Protein Structures

DOI: 10.1002/anie.200801352

3D Protein Structures by Solid-State NMR Spectroscopy: Ready for High Resolution**

*Anja Böckmann**

fibrils · membrane proteins · NMR spectroscopy · protein structures

The solution of high-resolution 3D protein structures is a major achievement of structural biology. Protein structures are the basis for a detailed understanding of how a protein functions; how enzymes do chemistry; how molecules are transported through membranes; and how mechanical forces can be exerted. They allow the identification of the key motifs in three-dimensional protein structures and as such allow the development of strategies to undermine protein function through the design of inhibitors and also to produce mutants to answer detailed questions. 3D structures form the prerequisite for other structural studies—such as the investigation of protein–protein interactions or interactions with nucleic acids, lipids, or solvent—and also allow protein dynamics to be mapped in space. Without 3D structures, this understanding of protein dynamics cannot be translated into spatially resolved information.

Magic-angle-spinning (MAS) solid-state NMR experiments for structure determination of uniformly ^{13}C,^{15}N isotopically enriched proteins have undergone major advances recently, and the method is on the way to joining X-ray crystallography and liquid-state NMR spectroscopy as a tool for structural biology. Solid-state NMR spectroscopy mainly aims at structural studies of insoluble proteins, such as fibrils and membrane proteins. These proteins are poorly represented in the Protein Data Bank,[1] as their structures are difficult to obtain by other methods. A decisive step was the development of studies on fully ^{13}C,^{15}N-labeled proteins by solid-state NMR spectroscopy at the end of the 1990s; in contrast to studies using few selective labels, these approaches allow information to be obtained from the entire protein sequence. The proof that sequential resonance assignments using solid-state NMR experiments are possible was a breakthrough and proved to be immediately useful for a plethora of studies based on these chemical shifts including protein interactions, dynamics, and folding.

[*] Dr. A. Böckmann
Institut de Biologie et Chimie des Protéines UMR 5086
CNRS/Université de Lyon 1
IFR 128 BioSciences
7 passage du Vercors, 69367 Lyon (France)
Fax: (+33) 4-7272-2649
E-mail: a.bockmann@ibcp.fr

[**] This work has been supported by the French ANR (JC05_44957, ANR-07-PCVI-0013-03). I thank Antoine Loquet for color figures.

For the determination of 3D high-resolution structures by solid-state NMR spectroscopy, however, two major obstacles must be overcome: How can spectra be recorded that yield distance information, and how can this information be extracted from the spectra? Distance information can be recorded in proteins between proton, carbon, and nitrogen spins. For protons, the difficulty lies in the poor resolution, whereas measurements between carbon spins were hampered by the omnipresent strong coupling between adjacent spins. This coupling makes polarization transfer to remote spins a challenge. The polarization is indeed trapped between the directly bonded spins and is barely transferred beyond the nitrogen atoms, limiting the spin system of each amino acid, a phenomenon called dipolar truncation. The large number of spins and the still relatively broad lines observed in solid-state NMR spectra leads to multiple possible assignments of the peaks, if only the chemical shift is considered; the assignment of cross-peaks thus remains spectrally ambiguous, which makes the extraction of distance information from the spectra difficult.

The first structure calculation by the group of Oschkinat in 2002[2] using extensively labeled protein samples was thus a breakthrough. Extensive labeling is achieved by adding [1,3-^{13}C]- or [2-^{13}C]glycerol to the growth medium, which results in "checkerboard" labeling of approximately one out of two carbon atoms, but with two ^{13}C spins rarely being found as nearest neighbors (Figure 1 a). This pattern relieves both the dipolar truncation, as well as the assignment problem. Removing nearest neighbors removes the strong one-bond couplings and divides the number of carbon resonances roughly by two, thus greatly reducing spectral crowding. Moreover, the signals are line narrowed since the J couplings are suppressed.

Despite the success of this technique, several differently labeled samples were nevertheless necessary to obtain high-resolution structures with this approach, and the wish for more general solutions using a unique, fully labeled sample drove several subsequent developments. Distance measurements between protons (Figure 1 d), in analogy to liquid-state NMR spectroscopy, become possible in solid proteins when they are either detected on heteronuclei showing better resolved spectra, or when proton detection is made possible by using the latest advances in technology and methodology, such as high magnetic fields, high frequency MAS, and sample deuteration.[3–6] First proposed by Baldus and co-workers[7] in

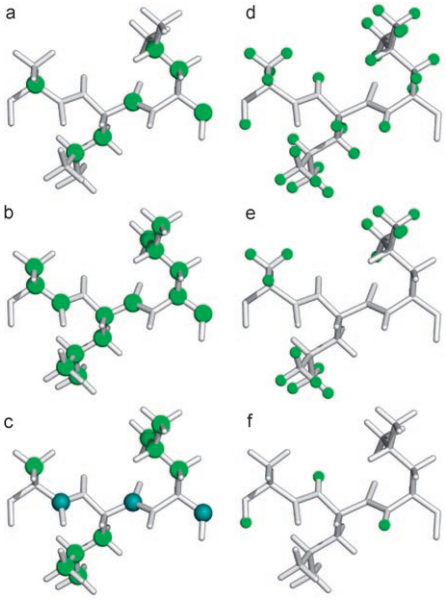

Figure 1. Spins for which distance restraints can be obtained from the different experiments described in the text. a) ^{13}C spins in extensively labeled samples; b) all carbon spins, between which PDSD, DARR, and PAR experiments may record distance restraints; c) spins of carbonyl (dark green) and aliphatic carbon atoms (light green) recoupled in selective recoupling experiments as HBR[2]; d) spins available for the measurement of interproton distances in CHHC and e) in J-CHHC experiments; f) spins addressed in direct proton-detected experiments on deuterated samples, whose ^{2}H atoms have been partially back-exchanged with ^{1}H atoms.

the field of protein structure determination, the detection of proton distance restraints on carbon or nitrogen spins yields a similar kind of restraint as obtained from NOESY spectra (NMR spectroscopy in solution), and also gives a $1/r^6$ distance dependence. A 2D spectra resulting from so-called CHHC experiment (proton-mediated, carbon-detected correlation experiment) is shown in Figure 2 for the microcrystalline Crh protein; it contains cross-signals reflecting local and long-range distance information and offers the necessary distance information for structure calculation.[8,9] The relationship between the cross-peak intensities and the distance is less quantitative than in NOESY spectra because of multispin effects, as well as the difficulties to transfer polarization from a carbon or nitrogen atom only to a selected proton neighbor. Still, semiquantitative measurements can be obtained, and the presence of a cross-peak generally is interpreted in the way that the two spins are located in a certain distance range from each other, without further analysis of the peak intensities. This approach has been successfully used for the measurement of distance restraints in the photosynthetic light-harvesting 2 protein complex,[10] and for the structure determination of drugs[11] and small proteins, including Kaliotoxin in its free[8] and bound[12] state.

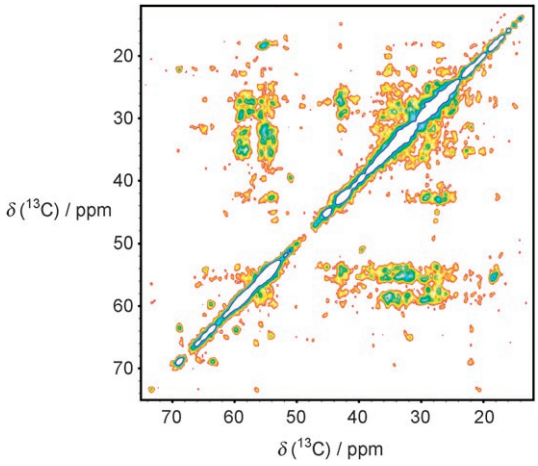

Figure 2. CHHC spectrum of the fully labeled 2×10.4 kDa Crh protein, measured at a proton field strength of 500 MHz. The cross-signals contain local, medium-, and long-range distance information.

Lesage and co-workers recently showed that a variant of the CHHC experiment (J-CHHC; J = *J*-coupling transferred)[13] efficiently overcomes one of the drawbacks of the CHHC experiment, being that contacts between methyl groups are difficult to observe owing to their unfavorable polarization transfer under the three CP steps (CP = cross polarization) in the sequence. By using scalar rather than dipolar polarization transfer, methyl groups can be efficiently selected (Figure 1 e), which reveals important restraints in the hydrophobic core of the protein.

A major drawback of the above-described heteronuclei-detected proton correlation experiments is their poor sensitivity. This problem can be resolved by using deuterated proteins whose ^{2}H atoms are partially back-exchanged with ^{1}H atoms; these proteins allow the direct measurement of proton distances.[4,14,15] Recently, proton spin dilution with ^{2}H, combined with high magnetic fields, high-frequency sample spinning, and triple-resonance experiments enabled proton detection with high resolution and sensitivity in the 55 residue GB1 protein.[16] Heteronuclear 3D experiments give the necessary spectral resolution to identify contacts, mainly between amide protons, but also with exchangeable side-chain protons from asparagine, glutamine, tryptophane, and threonine. The proton-detected experiments have to include a water-suppression sequence, which is presently technically less straightforward to implement then in liquid-state NMR spectroscopy, for which water suppression schemes by means of gradient pulses are routinely used. This feature is present only in a few MAS solid-state NMR probes today.[5,17] Alternative solutions are, however, under rapid progress.[18] A drawback of the measurement of restraints between mainly amide protons is that the dense network of protons in fully protonated samples is not present (Figure 1 f). This might present an obstacle for the structure determination, for example, of helical proteins. Also, side-chain conformations can only be poorly determined with this approach, and structures of bigger and more complex molecules may prove difficult to resolve.

The measurement of distances between carbon atoms was long thought to be very difficult in fully labeled protein samples because of dipolar truncation. The first indications[19,20] that simple proton-driven spin-diffusion experiments suffer less from dipolar truncation than recoupling experiments were confirmed by theoretical studies by Ernst

and co-workers.[21] Indeed, polarization transfer in proton-driven sequences relies on a second-order effect involving cross-terms between two dipolar couplings, which is less sensitive to dipolar truncation than the dipolar flip-flop term. Measurement of the distance between carbon atoms in fully labeled systems was recently demonstrated.[22,23] In the work of Meier and co-workers, the structure of ubiquitin was determined by using the DARR pulse sequence (DARR = dipolar-assisted rotational resonance), which is a variant of proton-driven experiments to measure restraints between carbon atoms. However, the distance measurements remain imprecise because of relayed magnetization transfer; the spectra contain very many signals, as all intraresidue and sequential cross-peaks are observed in addition to the signal containing the desired long-range information.

An interesting approach to more precise distance measurements is the use of selective recoupling methods. Chemical-shift-selective recoupling methods were already used for the structure determination of small molecules with resolved chemical shifts, by recoupling one spin pair[24] or one particular dipolar interaction[25] in each experiment. This process can be tedious for larger molecules. Multiple measurements of the restraints between nitrogen and carbon spins can be carried out by using variants of the TEDOR experiment (TEDOR = transfer echo double resonance).[26,27] This method was used for the structure determination of the amyloidogenic protein transthyretin.[28] To reduce the number of required carbon-distance measurements, Ladizhansky and co-workers recently introduced an experiment to measure distances between multiple carbonyl and side-chain carbon atoms with good precision (Figure 1c).[29] The underlying 3D experiments include polarization transfer from N to $C_{carbonyl}$ spins to give a 2D N–C′ correlation map, which resolved most of the spins in the 55 residue GB1 protein they investigated, followed by selective transfer to side-chain spins by using the HBR^2 sequence (homogeneously broadened version of R^2). Different spinning frequencies are then used to perform a series of experiments; distance information is then extracted from the spinning-speed dependence of the polarization transfer.[30] In contrast to chemical-shift-selective recoupling experiments, the homogeneously broadened version can recouple entire chemical-shift ranges and thus allows the extraction of multiple distances at a precision usually better than ± 1 Å. A slight drawback of this experiment is the need to record spectra at different spinning frequencies. Also, the number of constraints that can potentially be measured is smaller than when distances between all carbon or proton spins are measurable without restrictions, considering that a high number of restraints is still the major asset for high-resolution structure determination.

A new class of carbon correlation experiments has been recently developed by Griffin and co-workers.[31] In these experiments, carbon/nitrogen spin interactions are mediated by a third spin, in general a nearby proton spin, which has to be positioned in a favorable geometry with respect to the two other spins. This dependence on the position of the proton "helper" spin means that directly bound carbon spins have unfavorable recoupling conditions and that long-range transfers, for which it is more likely to find a proton spin in-between two carbon spins, are favored. The resulting spectra are thus much less contaminated by intraresidue and sequential transfers then traditional carbon correlation spectra such as DARR or PDSD (proton-driven spin diffusion), and the interesting long-range information is more easily accessible.

Angle measurements are complementary to distance measurements in protein-structure determination. The use of dihedral angle restraints predicted from information on chemical shifts is widely used in NMR spectroscopy.[34] Rienstra and co-workers recently demonstrated that the use of vector angle restraints can greatly improve structure quality.[35] These restraints define the relative orientation of ^1H–^{15}N and ^1H–^{13}C dipolar vectors, and can be accessed by dipolar shift experiments. Their combined use with ^1H- and ^{13}C-distance restraints from CHHC and DARR spectra allowed the atomic structure of GB1 to be determined to high precision.

A great leap has thus been made in NMR methodology in recent years; there are now numerous approaches for the measurement of distance restraints in fully labeled proteins. The most recent hallmark in this respect is the high-resolution 3D structure determination of a prion protein by solid-state NMR spectroscopy by Meier and co-workers. They used a combination of several of the approaches discussed above to deduce the structure of the C-terminal portion of the Het-s prion protein in its amyloid fibrillar form.[36]

In most proteins used as examples in the work described herein, the distance information could still be extracted from the spectra by hand. Considering the line-widths commonly observed in solid-state NMR spectra of protein samples (0.5–1 ppm), manual assignment rapidly becomes a difficult task as the number of spins increases. If narrower lines can be achieved, the size limit could be increased further, but at some point manual assignment becomes tedious or impossible. Ambiguities in cross-peak assignments can be resolved, not only by using homology models, but also by iterative assignment procedures,[37] as commonly implemented in programs such as ARIA or ATNOS/CANDID. These protocols handle the fact that most cross-peaks can have several possible assignments by discarding the least probable options by using structural models generated iteratively from the different possible restraints. Recently, close collaboration between solid-state NMR spectroscopists and developers of different programs used in solution NMR spectroscopy has enabled the protocols to be adapted to the special needs of solid-state NMR spectroscopy. Indeed, in ARIA it is now possible to input cross-peaks from CHHC/NHHN spectra reflecting restraints between protons (but with carbon/nitrogen chemical shifts) as well as between carbon/nitrogen spins. This approach has led to a high-definition structure of the Crh dimer (Figure 3).[32] The information contained in checkerboard labeling has been included in the program SOLARIA,[38] and distance calibration taking into account the fast polarization transfer along direct bonds was addressed by ATNOS/CANDID.[23]

Together, the development of complementary protocols to measure and assign distance restraints has recently enabled several high-resolution protein structures to be determined by solid-state NMR spectroscopy with a precision approaching

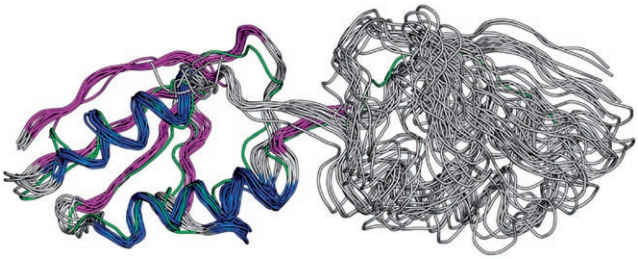

Figure 3. Structure bundle of the ten lowest energy structures of the microcrystalline 2×10.4 kDa Crh protein determined by CHHC and NHHC solid-state NMR experiments.[32] The structures are superimposed on one monomer, and the crystal structure[33] is shown in green.

that obtained by NMR spectroscopy in solution. Solid-state NMR spectroscopy as a technique is now preparing to tackle problems in structural biology concerning proteins beyond the scope of techniques such as X-ray crystallography and NMR spectroscopy in solution.

Published online: July 4, 2008

[1] H. M. Berman, K. Henrick, H. Nakamura, *Nat. Struct. Biol.* **2003**, *10*, 980.
[2] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, *Nature* **2002**, *420*, 98.
[3] A. E. McDermott, F. J. Creuzet, A. C. Kolbert, R. G. Griffin, *J. Magn. Reson.* **1992**, *98*, 408.
[4] B. Reif, C. P. Jaroniec, C. M. Rienstra, M. Hohwy, R. G. Griffin, *J. Magn. Reson.* **2001**, *151*, 320.
[5] V. Chevelkov, B. J. van Rossum, F. Castellani, K. Rehbein, A. Diehl, M. Hohwy, S. Steuernagel, F. Engelke, H. Oschkinat, B. Reif, *J. Am. Chem. Soc.* **2003**, *125*, 7788.
[6] E. K. Paulson, C. R. Morcombe, V. Gaponenko, B. Dancheck, R. A. Byrd, K. W. Zilm, *J. Am. Chem. Soc.* **2003**, *125*, 15831.
[7] A. Lange, S. Luca, M. Baldus, *J. Am. Chem. Soc.* **2002**, *124*, 9704.
[8] A. Lange, S. Becker, K. Seidel, K. Giller, O. Pongs, M. Baldus, *Angew. Chem.* **2005**, *117*, 2125–2129; *Angew. Chem. Int. Ed.* **2005**, *44*, 2089–2092.
[9] C. Gardiennet, A. Loquet, M. Etzkorn, H. Heise, M. Baldus, A. Böckmann, *J. Biomol. NMR* **2008**, *40*, 239–250.
[10] S. Ganapathy, A. J. van Gammeren, F. B. Hulsbergen, H. J. de Groot, *J. Am. Chem. Soc.* **2007**, *129*, 1504.
[11] A. Lange, T. Schupp, F. Petersen, T. Carlomagno, M. Baldus, *ChemMedChem* **2007**, *2*, 522.
[12] A. Lange, K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, *Nature* **2006**, *440*, 959.
[13] A. Loquet, S. Laage, C. Gardiennet, B. Elena, L. Emsley, A. Böckmann, A. Lesage, *J. Am. Chem. Soc.*, in press.
[14] B. Reif, B. J. Van Rossum, F. Castellani, K. Rehbein, A. Diehl, H. Oschkinat, *J. Am. Chem. Soc.* **2003**, *125*, 1488.
[15] E. K. Paulson, C. R. Morcombe, V. Gaponenko, B. Dancheck, R. A. Byrd, K. W. Zilm, *J. Am. Chem. Soc.* **2003**, *125*, 14222.
[16] D. H. Zhou, J. J. Shea, A. J. Nieuwkoop, W. T. Franks, B. J. Wylie, C. Mullen, D. Sandoz, C. M. Rienstra, *Angew. Chem.* **2007**, *119*, 8532; *Angew. Chem. Int. Ed.* **2007**, *46*, 8380.
[17] P. Charmont, A. Lesage, S. Steuernagel, F. Engelke, L. Emsley, *J. Magn. Reson.* **2000**, *145*, 334.
[18] D. H. Zhou, C. M. Rienstra, *J. Magn. Reson.* **2008**, *192*, 167–172.
[19] E. Crocker, A. B. Patel, M. Eilers, S. Jayaraman, E. Getmanova, P. J. Reeves, M. Ziliox, H. G. Khorana, M. Sheves, S. O. J. Smith, *J. Biomol. NMR* **2004**, *29*, 11.
[20] D. Marulanda, M. L. Tasayco, M. Cataldi, V. Arriaran, T. Polenova, *J. Phys. Chem. B* **2005**, *109*, 18135.
[21] A. Grommek, B. H. Meier, M. Ernst, *Chem. Phys. Lett.* **2006**, *427*, 404.
[22] A. Egawa, T. Fujiwara, T. Mizoguchi, Y. Kakitani, Y. Koyama, H. Akutsu, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 790.
[23] T. Manolikas, T. Herrmann, B. H. Meier, *J. Am. Chem. Soc.* **2008**, *130*, 3959–3966.
[24] C. M. Rienstra, L. Tucker-Kellogg, C. P. Jaroniec, M. Hohwy, B. Reif, M. T. McMahon, B. Tidor, T. Lozano-Perez, R. G. Griffin, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10260.
[25] K. Nomura, K. Takegoshi, T. Terao, K. Uchida, M. Kainosho, *J. Am. Chem. Soc.* **1999**, *121*, 4064.
[26] C. P. Jaroniec, C. Filip, R. G. Griffin, *J. Am. Chem. Soc.* **2002**, *124*, 10728.
[27] J. J. Helmus, P. S. Nadaud, N. Höfer, C. P. Jaroniec, *J. Chem. Phys.* **2008**, DOI: 10.1063/1.2817638.
[28] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, R. G. Griffin, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 711.
[29] R. Janik, X. Peng, V. Ladizhansky, *J. Magn. Reson.* **2007**, *188*, 129.
[30] X. Peng, D. Libich, R. Janik, G. Harauz, V. Ladizhansky, *J. Am. Chem. Soc.* **2008**, *130*, 359.
[31] J. R. Lewandowski, G. De Paepe, R. G. Griffin, *J. Am. Chem. Soc.* **2007**, *129*, 728.
[32] A. Loquet, B. Bardiaux, C. Gardiennet, C. Blanchet, M. Baldus, M. Nilges, T. Malliavin, A. Böckmann, *J. Am. Chem. Soc.* **2008**, *130*, 3579.
[33] M. Juy, F. Penin, A. Favier, A. Galinier, R. Montserret, R. Haser, J. Deutscher, A. Böckmann, *J. Mol. Biol.* **2003**, *332*, 767.
[34] G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **1999**, *13*, 289.
[35] W. T. Franks, B. J. Wylie, H. L. F. Schmidt, A. J. Nieuwkoop, R.-M. Mayrhofer, G. J. Shah, D. T. Graesser, C. M. Rienstra, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4621- 4626.
[36] C. Wasmer, A. Lange, H. V. Melckebeke, A. B. Siemer, R. Riek, B. H. Meier, *Science* **2008**, *319*, 1523.
[37] M. Nilges, *J. Mol. Biol.* **1995**, *245*, 645.
[38] M. Fossi, F. Castellani, M. Nilges, H. Oschkinat, B. J. van Rossum, *Angew. Chem.* **2005**, *117*, 6307; *Angew. Chem. Int. Ed.* **2005**, *44*, 6151.

