JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Structure Calculation from Unambiguous Long-Range Amide and Methyl $^{1}H-^{1}H$ Distance Restraints for a Microcrystalline Protein with MAS Solid-State NMR Spectroscopy

Rasmus Linser,[†] Benjamin Bardiaux,[‡] Victoria Higman,[§] Uwe Fink,[‡] and Bernd Reif[*,‡,∥,⊥]

[†]University of New South Wales (UNSW), Analytical Centre, Chemical Sciences Building, NSW 2052, Australia

[‡]Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), Robert-Rössle-Strasse 10, 13125 Berlin, Germany

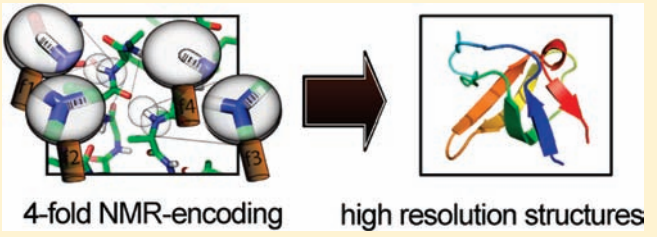
[§]Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom

[∥]Munich Center for Integrated Protein Science at Department Chemie, Technische Universität München (TUM), Lichtenbergstrasse 4, 85747 Garching, Germany

[⊥]Helmholtz-Zentrum München (HMGU), Deutsches Forschungszentrum für Gesundheit und Umwelt, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

Supporting Information

ABSTRACT: Magic-angle spinning (MAS) solid-state NMR becomes an increasingly important tool for the determination of structures of membrane proteins and amyloid fibrils. Extensive deuteration of the protein allows multidimensional experiments with exceptionally high sensitivity and resolution to be obtained. Here we present an experimental strategy to measure highly unambiguous spatial correlations for distances up to 13 Å. Two complementary three-dimensional experiments, or alternatively a four-dimensional experiment, yield highly unambiguous cross-peak assignments, which rely on four encoded chemical shift dimensions. Correlations to residual aliphatic protons are accessible via synchronous evolution of the ^{15}N and ^{13}C chemical shifts, which encode valuable amide−methyl distance restraints. On average, we obtain six restraints per residue. Importantly, 50% of all restraints correspond to long-range distances between residues i and j with $|i-j| > 5$, which are of particular importance in structure calculations. Using ARIA, we calculate a high-resolution structure for the microcrystalline 7.2 kDa α-spectrin SH3 domain with a backbone precision of ∼1.1 Å.



INTRODUCTION

Over the past few years, solid-state NMR has evolved as a valuable tool for structure determination of nonsoluble proteins. NMR structures of proteins are usually calculated using internuclear distance information. In the solid-state, this distance information is usually obtained from ^{13}C spin dilute samples[1,2] or from heteronucleus-detected NHHC or CHHC-like experiments utilizing spin diffusion among protons.[3−5] Direct determination of $^{1}H-^{1}H$ dipolar couplings and the quantification of the respective distances are achieved via dipolar recoupling experiments.[6] Detection of protons in the direct dimension is compromised by dipolar line broadening but can be enabled by a reduction of the overall proton density via perdeuteration and partial proton back exchange at labile sites.[7−9] This approach results in an increased lifetime of the NMR spin states and thus yields narrow resonance lines. To refine the structure of a microcrystalline sample of GB1, a three-dimensional (3D) version of a CONH experiment with a proton−proton radio frequency-driven dipolar recoupling (RFDR) mixing period has been suggested.[10] For deuterated proteins that contain 100% protons at labile sites, proton line widths still amount to ∼150 Hz, even at magic-angle spinning (MAS) rotation frequencies of 40 kHz. However, in order to separate a multitude of cross-peaks and for an accurate quantification of the peak volumes, a much higher spectral resolution is required. For larger proteins, cross-peak assignment becomes even more challenging, and experiments with higher dimensionality are highly recommended. In solution-state NMR, this is achieved by recording pairs of 3D experiments that yield ^{1}H and heteronuclear ($^{13}C/^{15}N$) chemical shift encoded amide−amide or amide−aliphatic contacts in nuclear Overhauser enhancement spectroscopy-heteronuclear single-quantum coherence (NOESY-HSQC)[11,12] and HSQC-NOESY-HSQC-type experiments.[13,14] This approach has been extended by synchronous encoding of the heteronuclear chemical shift,[15−17] leading, e.g., to 4D HSQC-NOESY-HSQC-type experiments employing time-shared ^{13}C and ^{15}N evolution periods.[18,19] In this manuscript, we present 3D and 4D experiments for protein structure determination in the solid state that allow an unambiguous assignment of magnetization exchange

Received: November 14, 2010
Published: March 24, 2011

   dx.doi.org/10.1021/ja110222h | J. Am. Chem. Soc. 2011, 133, 5905–5912

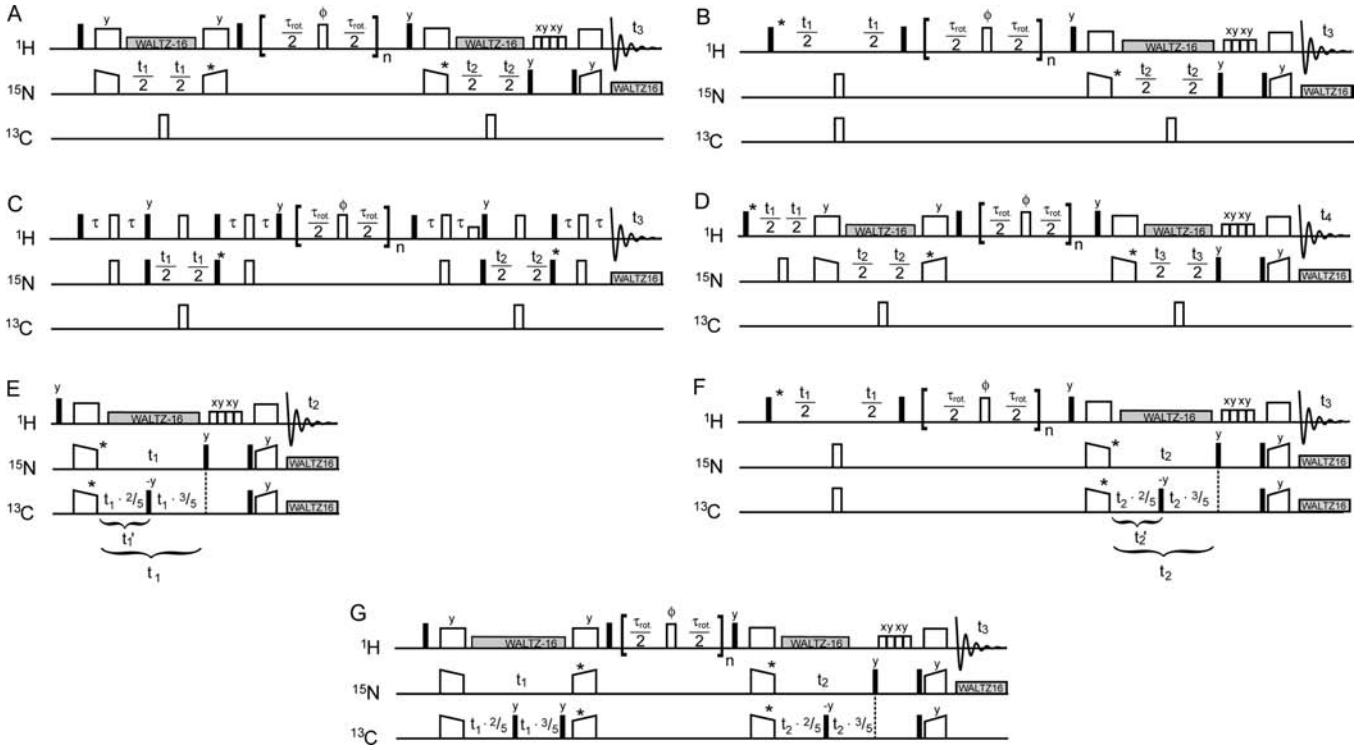


Figure 1. Pulse schemes for 3D $^{1}H-^{1}H$ RFDR experiments and the time-shared reference experiment with synchronous $^{13}C/^{15}N$ evolution. (A) 3D N···N−H correlation experiment. (B) 3D H···N−H correlation experiment. Both experiments use CP for magnetization transfer. CP transfers are more efficient for samples with short proton T_2 times and facilitate effective suppression of water artifacts. Water suppression in CP-based schemes is achieved using a scheme proposed by Zhou et al.[29] (C) INEPT-based version of experiment (A). INEPT-based sequences allow the observation of flexible residues that are obscured in CP-based experiments (see Figure 2, Supporting Information). Water suppression is achieved by application of a 1 ms spin-lock pulse. (D) 4D H−N···N−H correlation experiment. For this experiment, a double resonance probe was employed. Therefore, no ^{13}C pulses could be applied. (E) Time-shared $^{1}H-^{15}N/^{13}C$ correlation experiment for synchronous evolution of the ^{15}N and ^{13}C chemical shifts. Increments for ^{13}C chemical shift evolution (t_1') are set to 2/5 with respect to increments employed for ^{15}N to account for the different gyromagnetic ratios. Except for the 2D reference experiment, $t_1^{max}(^{15}N)$ = 25 ms and $t_1^{max}(^{13}C)$ = 10 ms are employed; t_1^{max} on the ^{13}C channel is set shorter, as evolution of $^{13}C-^{13}C$ scalar couplings would decrease sensitivity. Simultaneous $^{1}H-^{15}N$ and $^{1}H-^{13}C$ magnetization transfers are optimized by matching the ^{15}N and ^{13}C rf fields to the $(n-1)$ Hartmann−Hahn resonance condition. (F) Time-shared 3D H···N/C−H correlation experiment using cross-polarization for magnetization transfer. (G) Time-shared 3D N/C···N/C−H correlation experiment. All time-shared experiments are based on simultaneous $^{1}H-^{15}N/^{1}H-^{13}C$−CP transfers. Alternatively, INEPT can be employed for magnetization transfer.[19]

cross-peaks and an accurate quantification of peak volumes. The experiments encode chemical shifts in four dimensions in a time-shared fashion, yielding amide−amide and amide−methyl contacts. The proposed method is demonstrated using a microcrystalline sample of the 7.2 kDa SH3 domain of chicken α-spectrin.

EXPERIMENTAL SECTION

NMR. Partial proton back exchange at labile sites in otherwise deuterated proteins has been shown to yield long-lived proton spin coherences and excellent resolution.[9,20−22] The optimal proton content at exchangeable sites for the deuterated chicken α-spectrin SH3 domain is obtained by precipitating the protein from a buffer containing 25% H_2O/75% D_2O.[23] Expression and purification of the protein was carried out as described previously,[24] and 75 mM $[Cu(edta)]^{2-}$ was added to the mother liquor in order to achieve higher repetition cycles in the NMR experiments.[25−28] In this way, high sensitivity and resolution is obtained in a comparatively short time. Experimental ^{1}H and ^{15}N line widths are on the order of 20−25 and 10−12 Hz, respectively. Typically, a recycle delay of 220 ms is sufficient to recover magnetization.[26] All experiments (except the 4D) were recorded on a Bruker Avance spectrometer operating at a proton Larmor frequency of 600 MHz, equipped with a 3.2 mm triple resonance probe. The effective temperature was set to 22 and 25 °C and the MAS rotation frequency was adjusted to 24 and 20 kHz for standard and time-shared experiments, respectively. Approximately 18 mg of protein was center packed into a 3.2 mm MAS rotor. The 4D experiment was recorded on a Bruker Avance 700 MHz spectrometer at an effective temperature of 10 °C, using a 1.3 mm double resonance probe. The 1.3 mm rotor was filled with 2 mg of protein. The MAS rotation frequency was adjusted to 24 kHz. All pulse programs are depicted in Figure 1.

Structure Calculation from Unambiguous Manual Assignments. Data processing and peak assignment were conducted using Topspin-3 and CcpNmr Analysis,[30] respectively. An ensemble of structures was calculated with CNS[31] from manually assigned $^{1}H-^{1}H$ restraints derived from 3D H···N/C−H spectra. Chemical shifts for all assigned residues were taken from the literature[24,32] and used as input for TALOS+[33] to yield dihedral angle restraints. Assignments of exchange cross-peaks were obtained by combining the spectral information obtained from 3D H···N/C−H and N/C···N/C−H spectra. We employed the simulated annealing protocol of ARIA[34] to calculate 250 conformers with 216 000 cooling steps. Unlike previous work,[4] the ARIA protocol was used only as a platform for a standard structure calculation, i.e., the automated assignment module was not employed at this stage. The 20 lowest-energy structures were refined in a shell of water molecules,[35] and the 10 best structures were analyzed as the final

dx.doi.org/10.1021/ja110222h | J. Am. Chem. Soc. 2011, 133, 5905–5912

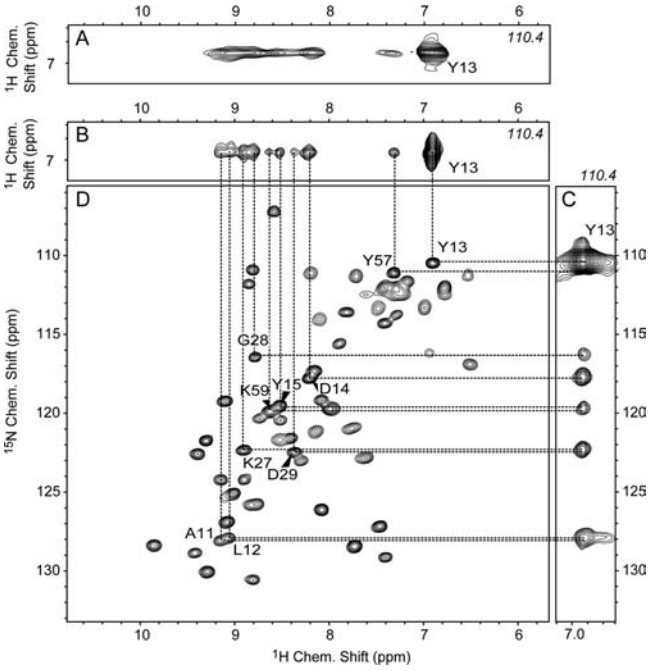


Figure 2. 3D RFDR spectra obtained for α-spectrin SH3. (A) Proton chemical shift encoded contacts for Y13 employing the pulse scheme shown in Figure 1B; t_1^{max} was limited to 10 ms in the ^{1}H evolution period yielding ^{1}H line widths of ∼100 Hz. The spectrum was obtained within 4 h. (B) Same as in (A) but with t_1^{max} = 55 ms (∼25 Hz ^{1}H line widths), which corresponds to the maximum achievable resolution. (C) ^{15}N chemical shift encoded contacts for Y13 obtained using the pulse scheme shown in Figure 1A (t_1^{max} = 20 ms, experimental time: 4 h). For all three experiments, t_2^{max} amounted to 10 ms. **D)** Reference H/N-HSQC, recorded in approximately 10 min. All spectra were taken by setting the effective sample temperature to 22 °C and by adjusting the MAS rotation frequency to 24 kHz. All strips are F1/F3 planes. Numbers in italics correspond to the Y13 ^{15}N chemical shift (F2).

structure ensemble. Manually assigned ^{1}H−^{1}H correlations were translated into distance restraints with large upper bounds (u, with u = 8 and 12 Å for mixing times of 3 and 8 ms, respectively). On the basis of the TALOS+ restraints alone, an ensemble of 250 conformers was first calculated with a simulated annealing protocol in torsion angle space. The maximum distance (m_{ij}) observed in this structure ensemble for nuclei i and j was then determined for each distance restraint. For the final structure calculation, default upper bounds (u) of distance restraints were replaced by the corresponding maximum distance (m_i) in all cases where $m_i < u$.

Automated Assignment and Distance Calibration from Ambiguous ^{1}H−^{1}H Contacts. The iterative ARIA protocol was further used to determine an ensemble of structures from ambiguous data, exclusively employing unassigned 3D H···N/C−H RFDR spectra. Instead of the default NOE-derived "r^{-6} summation", we used a r^{-3} relationship both for the analysis of ambiguous distance restraints and for the structure calculation. Chemical shift tolerances for the initial assignment were 0.05 ppm in the ^{1}H dimension and 0.4 ppm in ^{15}N. Upper bounds for the distance restraints were derived from the intensity I_{ij} of the cross-peak between atom i and j using a uniform calibration procedure. The calibration was carried out independently for the spectra recorded with mixing times of 3 and 8 ms, using the following relationship:

$$I_{ij} = kd_{ij}^{-3} \quad (1)$$

The calibration factor k was determined from the ratio of the average experimental and theoretical intensities:

$$k = \frac{\langle I_{ij} \rangle}{\langle d_{ij} \rangle^{-3}} \quad (2)$$

Without prior knowledge of the distances, $\langle d_{ij} \rangle$ was chosen to be 3.2 Å. Upper bounds were derived from the calibrated distances (eq 1) by adding a tolerance of 20%. Calibrated upper bounds larger than default upper bounds (8 Å for 3 ms, 12 Å for 8 ms) were capped to their default values. The final ensemble of the last ARIA iteration consists of 100 structures, of which the 10 lowest-energy conformers were refined in water and analyzed. In each case (manual and automated calculation), peaks that could be unambiguously assigned to intermolecular correlations between neighboring molecules in the crystal lattice were identified and removed. The Ramachandran plot statistics of the final structure ensembles was analyzed with the software PROCHECK.[36]

RESULTS

3D Experiments. Restraints for NMR structure calculations rely on the quantification of distances between pairs of nuclei. This information is typically encoded in the intensity of a single cross-peak per pair of nuclei. In addition to their correct assignments, cross-peaks need to be sufficiently resolved to allow for correct quantification of their intensities. For an unambiguous characterization of the distances between amino acids that are close in space, we developed a set of two complementary 3D experiments, yielding the ^{1}H and the ^{15}N chemical shifts of both involved pairs of nuclei for each amide−amide contact. In addition, a corresponding 4D experiment was implemented (see Figure 1 for all pulse programs). Figure 2 shows the respective strips for Y13 from the 3D H···N−H and N···N−H experiments. The availability of both the ^{1}H and ^{15}N chemical shift prior to the mixing step makes the identification of amide contacts fast and highly unambiguous. Figure 2A depicts the resolution obtained for the ^{1}H chemical shift encoded experiment after 4 h (using a two-step phase cycle), whereas Figure 2B shows the optimum achievable resolution. The ^{15}N encoded experiment in Figure 2C was recorded in 4 h as well. Due to a lower spectral width of only 1620 Hz (corresponding to 27 ppm at 600 MHz), the same spectral resolution is obtained much faster in comparison to the ^{1}H encoded experiment. Given the high sensitivity, both experiments are in the resolution-limited rather than the sensitivity-limited regime.

The ^{1}H- and ^{15}N-edited 3D RFDR experiments can also be employed to assist sequential assignments where ^{13}C-based backbone experiments[9] do not yield unambiguous assignments (Figure 1, Supporting Information). In addition to experiments that employ cross-polarization (CP) for magnetization transfer, scalar coupling based versions of these experiments were carried out (see Figure 1C for the pulse programs). Although water suppression is more of an issue in this case, flexible residues can only be detected if INEPT-based experiments are employed. This way, the N-terminal region of the protein (residues T4-L8) is readily assigned (Figure 2, Supporting Information), which is not detectable in CP-based experiments. The through-space contacts obtained for these residues confirmed the assignment obtained from transverse relaxation-optimized spectroscopy (TROSY)-type experiments.[37] INEPT-based experiments thus complement CP-based pulse schemes very well. In general, however, the dipolar coupling based strategies seem more widely applicable, since the transfer efficiency is then less dependent on

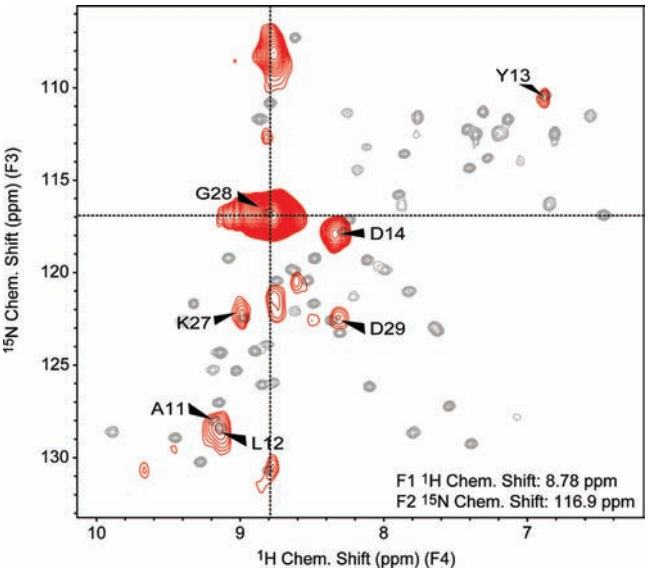


Figure 3. 2D F3/F4 plane (red) of the 4D H−N···N−H experiment extracted at the F1/F2 chemical shifts of G28 in the α-spectrin SH3 domain (8.78 and 116.9 ppm, indicated by dashed lines). The amides in spatial proximity appear with their ^{1}H and ^{15}N chemical shifts in F3 and F4 and can be easily identified employing a reference H/N correlation (gray). The experiment was recorded within 3.5 days of experimental time using 2 mg of perdeuterated protein. The pulse program is given in Figure 1D. See the Experimental Section for details.

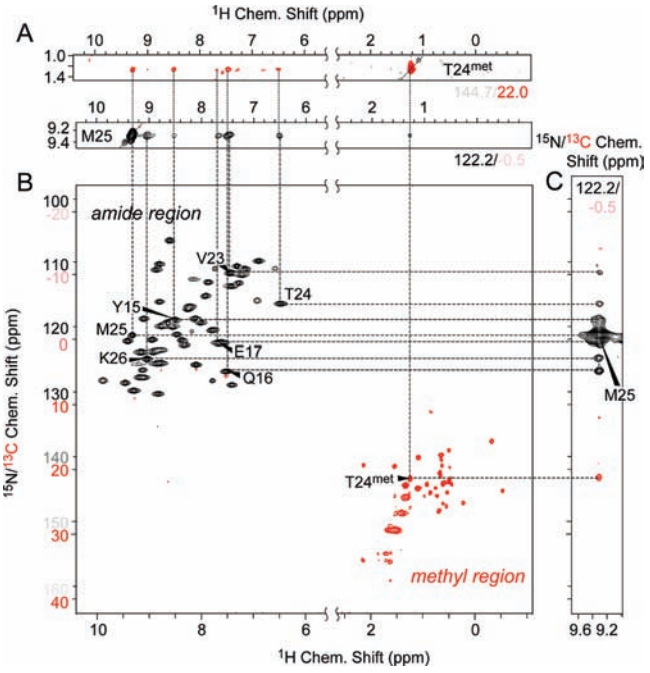


Figure 4. Complementary time-shared 3D correlation spectra with synchronous ^{13}C and ^{15}N evolution periods. (A) Strips of the H···N/C−H experiment that encodes ^{1}H chemical shift in F1 extracted at the F2, F3 frequencies of the methyl group of T24 (top) and of the amide moiety of M25, respectively (bottom). The pulse program is depicted in Figure 1F. The contact of T24−H^{met} to M25−H^{N} can be cross-validated using the F1 strip containing the M25−H^{N}/N diagonal peak (located in the upfield region in F2). (B) Reference spectrum containing the time-shared H^{N}/N and H^{met}/C^{met} correlation peaks. The pulse program is shown in Figure 1E. (C) F1 strip through the double-time shared ^{15}N/^{13}C encoded 3D (N/C···N/C−H) at the F2 and F3 frequency of the amide of M25 (pulse scheme as in Figure 1G). The combination of (C) and the bottom strip in (A) unambiguously yields the assignments for the contact between M25−H^{N} and T24−H^{met}. In both 3D experiments, an RFDR mixing time of 8 ms was employed. In all spectra, a negative sign (red contours) is implied in the pulse sequence for magnetization transfers via ^{13}C. Labels for the ^{15}N and the ^{13}C chemical shift axis are drawn in black and red, respectively.

the apparent ^{1}H T_2, which critically depends on the sample and the degree of protonation.[23]

4D Experiments. In order to minimize ambiguity, a 4D H−N···N−H RFDR experiment was implemented (see Figure 1D for the pulse program). Employing a two-step phase cycle, a uniformly sampled 4D data set with sufficient resolution can be recorded within 3.5 days with only 2 mg of protein. In the future, nonuniform sampling might allow a further decrease in the experimental time.[38−41] Figure 3 shows a 2D F3/F4 plane extracted at the F1/F2 chemical shifts of 116.9 and 8.78 ppm, corresponding to the resonance frequencies of G28. The acquisition times in the indirect dimensions t_1 (^{1}H), t_2 (^{15}N), and t_3 (^{15}N) were set to 5.0, 11.0, and 11.0 ms, respectively. The projection is superimposed with a reference HSQC, which was recorded under comparable conditions.

Time-Shared Experiments. The degree of deuteration of commercially available [^{2}D$_7$, ^{13}C$_6$] glucose is on the order of ∼97%. This allows methyl proton spectra with relatively high sensitivity and resolution to be recorded.[22] In order to include these aliphatic protons in the experiments, we implemented a time-shared ^{13}C evolution period (see Figure 1E−G), in which chemical shifts of ^{13}C and ^{15}N nuclei are allowed to evolve at the same time. Concomitant magnetization transfer in amides and methyls was achieved via a simultaneous ^{1}H−^{13}C and ^{1}H−^{15}N Hartmann−Hahn match. A time-shared evolution was implemented by scaling ^{13}C increments by 2/5 compared to ^{15}N increments to yield correct spectral ranges (in ppm) after processing. ^{13}C magnetization is prematurely flipped to the z-axis to yield a shorter t_1^{max} on carbons and thus avoids evolution of homonuclear J-couplings. The carrier frequencies on the ^{15}N and ^{13}C channels were set to the edges of the respective spectral windows (132 and 7 ppm, respectively), such that amide−methyl cross-peaks are resolved in the spectra (see Figure 4B). Magnetization pathways involving transfers through ^{13}C are phase shifted by 180° in order to differentiate between the two different types of correlations. For the concomitant CP steps, a duration of 1 ms was found to be a good compromise between sensitivity and minimization of H^{N}/C^{α} correlations, which get stronger for longer CP durations.[42] Figure 4B represents a 2D time-shared experiment without ^{1}H−^{1}H mixing. The corresponding pulse scheme is shown in Figure 1E. Figure 4A and C shows strips involving M25 extracted from the 3D experiments, which were recorded with one and two time-shared evolution periods for the ^{1}H and ^{15}N/^{13}C chemical shift encoded experiments, respectively. Pulse schemes for the H···N/C−H and N/C···N/C−H experiments are represented in Figure 1F and G. Amide−methyl cross-peak intensities are sufficiently large to employ these contacts in structure calculations. ILV-methyl group labeling together with partial amide proton back exchange in otherwise perdeuterated proteins would yield a further increase in sensitivity.[19,43] Similarly, structure calculations would benefit from methyl−methyl contacts, which are absent in our spectra due to the low degree of methyl protonation.

dx.doi.org/10.1021/ja110222h |J. Am. Chem. Soc. 2011, 133, 5905–5912

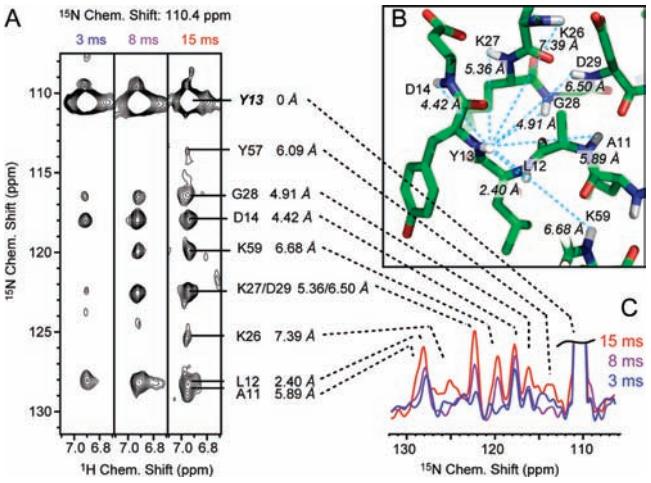


Figure 5. ^{15}N edited 3D N···N−H correlation experiment recorded with 3 different RFDR mixing times (3, 8, and 15 ms). The pulse scheme employed is shown in Figure 1A. (A) F1/F3-strips showing ^{1}H−^{1}H correlations involving the amide proton of Y13. (B) Structural model of the α-spectrin SH3 domain, highlighting the environment around residue Y13 (PDB: 2NUZ). (C) 1D columns extracted from (A). Peak heights are scaled to yield identical diagonal peak intensities. Each experiment was recorded within 4 h with $t_{1(2)}^{max}$ = 20 and 10 ms in the indirect ^{15}N evolution periods F1 and F2, respectively.

In all proton-encoded experiments, we observe strong cross-peaks to water for side chains that contain a hydroxyl group (see Figure 4, Supporting Information). This is consistent to what has been observed by Böckmann and co-workers.[44] Usually, water exchanges rapidly with the hydroxyl protons. Subsequently, magnetization is transferred via dipolar interactions to other protons, such as in this case to the methyl and amide group of T24 and M25.

Structural Restraints. Mixing of magnetization among different sites can be achieved in different ways such as by proton-driven spin diffusion (PDSD),[45] γ-encoded symmetry-based multipulse sequences,[6,46] RFDR,[7,47] etc. In this study, active recoupling of dipolar interactions by RFDR was used as spin diffusion alone is ineffective due to dilution of the proton spin system by deuteration. In order to quantitatively access ^{1}H−^{1}H dipolar couplings and the corresponding distances, cross-peak buildup rates need to be interpreted.[5,6] However, for weak dipolar interactions, an accurate quantification of the underlying distance is presumably compromised by spin diffusion. In solution-state NMR, structure calculations are usually performed using a single NOESY spectrum in which distances are grouped into three classes according to the cross-peak intensities in the initial rate regime. We analyzed cross-peak intensities semiquantitatively using 3D N···N−H experiments recorded with three RFDR mixing times (3, 8, and 15 ms). Figure 5A displays strips of the ^{15}N chemical shift encoded 3D N···N−H experiment for Y13. At even longer RFDR mixing times, signal intensities decay due to proton longitudinal relaxation and pulse imperfections. Strips in Figure 5C are therefore scaled to yield equal diagonal peak intensities. Figure 5 in the Supporting Information shows the cross-peak intensity buildup for G28 as a function of the mixing time.

Cross-peak buildup appears to be slightly slower than previously observed for crystalline proteins in which 100% of the exchangeable sites are protonated.[10] We assume that this is due to a reduction of secondary magnetization transfer pathways and spin diffusion. Distances for which cross-peaks are still observable (signal-to-noise ratio greater than 2:1 in the ^{1}H-edited experiment) are on the order of approximately 7−8 Å for τ_{mix} = 3 ms (experimental time ∼12 h) and 12−13 Å for τ_{mix} = 8 ms (experimental time ∼2 days). The experiment with τ_{mix} = 15 ms was only recorded shortly in a ^{15}N-edited fashion with an experimental time of 4 h. Due to the lower overall signal-to-noise ratio of this spectrum only distances of up to 12 Å can be observed.

Structure Calculation from Unambiguous Manual Assignments. The ARIA algorithm was originally developed for structure determination using ambiguous restraints, which involves iterative peak assignment.[4] In this section, however, ARIA was employed without the iterative peak assignment module to calculate an ensemble of structures on the basis of 294 defined distance restraints. These restraints were obtained from H···N/C−H correlations that had been assigned unambiguously by use of four-fold chemical shift encoding employing both H···N/C−H and N/C···N/C−H experiments. At the same time, backbone dihedral angle restraints predicted from secondary chemical shifts for 35 residues were employed. In the calculation, distances were restrained using generous upper bounds of 8 and 12 Å for correlations appearing in spectra recorded with 3 and 8 ms mixing times, respectively. Using the backbone torsion angle predictions, it was possible to refine the upper bounds in agreement with the allowed topology of the molecule. The final structural ensemble is well-defined, yielding a mean backbone root-mean-square deviation (RMSD) of 1.2 Å for the structured regions. The obtained model is close to the reference crystallographic structure (PDB: 2NUZ) with a RMSD of 1.4 Å (Figure 6A).[48] Almost half of all distance restraints consist of long-range restraints (Table 1), which are crucial for an accurate structure determination from sparse data. To compensate for the incompleteness of the data set and to help convergence, we used a particularly long simulated annealing protocol. In the most favored region of the Ramachandran plot, 82% of all residues are found, which is a typical value obtained for a high-resolution structure from solution-state NOESY data. The all atoms RMSD is higher, as backbone to side chain correlations (95 contacts) are only possible for Leu, Ile, and Val side chains. For all atoms (referring to heteronuclei only) we find a precision of 2.1 Å and a deviation from the X-ray structure of 2.2 Å (Table 1).

Automated Assignment and Distance Calibration from Ambiguous ^{1}H−^{1}H Contacts. To further evaluate the possibility of de novo structure determination from high-resolution ^{1}H−^{1}H correlations using one 3D experiment in which only three chemical shifts are encoded, we carried out an iterative assignment procedure using the program ARIA.[34] For this purpose, the information from the 3D ^{15}N chemical shift encoded N/C···N/C−H and the 4D H−N···N−H experiments was disregarded. A total of 540 cross-peaks was picked manually but assigned automatically, using the 3D H···N/C−H experiments recorded with RFDR mixing of 3 and 8 ms. Iterative cross-peak assignment and structure calculation made use of literature chemical shift values[24] and correspondingly derived torsion angle restraints. Initially, ARIA found an average number of six assignment possibilities per peak. With respect to the size of the chemical shift tolerance windows, this value appears reasonably low, demonstrating the high resolution of the spectra. In the next step, cross-peak intensities were used to derive quantitative upper distance restraints employing a naive calibration approach, in which we assumed correlations among

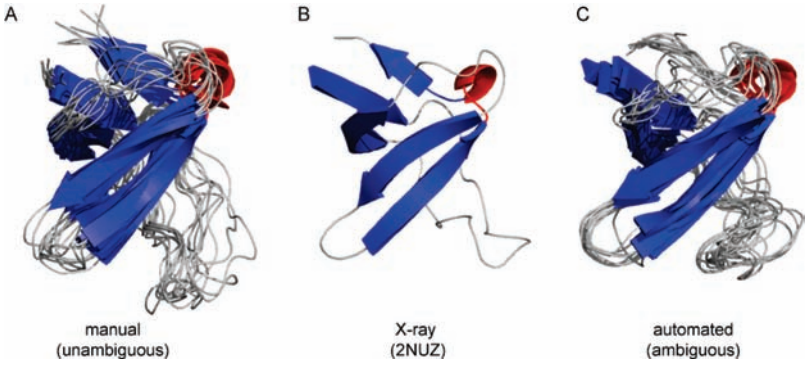


Figure 6. Structural ensembles obtained from the manual (A) and automated (C) structure calculation strategy. Structures are color-coded according to the secondary structure element. The reference X-ray structure (PDB: 2NUZ) is shown in the middle (B). Pictures were rendered using PyMOL.[50]

Table 1. Structural Restraints from the 3D H···N/C−H RFDR Experiment and Structure Determination Statistics for the Manual and the Automated ARIA Calculations of the α-Spectrin SH3 Domain[a]

	manual	automated
Distance Restraints		
total	294	413
unambiguous	294	259
intraresidue ($\|i-j\|=0$)	28	19
sequential ($\|i-j\|=1$)	58	90
medium ($2 \le \|i-j\| < 5$)	60	52
long-range ($\|i-j\| \ge 5$)	148	97
ambiguous	0	154
Ramachandran Statistics (%)		
most favored regions	82.4 ± 6.2	67.1 ± 3.8
allowed regions	13.5 ± 4.5	27.1 ± 3.5
generously allowed regions	1.6 ± 1.9	2.6 ± 2.2
disallowed regions	2.4 ± 1.6	3.0 ± 2.0
Ensemble Precision (Å)		
backbone atoms (8 −17, 23−58)	1.16 ± 0.32	1.11 ± 0.17
all atoms (8 −17, 23−58)	2.09 ± 0.45	2.02 ± 0.28
Bias to X-ray 2NUZ (Å)		
backbone atoms (8 −17, 23−58)	1.36	1.62
all atoms (8 −17, 23−58)	2.19	2.75

[a] Ramachandran statistics were determined using PROCHECK.[36] "All atoms" RMSDs refer to the nonhydrogen atoms only.

isolated spin pairs. The final conformer ensemble yielded a mean backbone RMSD of 1.1 Å, which is comparable to the precision obtained for the manually obtained set of restraints (Figure 6C). However, the bias to the X-ray structure is slightly larger (1.6 Å), and the structural quality seems reduced given the Ramachandran statistics (Table 1). The major difference in comparison to the X-ray structure is found for the RT loop (residues 18−23) (see Figure 6C and Figure 7, Supporting Information). In part, differences can be attributed to incorrect assignments of contacts which turn out to be intermolecular (the RT loop, residues 18−23, is close in space to residues 12−14 and 55−57 of a symmetry related molecule in the microcrystalline lattice). Moreover, the ARIA protocol may not be well adapted with regard to the elimination of unlikely assignment possibilities since the r^{-3} function is less discriminative in terms of distance. Indeed, ARIA discards unlikely assignments through a weighting procedure based on the distance found in the calculated structure ensemble. In addition, the r^{-3} summation used here for the determination of the effective distance of an ambiguous restraint may bias the imposed distance toward shorter values. Hence, many ambiguous assignments can still be accommodated in the employed generous bounds, thus hampering the determination of a more accurate structure. In addition, the crude approximation for the distance calibration gives rise to underestimated upper bounds that introduce distortions in the structure (see Figure 6, Supporting Information). A better calibration method that would take into account relayed transfers and correct the distance for spin diffusion could certainly improve the accuracy of the final structures.[49] As expected, the all atom RMSD (2.0 Å) and the deviation from the X-ray structure (by 2.8 Å) are high for the automated structure calculation in comparison to the backbone-only values (Table 1).

A second ARIA calculation, performed with a higher value for $\langle d_i \rangle$ in the distance calibration process, failed to converge (data not shown). Similarly to the protocol employed by Manolikas et al. for ^{13}C correlations experiments,[45] here we used only the average maximum distance (5.5 Å) deduced from the calculation based on TALOS+ restraints. We attribute this inability to converge to the fairly permissive nature of the derived upper bounds (average upper bound raised by 30%).

DISCUSSION

Although many strategies for structure determination of proteins in the solid state have been proposed in the past, the lack of spectral resolution, in particular for proton chemical shift-based approaches, resulted in ambiguities of cross-peak assignments and thus low convergence of structure calculations. This problem can be largely alleviated by the proposed approach, which yields unambiguous identification of amide−amide and amide−methyl contacts. In particular for larger proteins, this strategy will be essential, where ambiguities due to degenerate chemical shifts in one dimension are tremendous. Deuteration eliminates dipolar broadening and yields long-lived spin states and narrow resonance lines in all spectral dimensions without the need for high-power hetero- or homonuclear decoupling. Although ARIA can handle ambiguous cross-peak assignments and already yields convergence for our restricted data set, the achievable accuracy of the calculated structure is higher when the full data set containing only unambiguously assigned $^{1}H-^{1}H$

contacts is employed. Furthermore, subsequent computational efforts for evaluation of different cross-peak assignment possibilities are low. For larger proteins, the automated assignment module of ARIA is expected to be beneficial even with respect to the full data set. Although the overall fold is certainly the first and most important step toward a detailed picture of a protein, details about side chain geometries are indispensable for protein functionality. Even though the inclusion of amide-methyl distance restraints brings the all atom RMSDs close to the backbone RMSD wherever available (Figure 7, Supporting Information), the average RMSD for all atoms (excluding protons) amounts to 2.0 and 2.1 Å for automated and manual assignments, respectively. This issue therefore remains a major challenge for the applied methodology. Inclusion of further amide-aliphatic contacts might be a solution to this problem.[51] Work in this direction is currently in progress in our laboratory. Furthermore, protein structure calculations will benefit from the use of additional ^{13}C-based distance or torsion angle restraints.

The suggested approach allows long-range distances to be obtained, which have so far been very difficult to obtain in the solid state. The sparse distribution of protons prevents dipolar truncation, i.e., the masking of long-range contacts in the presence of large dipolar interactions due to closely spaced nuclei, and yields correlations among protons separated by up to 13 Å. These long-range contacts are of particular importance for obtaining a well-defined overall protein structure.

CONCLUSIONS

We have presented an approach to access unambiguous long-range distance information for protein structure calculations in MAS solid-state NMR. The experiments rely on uniformly deuterated samples in which labile deuterons are partially back substituted with protons. This labeling scheme is essential to yield favorable relaxation properties. In the experiments, contacts between amide and/or methyl groups are encoded using the chemical shifts of all four nuclei involved. This is achieved either by recording a pair of complementary 3D experiments or by acquiring a 4D data set. The experiments show magnificent spectral resolution in combination with very high sensitivity, yielding unambiguous distance restraints within a few hours. Distances of up to 13 Å are accessible, allowing protein structure calculations solely based on amide and methyl proton contacts. The approach presented provides a very potent tool for structure elucidation, particularly for larger proteins, which typically suffer from a high inherent chemical shift degeneracy.

ASSOCIATED CONTENT

Supporting Information. Sequential assignment via dipolar ^{1}H−^{1}H correlations, spectra of scalar transfer-based experiments, details of time-shared experiments, build-up curves, and correlations of distance and peak intensity. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
reif@tum.de, rlinser@unsw.edu.au

ACKNOWLEDGMENT

We are thankful to Dr. James M. Hook for his support of the project. This research was supported by the Leibniz-Gemeinschaft and the DFG (Re1435, SFB449, SFB740). R.L. is a Kekulé scholar and acknowledges financial support by the Verband der Chemischen Industrie (VCI). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 261863.

REFERENCES

(1) Hong, M.; Jakes, K. *J. Biomol. NMR* **1999**, *14*, 71–74.
(2) Castellani, F.; van Rossum, B.-J.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98–102.
(3) Lange, A.; Giller, K.; Hornig, S.; Martin-Eauclaire, M. F.; Pongs, O.; Becker, S.; Baldus, M. *Nature* **2006**, *440*, 959–962.
(4) Loquet, A.; Bardiaux, B.; Gardiennet, C.; Blanchet, C.; Baldus, M.; Nilges, M.; Malliavin, T.; Böckmann, A. *J. Am. Chem. Soc.* **2008**, *130*, 3579–3589.
(5) Reif, B.; van Rossum, B.-J.; Castellani, F.; Rehbein, K.; Diehl, A.; Oschkinat, H. *J. Am. Chem. Soc.* **2003**, *125*, 1488–1489.
(6) Reif, B.; Jaroniec, C. P.; Rienstra, C. M.; Hohwy, M.; Griffin, R. G. *J. Magn. Reson.* **2001**, *151*, 320–327.
(7) Paulson, E. K.; Morcombe, C. R.; Gaponenko, V.; Dancheck, B.; Byrd, R. A.; Zilm, K. W. *J. Am. Chem. Soc.* **2003**, *125*, 14222–14223.
(8) Reif, B.; Griffin, R. G. *J. Magn. Reson.* **2003**, *160*, 78–83.
(9) Linser, R.; Fink, U.; Reif, B. *J. Magn. Reson.* **2008**, *193*, 89–93.
(10) Zhou, D. H.; Shea, J. J.; Nieuwkoop, A. J.; Franks, W. T.; Wylie, B. J.; Mullen, C.; Sandoz, D.; Rienstra, C. M. *Angew. Chem., Int. Ed.* **2007**, *46*, 8380–8383.
(11) Fesik, S. W.; Zuiderweg, E. R. P. *J. Magn. Reson.* **1988**, *78*, 588–593.
(12) Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 1515–1517.
(13) Ikura, M.; Bax, A.; Clore, G. M.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 9020–9022.
(14) Diercks, T.; Truffault, V.; Coles, M.; Millet, O. *J. Am. Chem. Soc.* **2010**, *132*, 2138–2139.
(15) Mariani, M.; Tessari, M.; Boelens, R.; Vis, H.; Kaptein, R. *J. Magn. Reson., Ser. B* **1994**, *104*, 294–297.
(16) Farmer, B. T.; Mueller, L. *J. Biomol. NMR* **1994**, *4*, 673–687.
(17) Sattler, M.; Maurer, M.; Schleucher, J.; Griesinger, C. *J. Biomol. NMR* **1995**, *5*, 97–102.
(18) Xu, Y.; Long, D.; Yang, D. *J. Am. Chem. Soc.* **2007**, *129*, 7722–7723.
(19) Frueh, D.; Leed, A.; Arthanari, H.; Koglin, A.; Walsh, C. T.; Wagner, G. *J. Biomol. NMR* **2009**, *45*, 311–318.
(20) Chevelkov, V.; Rehbein, K.; Diel, A.; Reif, B. *Angew. Chem., Int. Ed.* **2006**, *45*, 3878–3881.
(21) Linser, R.; Fink, U.; Reif, B. *J. Biomol. NMR* **2010**, *47*, 1–6.
(22) Agarwal, V.; Reif, B. *J. Magn. Reson.* **2008**, *194*, 16–24.
(23) Akbey, Ü.; Lange, S.; Franks, W. T.; Linser, R.; Rehbein, K.; Diehl, A.; van Rossum, B.-J.; Reif, B.; Oschkinat, H. *J. Biomol. NMR* **2009**, *46*, 67–73.
(24) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. *ChemBioChem* **2001**, *2*, 272–281.
(25) Wickramasinghe, N. P.; Kotecha, M.; Samoson, A.; Paast, J.; Ishii, Y *J. Magn. Reson.* **2007**, *184*, 350–356.
(26) Linser, R.; Chevelkov, V.; Diehl, A.; Reif, B. *J. Magn. Reson.* **2007**, *189*, 209–216.
(27) Linser, R.; Fink, U.; Reif, B. *J. Am. Chem. Soc.* **2009**, *131*, 13703–13708.
(28) Wickramasinghe, N. P.; Parthasarathy, S.; Jones, C. R.; Bhardwaj, C.; Long, F.; Kotecha, M.; Mehboob, S.; Fung, L. W.-M.; Past, J.; Samoson, A.; Ishii, Y. *Nat. Methods* **2009**, *6*, 215–218.
(29) Zhou, D. H.; Rienstra, C. M. *J. Magn. Reson.* **2008**, *192*, 167–172.
(30) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, P.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins* **2005**, *59*, 687–696.

(31) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J. J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.
(32) Krushelnitsky, A.; deAzevedo, E.; Linser, R.; Reif, B.; Saalwchter, K.; Reichert, D. *J. Am. Chem. Soc.* **2009**, *131*, 12097–12099.
(33) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. *J. Biomol. NMR* **2009**, *44*, 213–23.
(34) Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, M.; Malliavin, T. E.; Nilges, M. *Bioinformatics* **2007**, *23*, 381–382.
(35) Linge, J. P.; Williams, M. A.; Spronk, C. A.; Bonvin, A. M.; Nilges, M. *Proteins* **2003**, *50*, 496–506.
(36) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(37) Linser, R.; Fink, U.; Reif, B. *J. Am. Chem. Soc.* **2010**, *132*, 8891–8893.
(38) Kupce, E.; Freeman, R. *J. Am. Chem. Soc.* **2003**, *125*, 13958–13959.
(39) Franks, T. W.; Atreya, H. S.; Szyperski, T.; Rienstra, C. M. *J. Biomol. NMR* **2010**, *48*, 213–223.
(40) Hyberts, S. G.; Takeuchi, K.; Wagner, G. *J. Am. Chem. Soc.* **2010**, *132*, 2145–2147.
(41) Matsuki, Y.; Eddy, M. T.; Griffin, R. G.; Herzfeld, J. *Angew. Chem., Int. Ed.* **2010**, *122*, 9401–9404.
(42) Agarwal, V.; Linser, R.; Fink, U.; Faelber, K.; Reif, B. *J. Am. Chem. Soc.* **2010**, *132*, 3187–3195.
(43) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868–13878.
(44) Lesage, A.; Gardiennet, C.; Loquet, A.; Verel, R.; Pintuacuda, G.; Emsley, L.; Meier, B. H.; Böckmann, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 5851–5854.
(45) Manolikas, T.; Herrmann, T.; Meier, B. H. *J. Am. Chem. Soc.* **2008**, *130*, 3959–3966.
(46) Nielsen, N. C.; Bildsoe, H.; Jakobsen, H. J.; Levitt, M. H. *J. Chem. Phys.* **1994**, *101*, 1805–1812.
(47) Bennett, A. E.; Ok, J. H.; Vega, S.; Griffin, R. G. *J. Chem. Phys.* **1992**, *96*, 8624–8627.
(48) Chevelkov, V.; Faelber, K.; Schrey, A.; Rehbein, K.; Diehl, A.; Reif, B. *J. Am. Chem. Soc.* **2007**, *129*, 10195–10200.
(49) Linge, J. P.; Habeck, M.; Rieping, W.; Nilges, M. *J. Magn. Reson.* **2004**, *167*, 334–342.
(50) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific San Carlos, CA, 2002.
(51) Asami, S.; Schmieder, P.; Reif, B. *J. Am. Chem. Soc.* **2010**, *132*, 15133–15135.

■ NOTE ADDED IN PROOF

A proton-detected 4D solid-state NMR experiment for Protein Structure Determination has been proposed simultaneously by Huber et al. 2011: Huber, M.; Hiller, S.; Schanda, P.; Ernst, M.; Bockmann, A.; Verel, R.; Meier. B. H. A proton-detected 4D solid-state NMR experiment for protein structure determination. *ChemPhysChem* **2011**, DOI: 10.1002/cphc.201100062.