were possible for this microcrystalline form of uniformly ^{13}C,^{15}N-enriched ubiquitin. Recently, the first solid-state NMR structure of a 62-residue protein, SH3 domain of α-spectrin, was reported.[43] In that work, ^{13}C−^{13}C distance restraints were obtained from spin diffusion buildup curves and used as an input for structure calculation. Elegant methodology has also been applied to a biologically relevant tripeptide, which resulted in the determination of its 3D structure.[44] Despite these landmark achievements, more methods are needed for efficient and precise determinations of structure and dynamics of proteins in the solid state. With the crystallization conditions well-mapped and the assignment table largely "filled in", ubiquitin promises to be an excellent system for probing new pulse sequence methods for structure and dynamics, just as it has been for solution NMR studies.

Acknowledgment. We thank Dr. Charles Babu (University of Pennsylvania) for his help with solution NMR experiments. A.E.M. and T.I.I. were supported by a grant from the National Science Foundation (MCB 9983581). A.J.W. was supported by NIH Grant DK39806. We thank Tatyana Polenova for help in the early stages of this project.

JA030547O

(41) Lienin, S. F.; Bremi, T.; Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, *120*, 9870−9879.
(42) Fushman, D.; Tjandra, N.; Cowburn, D. *J. Am. Chem. Soc.* **1999**, *121*, 8577−8582.
(43) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98−102.
(44) Rienstra, C. M.; Tucker-Kellogg, L.; Jaroniec, C. P.; Hohwy, M.; Reif, B.; McMahon, M. T.; Tidor, B.; Lozano-Perez, T.; Griffin, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10260−10265.
(45) Martin, R. W.; Zilm, K. W. *J. Magn. Reson.* **2003**, *165*, 162−174.