Determination of Solid-State NMR Structures of Proteins by Means of Three-Dimensional $^{15}N-^{13}C-^{13}C$ Dipolar Correlation Spectroscopy and Chemical Shift Analysis†

Federica Castellani,‡ Barth-Jan van Rossum,*,‡,§ Annette Diehl,‡ Kristina Rehbein,‡ and Hartmut Oschkinat*,‡,§

Forschungsinstitut für Molekulare Pharmakologie, Robert-Rössle-Strasse 10, 13125 Berlin, Germany, and Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany

Received May 28, 2003; Revised Manuscript Received August 4, 2003

ABSTRACT: In this paper, a three-dimensional (3D) NMR-based approach for the determination of the fold of moderately sized proteins by solid-state magic-angle spinning (MAS) NMR is presented and applied to the α-spectrin SH3 domain. This methodology includes the measurement of multiple $^{13}C-^{13}C$ distance restraints on biosynthetically site-directed ^{13}C-enriched samples, obtained by growing bacteria on [2-^{13}C]glycerol and [1,3-^{13}C]glycerol. 3D $^{15}N-^{13}C-^{13}C$ dipolar correlation experiments were applied to resolve overlap of signals, in particular in the region where backbone carbon−carbon correlations of the $C^{\alpha}-C^{\alpha}$, CO−CO, C^{α}−CO, and CO−C^{α} type appear. Additional restraints for confining the structure were obtained from ϕ and ψ backbone torsion angles of 29 residues derived from C^{α}, C^{β}, CO, NH, and H^{α} chemical shifts. Using both distance and angular restraints, a refined structure was calculated with a backbone root-mean-square deviation of 0.7 Å with respect to the average structure.

Many biological systems, such as membrane proteins and amyloid fibrils, remain a challenge in structural biology because of difficulties with crystallization and solubility. In the past years, solid-state NMR[1] has become a promising method for obtaining structural information about these systems, via the measurement of accurate distances (*1*−*7*), ϕ and φ backbone torsion angles (*8*−*10*), and chemical shift anisotropy (*11*, *12*). In these studies, samples labeled only in the positions of interest were investigated. For the determination of the complete protein folds, however, a different approach that allows the collection of a large number of structural restraints from a small number of samples has to be followed. The quality of the structures increases with the number of restraints, and the more that are measured, the lower the accuracy of the individual restraints may be. From a close analysis of the topology of helical and β-sheet structures, it transpires that carbon−carbon distances are very important in defining the fold of a protein. For example, distances between backbone carbons, i.e., α-carbons and carbonyl carbons, define the secondary structure of a protein (Table 1). Distances between backbone and side chain carbons or between side chain and side chain carbons provide information about the tertiary structure. The detection of structure-defining long-range carbon−carbon restraints is only possible when so-called dipolar truncation effects are suppressed (*13*, *14*). This can be accomplished by employing a reduced labeling scheme, in which chemically bonded carbons are not simultaneously labeled and hence the number of strong dipolar couplings between connected nuclei is reduced. For proteins expressed in bacterial systems, this can be achieved by using [2-^{13}C]glycerol or [1,3-^{13}C]glycerol as the only carbon source in the media (*15*−*17*). In combination with this labeling pattern, long-range $^{13}C-^{13}C$ distance restraints may be collected by using a broad-band recoupling method like the proton-driven spin-diffusion (PDSD) mixing scheme (*18*). This methodology has recently been proven to be successful on the α-spectrin SH3 domain (*19*), where two-dimensional (2D) spectroscopy was used to collect inter-residue $^{13}C-^{13}C$ restraints, yet to allow the determination of the fold of the protein.

Investigations on larger proteins or determination of structures with higher resolution would require us to resolve as many backbone−backbone or backbone−side chain cross-peaks as possible. As a general approach, we introduce here a set of three-dimensional (3D) $^{15}N-^{13}C-^{13}C$ experiments of the NCACX and NCOCX type (where CX stands for any carbon atom). In the 3D spectra recorded with these techniques, the improved resolution obtained by adding a ^{15}N dimension allowed the identification of backbone carbon−carbon restraints of the C^{α}−CO, CO−C^{α}, $C^{\alpha}-C^{\alpha}$, and CO−CO type and restraints involving side chain carbons, which were not accessible in the 2D experiments. For the SH3 domain, 374 new inter-residue correlations could be assigned from the 3D data, in addition to the earlier study (*19*). Additional information was obtained from ^{1}H, ^{13}C, and

† The work was supported in part by the DFG (Grant SFB 449).

* To whom correspondence should be addressed. E-mail: oschkinat@fmp-berlin.de and brossum@fmp-berlin.de. Phone: (+49)30-94793161. Fax: (+49)30-94793169.

‡ Forschungsinstitut für Molekulare Pharmakologie.

§ Freie Universität Berlin.

[1] Abbreviations: 3D, three-dimensional; NMR, nuclear magnetic resonance; MAS, magic-angle spinning; SH3, Src-homology 3; PDSD, proton-driven spin-diffusion; CP, cross-polarization; rf, radio frequency; TPPM, two-pulse phase modulation; FID, free induction decay; HSQC, heteronuclear single-quantum coherence; rmsd, root-mean-square deviation.

10.1021/bi034903r CCC: $25.00
Published on Web 09/12/2003

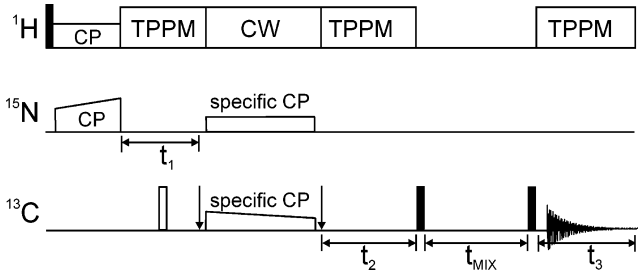

FIGURE 1: Pulse sequence for the 3D NCACX and NCOCX experiments. The 3D $^{15}N-^{13}C-^{13}C$ heteronuclear dipolar correlation experiment starts with $^{1}H-^{15}N$ CP transfer followed by a ^{15}N evolution period. The ^{15}N magnetization is then transferred to the C^{α} or CO carbons via a specific-CP step, achieved by applying weak r.f. fields, on resonance with the ^{15}N amides in the ^{15}N channel and off-resonance in the ^{13}C channel. The vertical arrows indicate a switch in the frequency offset, required for the specific-CP. After the carbon evolution period, the carbon magnetization is exchanged using a PDSD mixing unit. During evolution and acquisition periods, a TPPM decoupling is applied, while during specific-CP, continuous wave (CW) decoupling is used.

^{15}N chemical shift analysis (*20*), yielding 58 backbone dihedral angles. The 374 newly obtained distance restraints and the 58 angular restraints were used to refine the solid-state NMR structure of the α-spectrin SH3 domain.

We emphasize that the periodicity of the microcrystalline SH3 material is not a requirement for the presented methodology, but helps in providing a structurally homogeneous environment. For membrane proteins, structural or conformational homogeneity is supported by either embedding the protein into the membrane or embedding an agonist (or antagonist) into the receptor binding site which serves as a matrix and forces the ligand into a single conformation ensuring sufficiently narrow lines. Line widths of 30−40 Hz for singly ^{13}C-labeled sites in bacterial photosynthetic reaction centers and around 80−100 Hz, including *J* couplings, in ^{13}C spectra of bovine rhodopsin were achieved at moderate MAS frequencies (*21*, *22*). Likewise, sufficiently narrow lines were observed in 2D carbon−carbon correlation spectra recorded from multiply labeled ligands or cofactors bound to membrane proteins (*22*−*25*), as necessary for the assignment of side chain resonances or the measurement of distances. The size of many ligands triggering G-protein-coupled receptors ranges from 4 to 40 residues, which is well within the range for chemical shift assignment and structure determination by solid-state MAS NMR (*19*, *26*−*30*).

MATERIALS AND METHODS

Sample Preparation and Solid-State NMR Spectroscopy. The SH3 protein was expressed in *Escherichia coli* BL21(DE3), using M9 minimal media. The [[2-^{13}C]-glycerol,^{15}N]SH3 and [[1,3-^{13}C]glycerol,^{15}N]SH3 samples were prepared as described previously (*19*). For the MAS measurements, approximately 10−12 mg of protein was used and confined to the center of 4 mm rotors by using spacers. The 3D $^{15}N-^{13}C-^{13}C$ heteronuclear dipolar correlation

Table 1: Summary of the Most Characteristic Structure-Defining $^{13}C-^{13}C$ Distances in Proteins[a]

		structure-defining $^{13}C-^{13}C$ distance in proteins					
secondary structure	residue	$C^{\alpha}-C^{\alpha}$ (Å)	CO−CO (Å)	C^{α}−CO (Å)	CO−C^{α} (Å)	$C^{\alpha}-C^{\beta}$ (Å)	CO−C^{β} (Å)
α-helix	$i, i+1$	3.8 ± 0.1	3.0 ± 0.1	4.4 ± 0.1	2.4 ± 0.1	4.9 ± 0.1	3.7 ± 0.1
	$i, i+2$	5.5 ± 0.3	4.5 ± 0.1	5.4 ± 0.2	4.4 ± 0.2	6.0 ± 0.5	5.2 ± 0.3
	$i, i+3$	5.0 ± 0.2	4.7 ± 0.1	5.5 ± 0.2	4.4 ± 0.2	4.4 ± 0.5	4.3 ± 0.4
	$i, i+4$	6.0 ± 0.2	5.9 ± 0.2	7.2 ± 0.2	4.9 ± 0.2	5.7 ± 0.5	4.5 ± 0.3
3_{10}-helix	$i, i+1$	3.8 ± 0.1	3.1 ± 0.2	4.5 ± 0.2	2.4 ± 0.1	4.9 ± 0.1	3.7 ± 0.1
	$i, i+2$	5.3 ± 0.3	4.6 ± 0.2	5.6 ± 0.2	4.3 ± 0.3	5.6 ± 0.5	5.0 ± 0.5
	$i, i+3$	5.6 ± 0.4	5.5 ± 0.3	6.6 ± 0.5	4.7 ± 0.3	5.0 ± 0.5	4.5 ± 0.5
	$i, i+4$	8.1 ± 0.4	7.9 ± 0.4	9.2 ± 0.5	6.8 ± 0.4	8.3 ± 0.5	6.9 ± 0.5
antiparallel β-sheet	i, j	4.6 ± 0.3		5.4 ± 0.5		5.4 ± 0.7	
	$i-1, j+1$	5.4 ± 0.3	5.0 ± 0.4				
	$i-1, j$		4.9 ± 0.3		4.7 ± 0.5		5.4 ± 0.7
parallel β-sheet	i, j	4.9 ± 0.5	4.9 ± 0.3	4.6 ± 0.6	5.6 ± 0.3	4.6 ± 0.6	5.6 ± 0.7

[a] The four most commonly encountered secondary structure motifs are considered. For each secondary structure element, a set of structures was selected and distances between backbone and backbone and between backbone and side chain carbons were measured. The errors are estimated from the dispersion of the distances in the selected structures. The indices of the residues refer to the schematic representation of secondary structures shown in Figure 2.

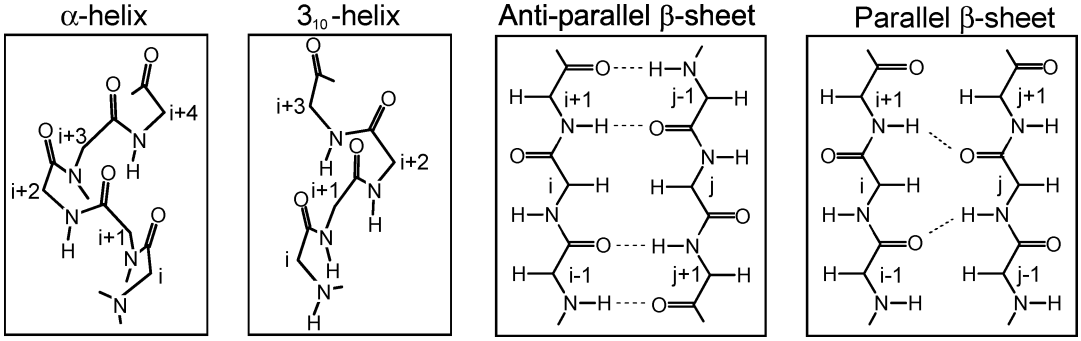

FIGURE 2: Schematic representation of α-helix, 3_{10}-helix, antiparallel β-sheet, and parallel β-sheet secondary structure motives. The numbering of the residues (*i* and *j*) corresponds to the numbering in Table 1.

experiments were performed on a DMX-400 spectrometer (Bruker, Karlsruhe, Germany) operating at a field of 9.4 T and equipped with a 4 mm triple-resonance CP/MAS probe (Bruker). The data sets were recorded at 280 K and at a MAS frequency $\omega_R/2\pi$ of 8.0 kHz. On the [[2-^{13}C]glycerol,^{15}N]-SH3 sample, an experiment of the NCACX type was performed, where the magnetization is transferred first from nitrogen to the C^α atoms and subsequently to the other carbons. In the 3D experiment carried out with the [[1,3-^{13}C]glycerol,^{15}N]SH3, the magnetization transfers from ^{15}N to the carbonyl atoms and then to the other carbons. For both experiments, the same pulse sequence has been applied, depicted in Figure 1. Following ^{1}H excitation, a ramped cross-polarization contact of 500 μs between ^{1}H and ^{15}N created the initial ^{15}N magnetization; spin lock fields were 33 kHz for ^{1}H and 19−26 kHz for the ^{15}N ramp. In the center of the first ^{15}N evolution period, a π pulse on the carbon channel was applied to refocus heteronuclear ^{15}N−^{13}C scalar couplings. Following the evolution of nitrogen, specific-CP (*31*) was employed to selectively transfer magnetization from ^{15}N to $^{13}C^\alpha$, in the case of the [[2-^{13}C]glycerol,^{15}N]SH3 sample, and from ^{15}N to ^{13}CO, in the case of the [[1,3-^{13}C]-glycerol,^{15}N]SH3 sample. In both experiments, weak rf powers corresponding to nutation frequencies of ∼10 kHz for the ^{15}N and ∼20 kHz for the ^{13}C were used and the amide ^{15}N signals were irradiated close to resonance while the $^{13}C^\alpha$ and ^{13}CO signals off-resonance. The ^{13}C frequency offset was varied to match the appropriate NCA or NCO specific-CP condition. After the specific-CP, the ^{13}C carrier frequency was placed back in the aliphatic region for the NCACX experiment, and in the carbonyl region for the NCOCX experiment. After the ^{13}C evolution period, the carbon magnetization was exchanged by using a PDSD mixing scheme, with a mixing time of 500 ms (*18*). During all evolution periods, proton decoupling was applied, using the two-pulse phase modulation technique (TPPM) (*32*), with a pulse width of 7.3 μs and a phase modulation of 10°. The 3D data sets were acquired using 1794 × 96 × 96 points and dwell times of 10, 60, and 120 μs for f_3, f_2, and f_1, respectively; each FID was averaged from 32 scans, using a recycle delay of 2.6 s, yielding a total measurement time of ∼10 days for each 3D experiment.

The data were processed with the XWINNMR software, version 2.6 (Bruker), and subsequently analyzed using the program Sparky, version 3.100 (T. D. Goddard and D. G. Kneller, University of California, San Francisco, CA).

Structure Calculation. Structures were calculated with the program CNS (*33*), version 1.0. Calculations were performed using the simulated annealing protocol with torsion-angle dynamics, starting with 200 randomized conformers. As described previously, the ^{13}C−^{13}C restraints obtained from 2D spectroscopy were classified as strong (2.5−4.5 Å), medium (2.5−5.5 Å), weak (2.5−6.5 Å), or very weak (2.5−7.5 Å). The six ^{15}N−^{15}N correlations were restrained to 3−6 Å. These data are supplemented with the 374 inter-residue ^{13}C−^{13}C restraints obtained from the 3D spectra, categorized as very weak. Prediction of torsion angles was carried out using the TALOS software (*20*). In total, 58 predictions for the backbone torsion angles were obtained and included in the restraint list for the CNS calculation. The error margins for the angular restraints were set to a minimum value of 20°. The 10 lowest-energy structures calculated using both distance and angular restraints showed no distance violations greater than 0.3 Å.

RESULTS AND DISCUSSION

Structure-Defining $^{13}C-^{13}C$ Distances. For structure determination, a large number of distances in the range of 2−7 Å are needed. Analysis of the fold of α-helical and β-sheet proteins shows that many of these structure-defining distances can potentially be extracted from carbon−carbon correlation spectra of the NCACX and NCOCX type (see Table 1 and Figure 2). Table 1 gives average values of distances plus variation derived from a set of structures, which contain the different types of secondary structure elements, as depicted in Figure 2.

In α-helices, distances between backbone carbons of residues i and $i + n$, where $n = 1$, 2, 3, or 4, are the shortest and most characteristic in defining this type of secondary structure. Magnetization transfer between C^α_i and $CO_{i+2,3}$ covers distances on the order of 5.5 ± 0.2 Å and should lead to small but detectable cross-peaks in carbon−carbon correlation spectra. Distances of the C^α_i−CO_{i+4} type are around 7 Å and should not always be observable. The CO_i−$C^\alpha_{i+2,3,4}$ distances are in the range of 4.3−4.9 Å and hence valuable for defining this secondary structure. In the set of distances between C^α, the C^α_i−C^α_{i+3} distance is always the shortest and together with CO_i−$CO_{i+2,3}$ distances most indicative of the α-helix. The distances between the backbone carbons i and $C^\beta_{i+2,3}$ are also short enough to give detectable cross-peaks.

The characteristic feature of a 3_{10}-helix is that this secondary structure motif is more elongated than an α-helix. For the 3_{10}-helix, distances between backbone carbons of residues i and $i + 4$ are in general too long to provide detectable cross-peaks. Stronger cross-peaks are expected for correlations between CO_i−CO_{i+2} and CO_i−$C^\alpha_{i+2,3}$ signals, due to transfer over distances in the range of 4.3−4.7 Å. Weaker but still detectable peaks may occur because of correlations of the C^α−CO type of residues i, $i + 2$, and $i + 3$. Also important for defining the 3_{10}-helix are correlations between the backbone carbons i and side chains $C^\beta_{i+2,3}$ atoms.

In contrast to those of helices, β-sheet topologies are mostly defined by contacts between backbone carbons located in different strands (Figure 2). In the case of an antiparallel β-sheet, we denote residues of different strands that have the C^α closer together with i and j, and residues which have the C^α further apart with $i + 1$ and $j - 1$ (Figure 2). The C^α_i−C^α_j distances are relatively short (4.6 ± 0.3 Å), while those between C^α_{i+1} and C^α_{j-1} are on the order of 5.4 ± 0.3 Å. The shortest distances between carbonyl carbons are found for residues i and $j + 1$, and $i - 1$ and j, and are on the order of 5.0 ± 0.4 Å. Distances of the C^α_i−CO_{j+1} type are ∼5.4 ± 0.5 Å and those of the CO_{i-1}−C^α_j type are ∼4.7 ± 0.5 Å. The latter are very valuable for defining antiparallel β-sheet secondary structure. Cross-peaks may also be observed between backbone carbons of one strand and the β-carbons of a neighboring strand. The shortest distances of this type are on the order of 5.4 ± 0.7 Å. In the case of a parallel β-sheet, we define i and j as the closest residues located in the two parallel strands. All the distances between backbone carbons are in range of 4.6−5.6 Å and should give rise to detectable signals. In the same range are the distances between the backbone carbons and β-carbons.

Labeling Strategy: [2-^{13}C]Glycerol-Made SH3 and [1,3-^{13}C]Glycerol-Made SH3 Domains. To be able to detect a large number of ^{13}C−^{13}C long-range correlations for structure calculations, we developed a methodology that combines broad-band recoupling with a reduced level of ^{13}C labeling (*19*). Dilution of ^{13}C spins is achieved by growing the protein sample with [2-^{13}C]glycerol or [1,3-^{13}C]glycerol as the sole carbon source (*15*−*17*, *19*). Two different and complementary labeling schemes are obtained, where the resulting ^{13}C enrichment is both extensive and reduced. We refer to the two samples as 2-SH3 and 1,3-SH3 (*19*). The reduced level of labeling allows to a large extent the suppression of dipolar truncation effects that dominate recoupling experiments in uniformly labeled preparations, where the strong dipolar couplings between connected nuclei suppress the weaker ones. The latter comprise couplings between carbon spins close in space but not chemically bonded and contain the relevant structure information. A first estimate of the location of ^{13}C labels in the different residues of our protein preparations was obtained following their biosynthetic pathway, starting from [2-^{13}C]glycerol and [1,3-^{13}C]glycerol as carbon sources. Via their labeling schemes, the amino acids can be divided into two groups, reflecting the two major steps in the biosynthetic pathway. The first group contains amino acids that are synthesized from precursors made in the glycolysis pathway. This labeling scheme yields either 100 or 0% labeled sites. In the 2-SH3 domain, for example, glycines and alanines have the α-carbons fully labeled and the carbonyl positions unlabeled and, for alanines, no labeling at the β-position; the opposite labeling is obtained for the 1,3-SH3 preparation. Only in the case of valines and leucines, two chemically connected carbons (the α- and β-carbons for valine and the β- and γ-carbons for leucine), are both fully enriched in 1,3-SH3 and both unlabeled in 2-SH3. The second group contains amino acids produced from precursors made in the citric acid cycle, which exhibit a more complex labeling scheme. The strain used for the expression of the SH3 molecules (see Materials and Methods) contained all the enzymes that are involved in the citric acid cycle. As a consequence, the cycle can be followed several times, resulting in the production of a mixture of differently labeled molecules. On average, this results in amino acids with partially labeled nuclei. In Figure 3a, the possible isotopomers produced during the citric acid cycle in the case of isoleucine and threonine are shown, together with the average labeling. For 2-SH3, the labeled carbons are in the positions indicated in red, while for 1,3-SH3, the complementary labeling is obtained (represented in green). It is not possible to quantify the relative amounts of each isotopomer produced in the cycle; hence, the averaged labeling was estimated from analysis of solution NMR spectra. This was done by analyzing the topology and intensity of each peak in high-resolution ^{1}H−^{13}C HSQC spectra, recorded on 1,3-SH3 and 2-SH3 samples dissolved in aqueous solution.

An important and advantageous side effect of the alternating ^{13}C labeling is that the carbon lines in ^{13}C CP/MAS NMR spectra of 2-SH3 and 1,3-SH3 are significantly narrower than the carbon lines in spectra recorded on uniformly ^{13}C enriched samples. The reduced level of labeling partly suppresses or completely removes line broadening due to ^{13}C−^{13}C scalar couplings. These effects are illustrated in Figure 3, in which one-dimensional (1D) ^{13}C CP/MAS

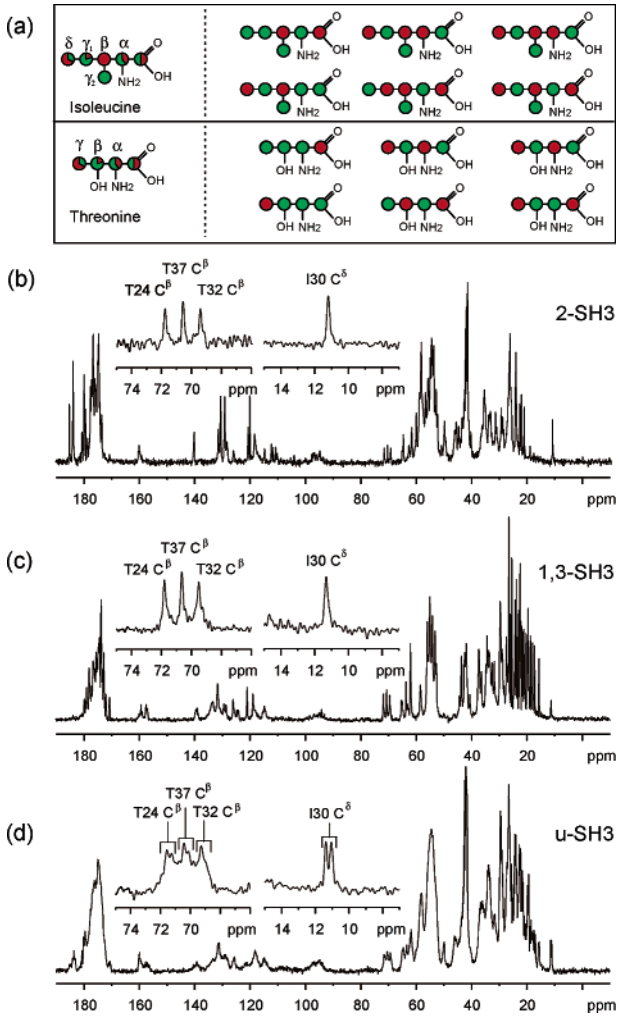

FIGURE 3: Extensive and alternating ^{13}C enrichment and the effect on the line width. (a) Schematic representation of the effective ^{13}C labeling for isoleucine and threonine, obtained by growth on [2-^{13}C]glycerol or [1,3-^{13}C]glycerol. For the 2-SH3 sample, the labeled sites are represented in red. For 1,3-SH3, the complementary labeling is obtained, depicted in green. On the left side, the average labeling is depicted, estimated from solution NMR data. On the right, the possible differently labeled isotopomers produced during the citric acid cycle are shown, derived following the reactions of the biosynthetic pathway. (b−d) 1D ^{13}C CP-MAS spectra of 2-SH3 (b), 1,3-SH3 (c), and [U-^{13}C]SH3 (d), recorded at 9.4 T and using a spinning frequency of 8.0 kHz. The spectra are recorded with a long acquisition time of 45 ms and 64 scans. In the expansions of the C$^{\beta}$ signals of threonines and the C$^{\delta}$ signal of I30, the line widths obtained with the two different ^{13}C labeling schemes can be compared. Worth noting in panel d is the splitting of the I30 C$^{\delta}$ signal of [U-^{13}C]SH3, due to ^{13}C−^{13}C scalar couplings.

spectra of 2-SH3 (Figure 3b), 1,3-SH3 (Figure 3c), and [U-^{13}C]SH3 (Figure 3d) are compared. The insets show the well-resolved C$^{\delta}$ signal of I30 at 11.3 ppm and the C$^{\beta}$ signals from three threonines around 70 ppm. In the 2-SH3 sample, these signals have line widths of 20−25 Hz (Figure 3b). The remarkable narrowing of these resonances in comparison to the fully enriched SH3 arises from the fact that in none of the isotopomers of the isoleucine are C$^{\delta}$ and C$^{\gamma 1}$ labeled at the same time, which completely removes the one-bond ^{13}C−^{13}C scalar coupling broadening. The same is observed for C$^{\beta}$ of threonine, which never is connected to a labeled neighbor. In the case of 1,3-SH3, a situation is observed where in some of the isotopomers the C$^{\delta}$ atom of I30 and

the C^{β} atoms of threonines are connected with labeled carbons. In the 1D spectrum, the line widths of these signals are 30−35 Hz (Figure 3c). Substantially broader lines are observed in the [U-^{13}C]SH3 sample (Figure 3d); for a relatively short acquisition time of ∼25 ms, the signals of I30 C^{δ} and threonine C^{β} have line widths of 60−70 Hz (data not shown). For longer acquisition times (e.g., 45 ms), the signal of I30 C^{δ} can be resolved into a doublet, because of the scalar coupling with the neighboring $C^{\gamma 1}$ (cf. Figure 3d). The signals of the threonine C^{β} present a more structured feature due to multiple couplings to both α- and γ-carbons (Figure 3d).

3D NCACX and NCOCX Experiments for Obtaining Complementary Restraints on 2- and 1,3-Glycerol-Labeled Samples. The broad-band recoupling method chosen for detection of inter-residue correlations is the proton-driven spin-diffusion technique (*18*). The advantages of this technique are its robustness due to its simplicity and aptness to be implemented for very long mixing times. Typically, mixing times of 500 ms for ^{13}C−^{13}C spectra and of several seconds for ^{15}N−^{15}N spectra can be used, without limitations on balancing radio frequency power input and sample heating. From a series of 2D spectra recorded with different mixing times, ∼500 ^{13}C−^{13}C and ^{15}N−^{15}N inter-residue correlations were obtained. The first structure calculation of the SH3 domain exhibited a backbone rmsd of 1.8 Å, with respect to the average structure (*19*).

A more general approach for structure determination of moderately sized proteins by solid-state MAS NMR can be envisaged using 3D ^{15}N−^{13}C−^{13}C dipolar correlation experiments as a basis (Figure 1). The 3D data allow extraction of structural information that cannot be easily derived from 2D data only. Two different 3D experiments, termed NCACX and NCOCX, can be performed on the two biosynthetically site-directed ^{13}C-enriched samples, to resolve C^{α}−C^{α}, CO−CO, C^{α}−CO, and CO−C^{α} correlations, which are important for defining the secondary structure (see also Table 1). Because of the biosynthetically derived labeling pattern, labels are present in either the C^{α} or CO position of most of the amino acids, while only few residues will have labels in both positions. Any of the four possible backbone carbon−carbon correlation cross-peaks involving C^{α} and CO needs therefore to be considered as a source for structure-defining restraints. Since the position of the C^{β} with respect to the backbone atoms of the same amino acid is only moderately conformationally dependent, distances of the C^{α}−C^{β} and CO−C^{β} type need to be considered as secondary structure-defining distances as well, also for compensating for the eventual scarcity of restraints between backbone carbons due to the lack of labels in appropriate places.

Restraints between C^{α}−C^{α} and CO−CO backbone carbons occur twice in NCACX and NCOCX spectra; hence, they may be resolved by two different nitrogen chemical shifts and assigned with less ambiguity. Interactions which occur between a backbone carbon and side chain carbons (e.g., between a C^{α} atom located in one helix and a methyl group located in another helix) occur only once in the set of spectra and can only be resolved by means of one nitrogen chemical shift.

As shown in Figure 1, the proposed 3D experiments consist of an indirect ^{15}N evolution period followed by specific-CP (*31*), to selectively transfer magnetization from

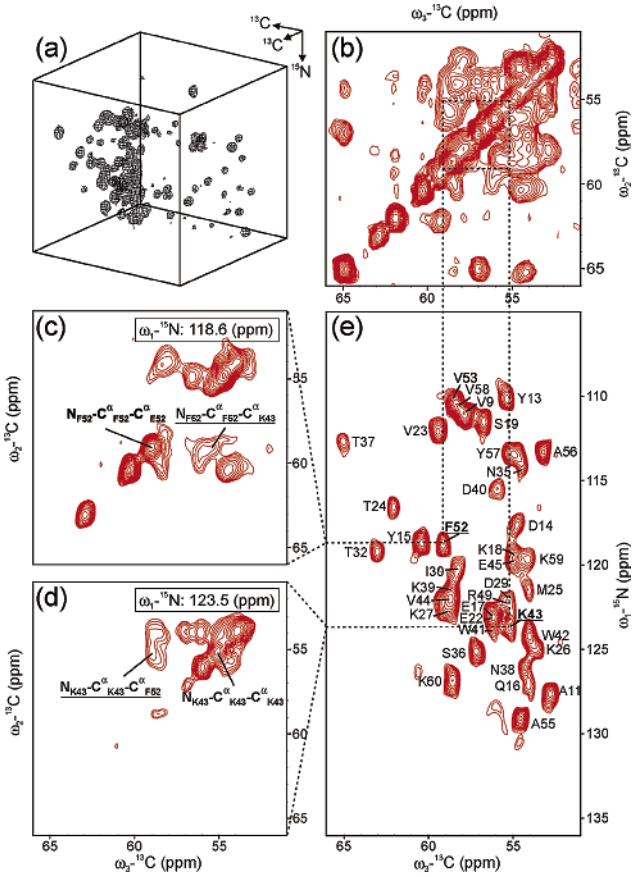

FIGURE 4: Assignment strategy for the 3D NCACX experiment. In panel a, a contour plot with a single contour of the 3D NCACX spectrum is shown. The region displayed in the plot is the aliphatic region. The spectrum was recorded for a 2-SH3 sample at 9.4 T and using a spinning frequency of 8.0 kHz. In panel b, the α-region of the 3D ^{13}C−^{13}C projection is displayed. The dotted lines indicate the positions of C^{α}_{K43}−C^{α}_{F52} and C^{α}_{F52}−C^{α}_{K43} correlations, not resolved in 2D spectra. The two residues have different ^{15}N chemical shifts, as shown in the 2D NCA spectrum in panel e, and in the 3D spectrum, the two long-range correlations are resolved in two different ^{15}N planes, displayed in panels c and d.

^{15}N to $^{13}C^{\alpha}$ or from ^{15}N to ^{13}CO. In a second evolution period, the ^{13}C magnetization evolves under the $^{13}C^{\alpha}$ or ^{13}CO isotropic chemical shifts and subsequently is exchanged to surrounding carbons via a PDSD mixing unit (*18*). To detect inter-residue correlations, a long mixing time of 500 ms is used.

In total, four different 3D experiments can be performed, combining the NCOCX and NCACX sequences with the two glycerol-made samples. However, in 2-SH3, most of the C^{α} atoms of hydrophobic and aromatic residues are labeled and the carbonyl carbons are unlabeled, while the opposite situation is observed for 1,3-SH3. As a consequence, two of the four combinations, the NCACX experiment carried out with 2-SH3 and the NCOCX experiment carried out with 1,3-SH3, will most effectively provide the largest number of ^{13}C−^{13}C structural restraints.

NCACX Experiment on 2-SH3. In Figure 4a, a contour plot of the 3D NCACX correlation experiment carried out with 2-SH3 is shown. Many signals are observed in the α-region (50−70 ppm), and in particular for each C^{α}-labeled residue, the $^{15}N_i$−$^{13}C^{\alpha}_i$−$^{13}C^{\alpha}_i$ intraresidue correlation is observed as well as signals of the $^{15}N_i$−$^{13}C^{\alpha}_i$−$^{13}C^{\alpha}_j$ type, which are due to magnetization exchange between different

C^{α} atoms. In addition, signals are observed in the β- and γ-region (0−50 ppm), because of magnetization exchange between C^{α} and labeled side chain carbons, and in the carbonyl and aromatic region (100−180 ppm, not shown), because of transfer to CO and aromatic side chains. As an example, the assignment of the long-range correlation between K43 and F52 is illustrated in Figure 4. The two long-range peaks ($C^{\alpha}_{K43}-C^{\alpha}_{F52}$ and $C^{\alpha}_{F52}-C^{\alpha}_{K43}$), not resolved in the 2D spectrum (Figure 4b), are separated in two different ^{15}N planes of the 3D NCACX spectrum. One correlation appears in the ^{15}N plane of K43 at 123.5 ppm, together with the intraresidue correlation peak of K43 (Figure 4d), and the other in the ^{15}N plane of F52 at 118.6 ppm, together with the intraresidue correlation peak of F52 (Figure 4c). Additionally, in the ^{15}N plane of K43, correlations with the aromatic C^{γ} and C^{ϵ} atoms of F52 are assigned, while in the ^{15}N plane of F52, a correlation with C^{γ}_{K43} can be observed (data not shown). Furthermore, correlations between residues V44−G51 and W42−V53 could be identified from the 3D spectrum, defining, together with the K43−F52 correlation, one of the antiparallel β-sheets of the SH3 domain. In total, 168 new inter-residue correlations could be assigned from the NCACX spectrum. From these correlations, 74 are sequential, 31 are medium-range, and 63 are long-range. Since the 3D data were recorded with a single $^{13}C-^{13}C$ PDSD mixing time, the cross-peak intensities could not be readily translated into a distance class (*19*); hence, we classified these correlations in the class with distances in the range of 2.5−7.5 Å (*19*).

NCOCX Experiment with 1,3-SH3. The NCOCX experiment carried out with 1,3-SH3 is shown in Figure 5a. The contour plot comprises both carbonyl and aliphatic regions. In the carbonyl region, peaks of the $^{15}N_{i+1}-^{13}CO_i-^{13}CO_i$ type are detected for all residues except leucines, which do not have a labeled CO in 1,3-SH3. In the same region, correlations of the $^{15}N_{i+1}-^{13}CO_i-^{13}CO_j$ type also appear, because of magnetization exchange between carbonyl atoms of different residues. The assignment of inter-residue correlations was done by searching for pairs of peaks ($^{15}N_{i+1}-^{13}CO_i-^{13}CO_j$ and $^{15}N_{j+1}-^{13}CO_j-^{13}CO_i$) appearing in the two ^{15}N planes of residues $i+1$ and $j+1$. As an example, the assignment of the long-range correlation between K27 and Y13 is illustrated in Figure 5. In the ^{15}N plane at 116.7 ppm, corresponding to the nitrogen chemical shift of G28, the $N_{G28}-CO_{K27}-CO_{K27}$ diagonal peak appears, together with the $N_{G28}-CO_{K27}-CO_{K26}$ sequential peak and the $N_{G28}-CO_{K27}-CO_{Y13}$ long-range correlation (Figure 5c). At the same nitrogen chemical shift, T24 also resonates; hence, other peaks appear in Figure 5c as well and are labeled with the respective assignments. In the ^{15}N plane of D14 at a unique chemical shift of 117.7 ppm, the $N_{D14}-CO_{Y13}-CO_{Y13}$ diagonal peak appears, together with a sequential and two long-range correlations, $CO_{Y13}-CO_{K27}$ and $CO_{Y13}-CO_{K26}$ (Figure 5d). In the aliphatic region of the spectrum, several correlations are observed. In particular in the C^{α} region, correlations of the $^{15}N_{i+1}-^{13}CO_i-^{13}C^{\alpha}_j$ type are detected. For example, in three different ^{15}N planes of the 3D experiment, peaks could be assigned to $N_{L31}-CO_{I30}-C^{\alpha}_{L10}$, $N_{L10}-CO_{V9}-C^{\alpha}_{L31}$, and $N_{L33}-CO_{T32}-C^{\alpha}_{L8}$ correlations, defining one of the antiparallel β-sheets of the SH3. Since in 1,3-SH3 the leucines have the carbonyl carbon unlabeled and the valines the C^{α} unlabeled, the $N_{A11}-CO_{L10}-C^{\alpha}_{I30}$, $N_{T32}-CO_{L31}-C^{\alpha}_{V9}$, and $N_{L33}-CO_{T32}-C^{\alpha}_{L8}$ long-range peaks, also expected because of the proximity of the two strands of the β-sheet, cannot be observed. In total, 206 new inter-residue correlations could be assigned from the 3D NCOCX data set. From these correlations, 69 are sequential, 24 are medium-range, and 115 are long-range. These correlations were grouped in the class with distances in the range of 2.5−7.5 Å (*19*).

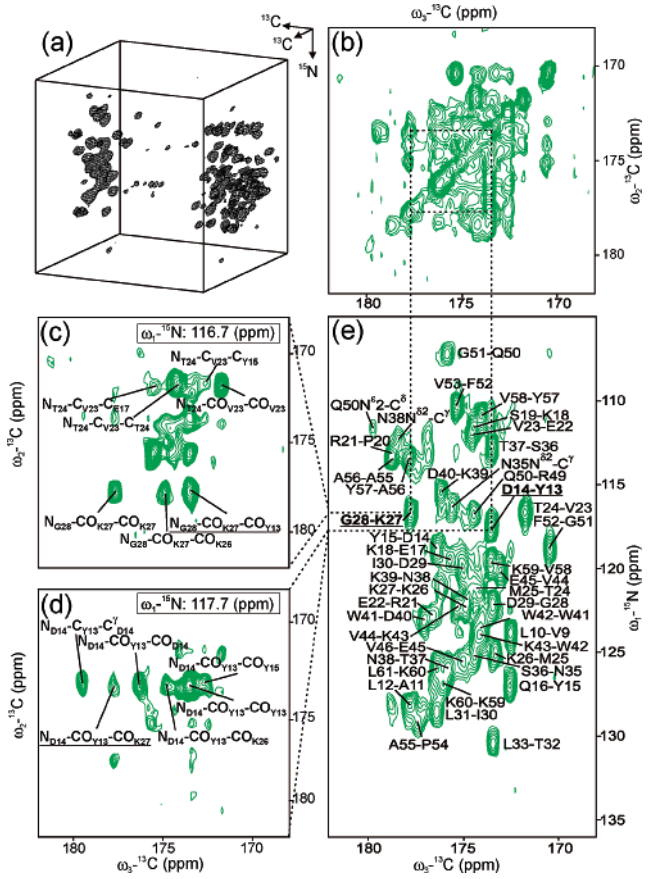

FIGURE 5: Assignment strategy for the 3D NCOCX experiment. In panel a, a contour plot with a single contour of the 3D NCOCX spectrum is shown. The plot includes the carbonyl and aliphatic regions of the spectrum. The spectrum was recorded on a 1,3-SH3 sample at 9.4 T and using a spinning frequency of 8.0 kHz. In panel b, the carbonyl region of the $^{13}C-^{13}C$ projection is displayed. The dotted lines indicate the positions of $CO_{Y13}-CO_{K27}$ and $CO_{K27}-CO_{Y13}$ correlations, not resolved in 2D spectroscopy. Because of the different ^{15}N chemical shifts of residues G28 and D14 of the 2D NCO spectrum in panel e, the long-range correlation between K27 and Y13 could be identified from the 3D NCOCX experiment. In panels c and d, the planes of the NCOCX experiment at 116.7 and 117.7 ppm are displayed, with the assignment of the $N_{G28}-CO_{K27}-CO_{Y13}$ and $N_{D14}-CO_{Y13}-CO_{K27}$ peaks, defining this long-range correlation.

Angular Restraints Obtained from Chemical Shift Analysis. The signals observed in a MAS solid-state NMR spectrum resonate at the isotropic chemical shifts, similar to solution NMR. This has two important consequences. First, chemical shift databases such as the BioMagResBank, compiled from solution NMR data, can be assessed for the resonance assignment in solid-state NMR. This has been applied for the side chain assignment of amino acids for various proteins (*26*, *27*, *29*, *30*). Second, the secondary chemical shift, i.e., the difference between the measured isotropic chemical shift and the corresponding random coil value, contains information about secondary structure motifs (*5*, *20*, *34*, *35*). Taking full advantage of the almost complete solid-state MAS NMR

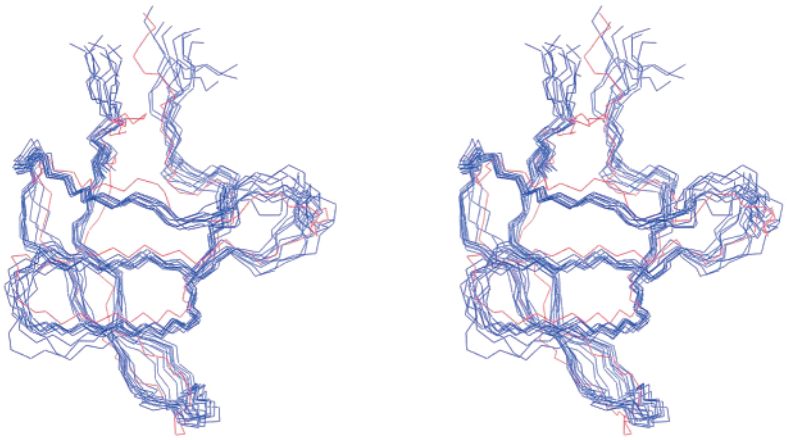

FIGURE 6: Solid-state structure of the α-spectrin SH3 domain. Stereoview of the 10 lowest-energy structures, representing the fold of the α-spectrin SH3 domain, colored blue. For comparison, the X-ray structure is included (*38*), displayed in red, and overlaid with the family of 10 solid-state structures by fitting the backbone C^{α} atoms to the average solid-state structure.

^{1}H, ^{13}C, and ^{15}N resonance assignments for the SH3 domain (*30*, *36*, *37*), we can identify additional restraints for improvement of the structure using the program TALOS (*20*) to predict backbone torsion angles, from the C^{α}, C^{β}, CO, NH, and H^{α} chemical shifts of the SH3 domain. For 29 residues, ϕ and φ backbone torsion-angle predictions with sufficient reliability were made and used in the structure calculation, together with the 889 distance restraints.

Structure Calculation. We have performed two structure calculations for the α-spectrin SH3 domain. For a first calculation, only distance restraints were used. A list of 889 inter-residue restraints was generated for residues 7−61 of the SH3 domain, obtained from 2D (*19*) and 3D spectroscopy. A conventional structure calculation protocol with simulated annealing and torsion-angle dynamics was applied, and the 10 lowest-energy structures were selected. The C^{α} coordinates of the regular structure elements exhibited an rmsd of 1.1 Å with respect to the average structure and of 1.6 Å to the X-ray structure (*38*). By addition of the 58 dihedral-angle restraints from TALOS (*20*) in a second structure calculation, the rmsd of the C^{α} coordinates of the regular structure elements with respect to the average structure is reduced to 0.7 Å. This structure is also closer to the X-ray structure, with a deviation of 1.2 Å. In Figure 6, the 10 lowest-energy structures selected to represent the fold of the α-spectrin SH3 domain are shown and superimposed on the X-ray structure (*38*). The MAS NMR structure is in general well defined. The most refined part is the β-sheet, while a lower degree of convergence is observed in the loop regions. A difference is found for residue E7. The first six residues of the protein are not detected in our MAS spectra, because of the flexibility of the N-terminus, and are not displayed in the figure.

Using 3D NMR spectroscopy and chemical shift analysis, a refined structure of the α-spectrin SH3 domain was obtained, exclusively derived from solid-state MAS NMR data. The fold of the SH3 domain consists of seven strands that form two orthogonal antiparallel β-sheets. A large number of backbone carbon restraints defining the β-sheet structure could be extracted from $^{15}N-^{13}C-^{13}C$ dipolar correlation experiments, due to the enhanced resolution achieved by adding a ^{15}N dimension to the experiment. In particular, cross-peaks of the $C^{\alpha}-C^{\alpha}$, CO−CO, C^{α}−CO, and CO−C^{α} type could be resolved, and it was not possible to extract them from 2D spectra only, because of the limited C^{α} and CO dispersion. We expect that the same methodology can be applied to probe other secondary structure motives such as parallel β-sheet or helices, which are characterized by structure-defining $^{13}C-^{13}C$ distances in the measurable range up to ∼7 Å.

ACKNOWLEDGMENT

We thank Bernd Reif, Peter Schmieder, and Ronald Kühne for helpful discussions.

REFERENCES

1. Balbach, J. J., Petkova, A. T., Oyler, N. A., Antzutkin, O. N., Gordon, D. J., Meredith, S. C., and Tycko, R. (2002) *Biophys. J. 83*, 1205−1216.
2. Isaac, B., Gallagher, G. J., Balazs, Y. S., and Thompson, L. K. (2002) *Biochemistry 41*, 3025−3036.
3. Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W., Groesbeek, M., Metz, G., Ziliox, M., and Aimoto, S. (2002) *Biophys. J. 82*, 2476−2486.
4. Nishimura, K., Kim, S., Zhang, L., and Cross, T. A. (2002) *Biochemistry 41*, 13170−13177.
5. Jaroniec, C. P., MacPhee, C. E., Astrof, N. S., Dobson, C. M., and Griffin, R. G. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 16748−16753.
6. Verdegem, P. J. E., Helmle, M., Lugtenburg, J., and deGroot, H. J. M. (1997) *J. Am. Chem. Soc. 119*, 169−174.
7. Petkova, A. T., Hatanaka, M., Jaroniec, C. P., Hu, J. G. G., Belenky, M., Verhoeven, M., Lugtenburg, J., Griffin, R. G., and Herzfeld, J. (2002) *Biochemistry 41*, 2429−2437.
8. van Beek, J. D., Hess, S., Vollrath, F., and Meier, B. H. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 10266−10271.
9. Blanco, F. J., Hess, S., Pannell, L. K., Rizzo, N. W., and Tycko, R. (2001) *J. Mol. Biol. 313*, 845−859.
10. Feng, X., Verdegem, P. J. E., Eden, M., Sandstrom, D., Lee, Y. K., Bovee-Geurts, P. H. M., de Grip, W. J., Lugtenburg, J., de Groot, H. J. M., and Levitt, M. H. (2000) *J. Biomol. NMR 16*, 1−8.
11. Hong, M., McMillan, R. A., and Conticello, V. P. (2002) *J. Biomol. NMR 22*, 175−179.
12. Yao, X., Yamaguchi, S., and Hong, M. (2002) *J. Biomol. NMR 24*, 51−62.
13. Hodgkinson, P., and Emsley, L. (1999) *J. Magn. Reson. 139*, 46−59.
14. Kiihne, S., Mehta, M. A., Stringer, J. A., Gregory, D. M., Shiels, J. C., and Drobny, G. P. (1998) *J. Phys. Chem. A 102*, 2274−2282.

15. LeMaster, D. M., and Kushlan, D. M. (1996) *J. Am. Chem. Soc. 118*, 9255−9264.
16. Hong, M. (1999) *J. Magn. Reson. 139*, 389−401.
17. Hong, M., and Jakes, K. (1999) *J. Biomol. NMR 14*, 71−74.
18. Szeverenyi, N. M., Sullivan, M. J., and Maciel, G. E. (1982) *J. Magn. Reson. 47*, 462−475.
19. Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) *Nature 420*, 98−102.
20. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
21. Shochat, S., Gast, P., Hoff, A. J., Boender, G. J., Vanleeuwen, S., Vanliemt, W. B. S., Vijgenboom, E., Raap, J., Lugtenburg, J., and deGroot, H. J. M. (1995) *Spectrochim. Acta, Part A 51*, 135−144.
22. Verhoeven, M. A., Creemers, A. F. L., Bovee-Geurts, P. H. M., de Grip, W. J., Lugtenburg, J., and de Groot, H. J. M. (2001) *Biochemistry 40*, 3282−3288.
23. EgorovaZachernyuk, T. A., vanRossum, B., Boender, G. J., Franken, E., Ashurst, J., Raap, J., Gast, P., Hoff, A. J., Oschkinat, H., and deGroot, H. J. M. (1997) *Biochemistry 36*, 7513−7519.
24. Griffiths, J. M., Lakshmi, K. V., Bennett, A. E., Raap, J., Vanderwielen, C. M., Lugtenburg, J., Herzfeld, J., and Griffin, R. G. (1994) *J. Am. Chem. Soc. 116*, 10178−10181.
25. Creemers, A. F. L., Kiihne, S., Bovee-Geurts, P. H. M., deGrip, W. J., Lugtenburg, J., and de Groot, H. J. M. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 9101−9106.
26. Hong, M. (1999) *J. Biomol. NMR 15*, 1−14.
27. McDermott, A., Polenova, T., Bockmann, A., Zilm, K. W., Paulson, E. K., Martin, R. W., Montelione, G. T., and Paulsen, E. K. (2000) *J. Biomol. NMR 16*, 209−219.
28. Pauli, J., van Rossum, B., Forster, H., de Groot, H. J. M., and Oschkinat, H. (2000) *J. Magn. Reson. 143*, 411−416.
29. Egorova-Zachernyuk, T. A., Hollander, J., Fraser, N., Gast, P., Hoff, A. J., Cogdell, R., de Groot, H. J., and Baldus, M. (2001) *J. Biomol. NMR 19*, 243−253.
30. Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001) *ChemBioChem 2*, 272−281.
31. Baldus, M., Petkova, A. T., Herzfeld, J., and Griffin, R. G. (1998) *Mol. Phys. 95*, 1197−1207.
32. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995) *J. Chem. Phys. 103*, 6951−6958.
33. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
34. Laws, D. D., Bitter, H. M., Liu, K., Ball, H. L., Kaneko, K., Wille, H., Cohen, F. E., Prusiner, S. B., Pines, A., and Wemmer, D. E. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 11686−11690.
35. Luca, S., Filippov, D. V., van Boom, J. H., Oschkinat, H., de Groot, H. J., and Baldus, M. (2001) *J. Biomol. NMR 20*, 325−331.
36. van Rossum, B. J., Castellani, F., Rehbein, K., Pauli, J., and Oschkinat, H. (2001) *ChemBioChem 2*, 906−914.
37. van Rossum, B. J., Castellani, F., Pauli, J., Rehbein, K., Hollander, J., de Groot, H. J., and Oschkinat, H. (2003) *J. Biomol. NMR 25*, 217−223.
38. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992) *Nature 359*, 851−855.

BI034903R