JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

^{13}C Spin Dilution for Simplified and Complete Solid-State NMR Resonance Assignment of Insoluble Biological Assemblies

Antoine Loquet, Guohua Lv, Karin Giller, Stefan Becker, and Adam Lange*

Department of NMR-Based Structural Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Supporting Information

ABSTRACT: A strategy for simplified and complete resonance assignment of insoluble and noncrystalline proteins by solid-state NMR (ssNMR) spectroscopy is presented. Proteins produced with [1-^{13}C]- or [2-^{13}C]glucose are very sparsely labeled, and the resulting 2D ssNMR spectra exhibit smaller line widths (by a factor of ∼2 relative to uniformly labeled proteins) and contain a reduced number of cross-peaks. This allows for an accelerated and straightforward resonance assignment without the necessity of time-consuming 3D spectroscopy or sophisticated pulse sequences. The strategy aims at complete backbone and side-chain resonance assignments based on bidirectional sequential walks. The approach was successfully demonstrated with the de novo assignment of the Type Three Secretion System PrgI needle protein. Using a limited set of simple 2D experiments, we report a 97% complete resonance assignment of the backbone and side-chain ^{13}C atoms.

Recent developments in magic-angle-spinning (MAS) solid-state NMR (ssNMR) methodology,[1−4] isotope labeling schemes,[5,6] structure calculation protocols,[7,8] and access to high-field instrumentation have allowed details at atomic resolution and, in the most favorable cases, high-resolution structures of microcrystalline proteins,[6−11] fibrillar aggregates,[12−17] oligomeric complexes,[18−22] and membrane proteins in nativelike environments[23−26] to be obtained. However, achieving a sufficient (and if possible, nearly complete) assignment of the NMR signals still remains as a major obstacle in obtaining site-specific structural information. The lack of resolution and the spectral overlap observed in uniformly ([U-^{13}C]glucose) labeled proteins requires the use of several 2D experiments[27] and, in the case of remaining assignment ambiguities, 3D or 4D spectroscopy. This renders the assignment step extremely time-consuming and demanding in terms of instrument performance and access to high-field spectrometers. An alternative to uniform labeling is the use of ^{13}C alternate labeling schemes,[5,28,29] which reduce spectral crowding and facilitate the assignment and the collection of distance restraints, as previously demonstrated[5,6] with glycerol-based labeling schemes.

We recently reported[30] the use of mixtures of [1-^{13}C]-[5] and [2-^{13}C]glucose (Glc)[29] for ssNMR studies of supramolecular protein interfaces. The high spectral quality observed for [1-^{13}C]- and [2-^{13}C]Glc-labeled α-synuclein already suggested that this labeling scheme could be useful for resonance assignments. Here we demonstrate that [1-^{13}C]- and [2-^{13}C]Glc-labeled proteins indeed exhibit excellent spectroscopic features that substantially simplify and accelerate the assignment step. As an illustration, we present the complete de novo assignment of a noncrystalline, insoluble biological assembly, the *Salmonella typhimurium* Type Three Secretion System (TTSS) needle. The TTSS (Figure 1a) is a complex nanomachine that allows pathogenic bacteria to deliver virulence effectors to host cells[31] through a hollow needle.[32,33] TTSS needles (Figure 1b) are formed by the assembly of multiple copies of PrgI.

Figure 1(c−e) shows the 1D cross-polarization (CP) spectrum of PrgI needles produced from a medium containing [2-^{13}C]Glc (for [1-^{13}C]Glc, see Figure S1 in the Supporting Information (SI)). Relative to [U-^{13}C]Glc-labeled PrgI, [2-^{13}C]Glc labeling leads to an improvement in spectral resolution due to the removal of one-bond ^{13}C−^{13}C dipolar and *J* couplings. For the isolated Ile71-Cγ2 peak (marked by an asterisk in Figure 1c,e), we measured a ^{13}C line width of 28 Hz (0.13 ppm; full-width at half height). This is comparable to the resolution observed for the HETs(218−289) prion[14] and nanocrystalline GB1,[10] which to date are the benchmarks in terms of ssNMR spectral quality. To evaluate the gain in spectral resolution systematically, we recorded 2D NCα spectra of both [2-^{13}C]Glc- and [U-^{13}C]Glc-labeled PrgI needles. As illustrated with six different residue types (Figure 1f), the improvement in ^{13}Cα resolution was ∼1.8 fold. For CO, similar values of ∼2.1 fold were measured in 2D NCO spectra (data not shown). In the ^{15}N dimension, no significant resolution enhancement between the two labeling schemes was observed. The line width for Cα was reduced by ∼45 Hz (Table S2). This corresponds roughly to the removal of one-bond *J* couplings. More significant improvements were observed for the aromatic region from spectra of [1-^{13}C]Glc- and [2-^{13}C]Glc-labeled PrgI (see Figures S2 and S6).

[1-^{13}C]- and [2-^{13}C]Glc are carbon sources that are isotopically ^{13}C-labeled at only one carbon position. This results in proteins in which only one out of six carbons are ^{13}C-labeled. Notably, the labeled positions are not random. Instead, similar to proteins produced from selectively labeled glycerol, an alternating "checkerboard" labeling scheme is achieved.

However, the Glc-based labeling is much more sparse than that obtained using glycerol as the sole carbon source, where one ([2-^{13}C]glycerol) or two ([1,3-^{13}C]glycerol) out of three carbons are labeled. The PDSD spectrum of [1-^{13}C]Glc-labeled

Received: January 4, 2011
Published: March 14, 2011

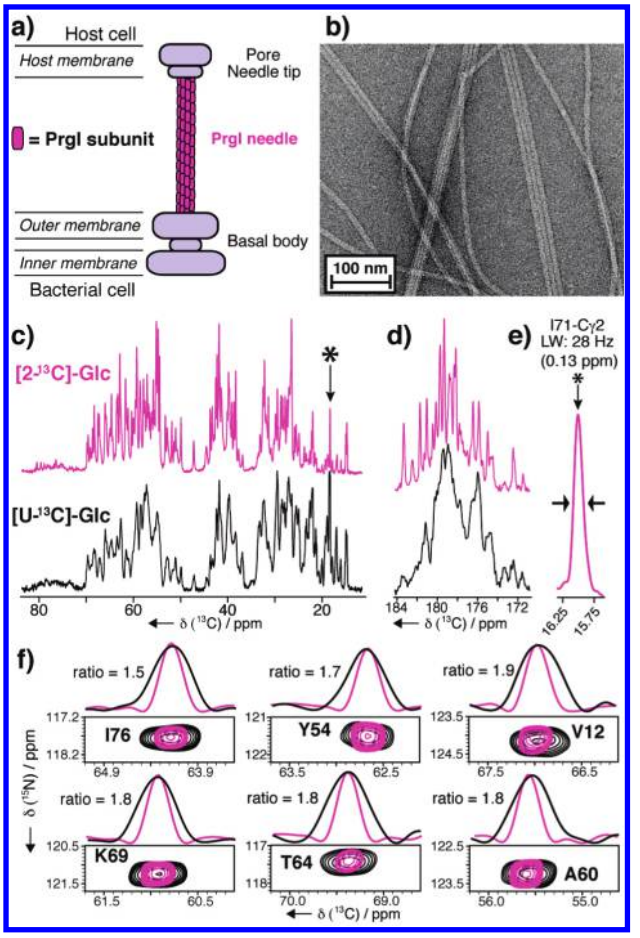

Figure 1. (a) Schematic representation of the TTSS. (b) Electron microscopy image of TTSS PrgI needles. (c–e) Excerpts of 1D CP spectra of (magenta) [2-^{13}C]Glc- and (black) [U-^{13}C]Glc-labeled needles: (c) aliphatic; (d) carbonyl region; (e) Ile71-Cγ2. No apodization functions were used. (f) Excerpts of 2D NCα spectra (full spectra in Figure S3) of the (magenta) [2-^{13}C]Glc- and (black) [U-^{13}C]Glc-labeled needles. 1D traces of the ^{13}C F_2 dimension are shown. The two spectra were recorded under the same conditions, and line widths were extracted after a sine bell apodization (90° shift) and zero-filling to 2048 (indirect) and 4096 (direct dimension) points.

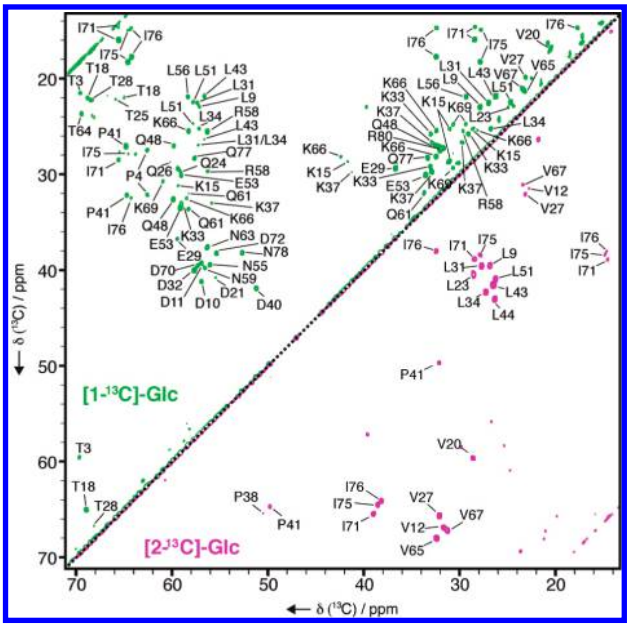

Figure 2. 2D PDSD spectra of (green) [1-^{13}C]Glc- and (magenta) [2-^{13}C]Glc-labeled PrgI needles recorded with mixing times of 100 and 50 ms, respectively. For the sake of clarity, each spectrum is shown on one side of the diagonal only. Full spectra in large format are shown in Figures S4 and S5.

PrgI needles (Figure 2) allows for the rapid identification of N, D, Q, E, R, K, I, and T. Similarly, I, P, V, L, and all of the aromatic residues (Figure S6) can rapidly be assigned in the [2-^{13}C]Glc-labeled spectrum. The two labeling schemes are quasi-complementary, as illustrated by the absence of cross-peaks detectable in both spectra. The labeled positions correspond to the labeling pattern reported by Kay and co-workers[29] on the basis of solution NMR spectra but may also depend on the details of protein expression (duration of expression, *Escherichia coli* strain, and expressed protein).

On the basis of the reduced number of detectable correlations and the observed improvement in ^{13}C resolution, we propose a strategy to sequentially assign all of the amino acid residues in a bidirectional way (i.e., from residue i to $i \pm 1$). For [2-^{13}C]Glc-labeled proteins, the polarization is mainly stored on the Cα's (no Cβ's are labeled except for V and I), as indicated by a strong diagonal in the 50 ms PDSD spectrum. With an increase in the PDSD mixing time, the polarization is then efficiently transferred to neighboring Cα's (from residue i to $i \pm 1$). Figure 3 shows the PDSD spectra of [2-^{13}C]Glc-labeled PrgI needles recorded with longer mixing times of 200 and 400 ms. The Cα−Cα region (Figure 3c), which is empty at short mixing times (Figure 2), is now composed of numerous sequential Cα−Cα correlations. Increasing the mixing time to 400 ms enables the collection of suprasequential information (labeled in blue) such as i to $i \pm 2,3$. Suprasequential contacts confirm the resonance assignment with higher fidelity by correlating i to $i \pm 2$. In addition, they provide precise site-specific structural information such as the i to $i \pm 3$ contacts, which are characteristic of an α-helical conformation. Following a similar analysis, sequential side-chain contacts and suprasequential distances were also detected in other regions of the spectra (e.g., see Figure S9 for the stretch E29−L43).

Complications may arise in the assignment of spectra from proteins with β-sheet secondary structure, where long-range Cα−Cα correlations in addition to sequential Cα−Cα correlations could be observed. In some instances, the resulting ambiguity may be resolved on the basis of residue-specific Cα chemical shifts. In other cases, additional experiments can be performed. For example, in NHHC spectra[34] of [2-^{13}C]Glc-labeled samples recorded with a short ^{1}H−^{1}H mixing time, sequential NC correlations [with a corresponding HN(i) to Hα($i-1$) distance of ∼2.1 Å in a β-sheet] are expected to be slightly stronger than intraresidue correlations and much stronger than long-range β-sheet correlations. In general, our approach works well for proteins with a mostly α-helical secondary structure, such as PrgI, or for amyloid fibrils with a parallel, in-register arrangement, such as the ones formed by human α-synuclein.[30] For other proteins, the data on [1-^{13}C]- and [2-^{13}C]Glc-labeled samples is ideally complemented by NCαCX/NCOCX-type experiments on uniformly labeled samples.

The high quality of the PrgI spectra obtained here allows for a precise and accurate detection of the resonance frequencies (±0.1 ppm) that facilitates an unambiguous sequential walk. At

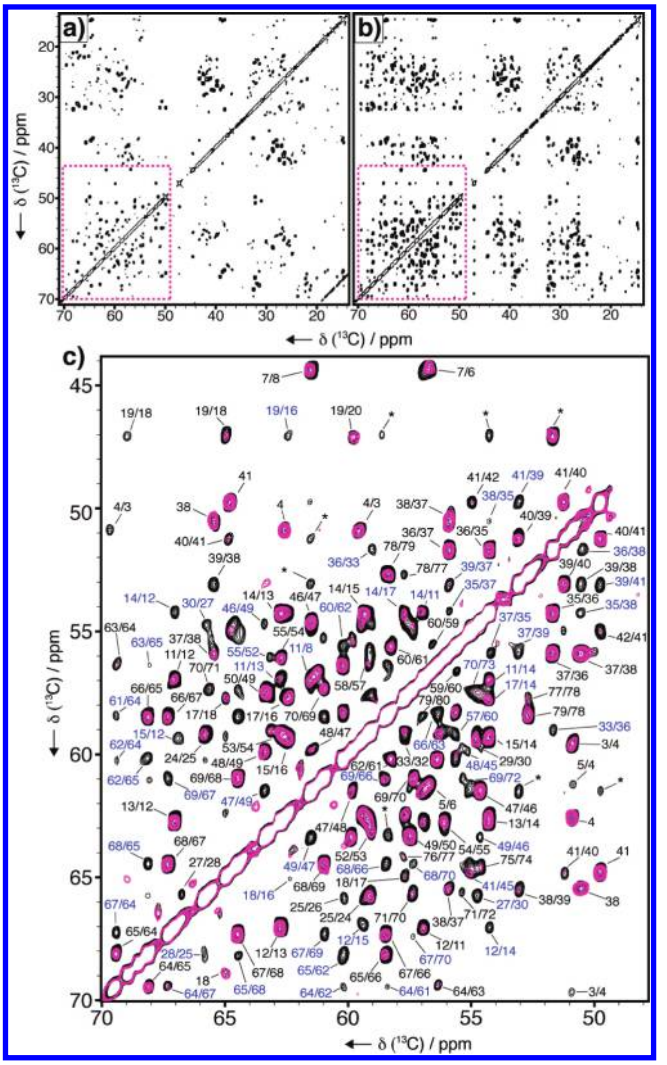

Figure 3. (a, b) 2D PDSD spectra of [2-^{13}C]Glc-labeled PrgI needles recorded with mixing times of (a) 200 and (b) 400 ms. (c) Cα−Cα region (magenta, 200 ms; black, 400 ms). Suprasequential distances are labeled in blue. Cross-peaks encoding for long-range distances are marked by asterisks.

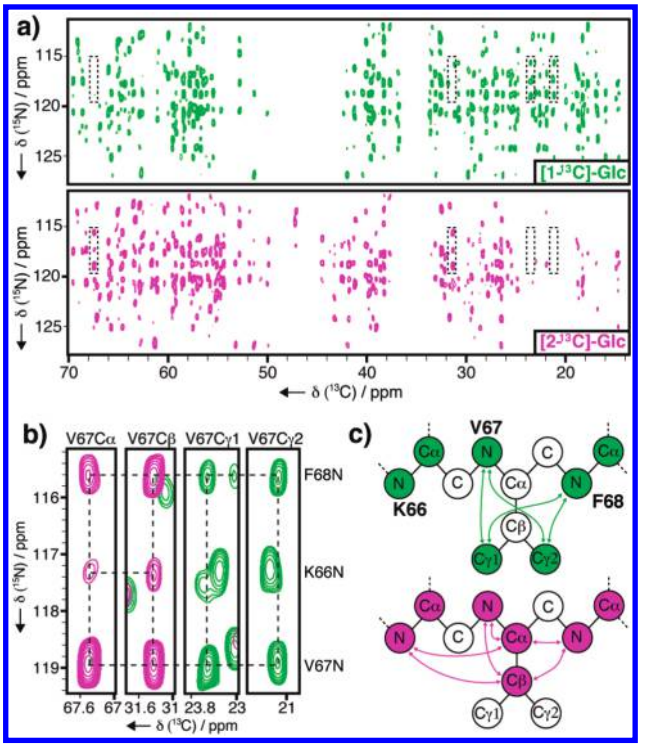

Figure 4. (a) 2D ^{15}N−^{13}C PAIN-CP spectra of (green) [1-^{13}C]Glc- and (magenta) [2-^{13}C]Glc-labeled PrgI needles. (b) Excerpts showing intraresidue and sequential correlations. (c) These connections are schematized for both labeling patterns. Notably, the direction of the transfer can be obtained by a comparison of peak intensities, where the N(i) to Cα($i-1$) correlations are stronger than the N(i) to Cα($i+1$) correlations.

this stage of the assignment process, a small set of 2D PDSD and NCα spectra of [1-^{13}C]Glc-, [2-^{13}C]Glc-, and [U-^{13}C]Glc-labeled PrgI needles was used (see the SI for details). The available data allowed for a sequential resonance assignment of 97% of the ^{13}C backbone and side-chain atoms as well as 79 out of 80 backbone ^{15}N atoms (Table S1).

A drawback of our approach is that three samples have to be produced. Furthermore, the maximum labeling level of 50% leads to an overall sensitivity drop in comparison with uniformly labeled samples or selectively glycerol-labeled samples (which are less sparse). Nevertheless, the reduction of intraresidue magnetization transfer due to the sparse labeling compensates for the loss in sensitivity, facilitating the observation of interresidue correlations. In the case of PrgI, which has a very high degree of local order, the decreased line width also partially compensates for the loss in sensitivity.

After this demonstration for TTSS PrgI needles, we anticipate the application of [1-^{13}C]- and [2-^{13}C]Glc-labeling to larger systems by making use of the ^{15}N dimension during the sequential walk. In this regard, we found that the TSAR mechanism[35] constitutes an excellent choice to transfer the polarization from ^{15}N (residue i) to neighboring ^{13}C atoms in a bidirectional way (i.e., to residues $i-1$ and $i+1$). 2D PAIN-CP spectra (Figure 4a) can be acquired in a relatively short experimental time (1−2 days), leading to highly resolved ^{15}N−^{13}C spectra. The low sensitivity of the TSAR mechanism to the dipolar truncation phenomenon[36] combined with the sparse labeling allows for efficient polarization transfer simultaneously to intraresidue and sequential ^{13}C atoms. The vast majority of the sequential contacts can already be observed at short mixing times (5 ms). The polarization pathways in the PAIN-CP spectra of Glc-labeled samples reassemble information traditionally obtained by means of $N_iC\alpha_iCX_i$ and $N_iCO_{i-1}CX_{i-1}$ experiments and additionally provide $N_iC\alpha_{i+1}CX_{i+1}$ contacts (Figure 4b,c and Figure S11) with better sensitivity in the one-step transfer in the PAIN-CP experiments than in the two-step transfer (SPECIFIC CP[37] followed by ^{13}C−^{13}C recoupling) required for the above-cited experiments.

In conclusion, we have reported substantial improvements in terms of increased spectral resolution and simplicity for [1-^{13}C]Glc- and [2-^{13}C]Glc-labeled proteins that can considerably improve the structural characterization of insoluble, noncrystalline biological assemblies. We have presented a strategy to reach a complete resonance assignment via a bidirectional sequential walk, and the method is simple and sensitive because it is based only on one-step polarization transfer (PDSD and PAIN-CP). Extension to recent assignment protocols[38,39] based on sophisticated 3D

experiments could be envisioned for larger systems. Long-range distances already appear in the 400 ms PDSD spectrum of the [2-^{13}C]Glc-labeled needles (marked by asterisks in Figure 3c); the combined effect of the almost 2-fold improved resolution and the reduced number of assignment possibilities should allow for their assignment at a high confidence level, which to date has been the bottleneck for 3D structure determination by ssNMR analysis. The use of PDSD spectra with longer mixing times to collect intramolecular distance restraints and the determination of the intermolecular interface of PrgI in the needle[30] are currently underway in our laboratory with the aim of obtaining an atomic-resolution structure of the TTSS needle.

■ ASSOCIATED CONTENT

Supporting Information. Details about ssNMR experiments, PrgI assignment table, line width measurements, and enlarged versions of all spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
adla@nmr.mpibpc.mpg.de

■ ACKNOWLEDGMENT

We are grateful to C. Griesinger for continuous support of our work, B. Angerstein for expert technical assistance, and D. Riedel for the electron microscopy image. We thank the Max Planck Society, the DFG (Emmy Noether Fellowship to A. Lange), the Fondation Bettencourt Schueller, EMBO (long-term fellowship to A. Loquet), and the China Scholarship Council (Ph.D. scholarship to G.L.) for financial support.

■ REFERENCES

(1) Griffin, R. G. *Nat. Struct. Biol.* **1998**, *5*, 508.
(2) Böckmann, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 6110.
(3) McDermott, A. E. *Annu. Rev. Biophys.* **2009**, *38*, 385.
(4) Renault, M.; Cukkemane, A.; Baldus, M. *Angew. Chem., Int. Ed.* **2010**, *49*, 8346.
(5) Hong, M. *J. Magn. Reson.* **1999**, *139*, 389.
(6) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98.
(7) Loquet, A.; Bardiaux, B.; Gardiennet, C.; Blanchet, C.; Baldus, M.; Nilges, M.; Malliavin, T.; Bockmann, A. *J. Am. Chem. Soc.* **2008**, *130*, 3579.
(8) Manolikas, T.; Herrmann, T.; Meier, B. H. *J. Am. Chem. Soc.* **2008**, *130*, 3959.
(9) Zech, S. G.; Wand, A. J.; McDermott, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 8618.
(10) Franks, W. T.; Wylie, B. J.; Schmidt, H. L.; Nieuwkoop, A. J.; Mayrhofer, R. M.; Shah, G. J.; Graesser, D. T.; Rienstra, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4621.
(11) Bertini, I.; Bhaumik, A.; De Paepe, G.; Griffin, R. G.; Lelli, M.; Lewandowski, J. R.; Luchinat, C. *J. Am. Chem. Soc.* **2010**, *132*, 1032.
(12) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742.
(13) Jaroniec, C. P.; MacPhee, C. E.; Bajaj, V. S.; McMahon, M. T.; Dobson, C. M.; Griffin, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 711.
(14) Wasmer, C.; Lange, A.; Van Melckebeke, H.; Siemer, A. B.; Riek, R.; Meier, B. H. *Science* **2008**, *319*, 1523.
(15) Helmus, J. J.; Surewicz, K.; Nadaud, P. S.; Surewicz, W. K.; Jaroniec, C. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6284.
(16) Loquet, A.; Bousset, L.; Gardiennet, C.; Sourigues, Y.; Wasmer, C.; Habenstein, B.; Schutz, A.; Meier, B. H.; Melki, R.; Bockmann, A. *J. Mol. Biol.* **2009**, *394*, 108.
(17) Nielsen, J. T.; Bjerring, M.; Jeppesen, M. D.; Pedersen, R. O.; Pedersen, J. M.; Hein, K. L.; Vosegaard, T.; Skrydstrup, T.; Otzen, D. E.; Nielsen, N. C. *Angew. Chem., Int. Ed.* **2009**, *48*, 2118.
(18) Chimon, S.; Shaibat, M. A.; Jones, C. R.; Calero, D. C.; Aizezi, B.; Ishii, Y. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1157.
(19) Ahmed, M.; Davis, J.; Aucoin, D.; Sato, T.; Ahuja, S.; Aimoto, S.; Elliott, J. I.; Van Nostrand, W. E.; Smith, S. O. *Nat. Struct. Mol. Biol.* **2010**, *17*, 561.
(20) Mainz, A.; Jehle, S.; van Rossum, B. J.; Oschkinat, H.; Reif, B. *J. Am. Chem. Soc.* **2009**, *131*, 15968.
(21) Han, Y.; Ahn, J.; Concel, J.; Byeon, I. J.; Gronenborn, A. M.; Yang, J.; Polenova, T. *J. Am. Chem. Soc.* **2010**, *132*, 1976.
(22) Jehle, S.; Rajagopal, P.; Bardiaux, B.; Markovic, S.; Kuhne, R.; Stout, J. R.; Higman, V. A.; Klevit, R. E.; van Rossum, B.-J.; Oschkinat, H. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1037.
(23) Lange, A.; Giller, K.; Hornig, S.; Martin-Eauclaire, M. F.; Pongs, O.; Becker, S.; Baldus, M. *Nature* **2006**, *440*, 959.
(24) Cady, S. D.; Schmidt-Rohr, K.; Wang, J.; Soto, C. S.; Degrado, W. F.; Hong, M. *Nature* **2010**, *463*, 689.
(25) Shi, L.; Ahmed, M. A.; Zhang, W.; Whited, G.; Brown, L. S.; Ladizhansky, V. *J. Mol. Biol.* **2009**, *386*, 1078.
(26) Schneider, R.; Etzkorn, M.; Giller, K.; Daebel, V.; Eisfeld, J.; Zweckstetter, M.; Griesinger, C.; Becker, S.; Lange, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 1882.
(27) McDermott, A.; Polenova, T.; Böckmann, A.; Zilm, K. W.; Paulsen, E. K.; Martin, R. W.; Montelione, G. T. *J. Biomol. NMR* **2000**, *16*, 209.
(28) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255.
(29) Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. *J. Biomol. NMR* **2007**, *38*, 199.
(30) Loquet, A.; Giller, K.; Becker, S.; Lange, A. *J. Am. Chem. Soc.* **2010**, *132*, 15164.
(31) Cornelis, G. R.; Wolf-Watz, H. *Mol. Microbiol.* **1997**, *23*, 861.
(32) Kubori, T.; Matsushima, Y.; Nakamura, D.; Uralil, J.; Lara-Tejero, M.; Sukhan, A.; Galan, J. E.; Aizawa, S. I. *Science* **1998**, *280*, 602.
(33) Poyraz, O.; Schmidt, H.; Seidel, K.; Delissen, F.; Ader, C.; Tenenboim, H.; Goosmann, C.; Laube, B.; Thunemann, A. F.; Zychlinsky, A.; Baldus, M.; Lange, A.; Griesinger, C.; Kolbe, M. *Nat. Struct. Mol. Biol.* **2010**, *17*, 788.
(34) Lange, A.; Luca, S.; Baldus, M. *J. Am. Chem. Soc.* **2002**, *124*, 9704.
(35) Lewandowski, J. R.; De Paepe, G.; Griffin, R. G. *J. Am. Chem. Soc.* **2007**, *129*, 728.
(36) Hodgkinson, P.; Emsley, L. *J. Magn. Reson.* **1999**, *139*, 46.
(37) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G. *Mol. Phys.* **1998**, *95*, 1197.
(38) Franks, W. T.; Kloepper, K. D.; Wylie, B. J.; Rienstra, C. M. *J. Biomol. NMR* **2007**, *39*, 107.
(39) Schuetz, A.; Wasmer, C.; Habenstein, B.; Verel, R.; Greenwald, J.; Riek, R.; Bockmann, A.; Meier, B. H. *ChemBioChem* **2010**, *11*, 1543.