Biomolecular NMR Spectroscopy

Complete Assignment of Heteronuclear Protein Resonances by Protonless NMR Spectroscopy**

Wolfgang Bermel, Ivano Bertini, Luminita Duma, Isabella C. Felli, Lyndon Emsley, Roberta Pierattelli, and Paul R. Vasos*

The complete assignment of the resonances of a protein is key to the determination of its solution structure by NMR spectroscopy and for the study of protein–protein and protein–ligand interactions. The proton-based assignment strategy usually starts with the correlation of individual resonances of each amino acid residue through scalar connectivities followed by linking them one after the other.[1,2] Although many different triple-resonance NMR spectroscopy experiments have been designed for full assignments,[2] spectral overlap can still lead to ambiguities. This poses a significant limiting factor in the cases of large and/or paramagnetic biomolecules.[3]

After the pioneering report of ^{13}C NMR spin-system assignments of ^{13}C-enriched *Anabaena* 7120 ferredoxin by Markley and co-workers,[4] heteronuclear NMR spectroscopy experiments were progressively abandoned in favor of ^{1}H-detection experiments. However, as was recently pointed out, heteronuclear NMR spectroscopy decreases the effect of detrimental transverse relaxation, which is typical of large or paramagnetic proteins.[5–17] For this reason, several heteronuclear NMR spectroscopy experiments for backbone assignment have been proposed for fully ^{13}C- and ^{15}N-enriched proteins.[13,14,17] Furthermore, backbone sequence-specific assignment by the recently-designed CANCO experiment has also been reported.[18] We present herein an extension of the set of *exclusively heteronuclear* experiments to protein side chain resonances for the complete heteronuclear assignment of a protein. With a novel CBCACO experiment the carbonyl carbon (CO) is linked to the C^{β} and to the C^{α} nuclei; the connection to the rest of the amino acid side chain is achieved through a ^{13}C–^{13}C TOCSY experiment with C^{α} detection. In these experiments, we have successfully implemented spin-state selection methods for the removal of signal splitting in the acquisition dimension which is caused by multiple ^{13}C–^{13}C scalar couplings. This makes ^{13}C detection an amenable tool for high-resolution NMR spectroscopy. The proposed assignment strategy is summarized in Figure 1. A

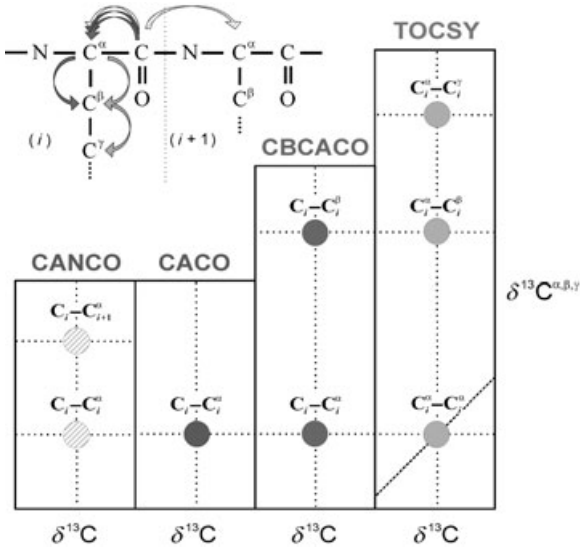

Figure 1. Illustration of the assignment procedure for ^{13}C NMR spectroscopy experiments. The assignment starts with analysis of the CACO experiment, which provides the correlation between the carbonyl carbon (CO) and the C^{α} nuclei of each amino acid. The spin-system assignment is extended to the C^{β} nuclei with the CBCACO experiment, and the process is completed with the TOCSY experiment, which provides correlation between the C^{α} and the other carbon nuclei of the amino acid side chain. The amino acid spin systems are finally assigned in a sequence-specific manner with the aid of a CANCO experiment,[18] which provides the correlation of each CO to the two neighboring C^{α} nuclei.

[*] Prof. I. Bertini
CERM and Department of Chemistry
University of Florence
Via Luigi Sacconi 6, 50019 Sesto Fiorentino (Italy)
Fax: (+39) 055-457-4271
E-mail: bertini@cerm.unifi.it
Dr. W. Bermel
Bruker BioSpin GmbH, Rheinstetten (Germany)
Dr. L. Duma, Prof. L. Emsley
Laboratoire de Chimie, UMR 5182 CNRS
Ecole Normale Superieure de Lyon (France)
Dr. I. C. Felli, Prof. R. Pierattelli, Dr. P. R. Vasos
CERM and Department of Chemistry
University of Florence (Italy)

[**] The authors are grateful to Dr. Alan Stern and Dr. Frank Delaglio for useful comments. L.D. was supported by the EC Marie Curie Fellowship (Contract HPMT-2000-000137). This work was supported in part by the EC Contracts HPRI-CT-2001-50026, HPRN-CT-2000-00092, and QLG2-CT-2002-00988. All NMR pulse sequences written for the Bruker Avance series spectrometers are available upon request.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200461794

detailed description of CBCACO and other novel experiments presented herein are provided in the Supporting Information.

Figure 2c shows the 2D version of the CBCACO spectrum recorded with isotopically enriched (70% ^{2}H, 98% ^{13}C and ^{15}N), reduced human Cu,Zn superoxide dismutase (SOD),[19] a dimeric protein of $M_W = 32000$. The 55-Hz CO line splitting due to the C^{α}–CO scalar coupling is removed by separation of the two multiplet components with the in-phase/anti-phase (IPAP) scheme,[20,21] followed by their recombination to increase the sensitivity of the spectrum[22] (Supporting Information). Upon comparison with the IPAP–CACO spectrum (Figure 2a), the additional cross-peaks in the IPAP–CBCACO spectrum allow identification of the types of amino acid groups present in the protein on the basis of the chemical shift of C^{β}.[23] Moreover, the resolution is enhanced through the chemical-shift dispersion of C^{β} resonances ($\approx$70 ppm), which are larger than those of C^{α} ($\approx$35 ppm). Notably, this single spectrum reveals all C^{α} and C^{β} signals that have been previously assigned.[24] Furthermore, the correlations of carbonyl and carboxylate nuclei of Asp, Asn, Glu, and Gln residues, which may be important in the study of surface properties in protein–protein interactions, are easily assigned, as they are observed in a specific region of the spectrum (centered at $\omega_2 = 180$ ppm and $\omega_1 = 35$ ppm). The IPAP–CBCACO experiment is designed for easy expansion

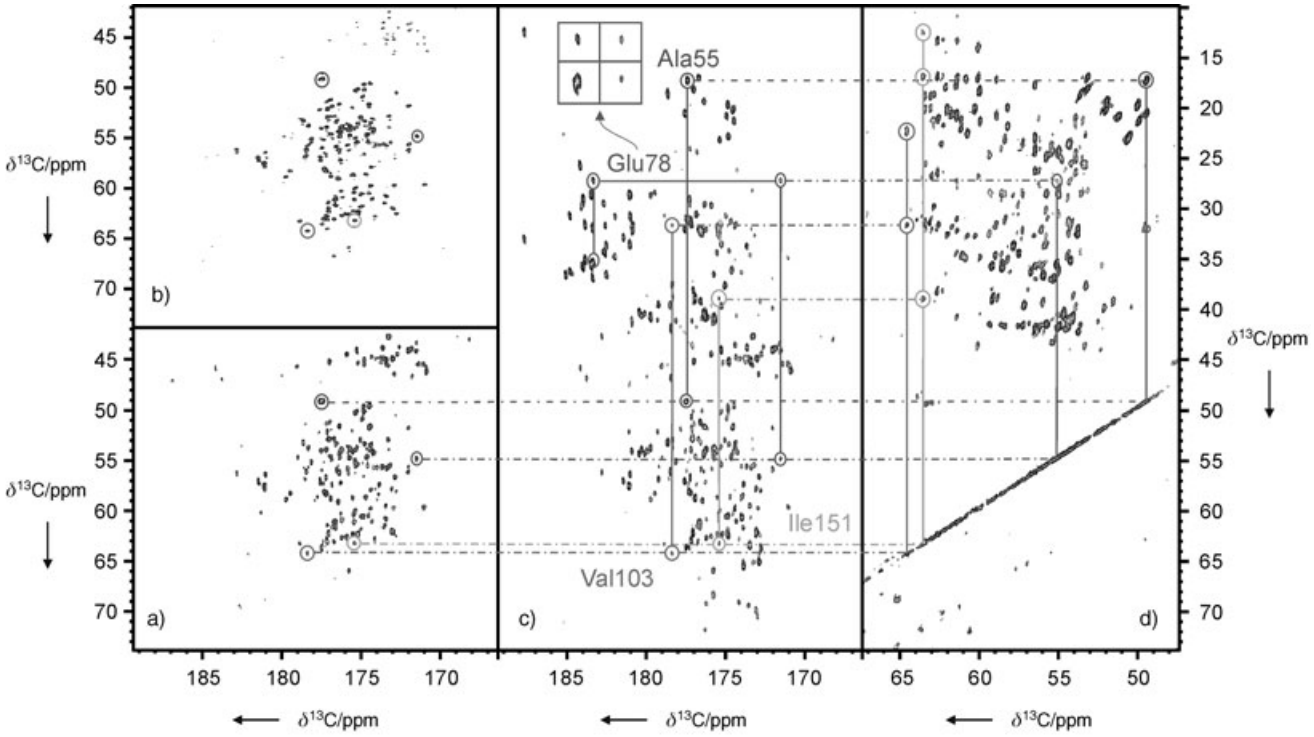

Figure 2. a) IPAP–CACO spectrum, b) DIPAP–COCA spectrum, c) 2D version (^{13}CO–$^{13}C^{\alpha/\beta}$) of the IPAP–CBCACO spectrum, and d) ^{13}C–^{13}C DIPAP–TOCSY spectrum recorded at 298 K with reduced Cu,Zn SOD (70% ^{2}H, 98% ^{13}C and ^{15}N) on a Bruker Avance NMR spectrometer (16.4 T, 700.06 MHz for ^{1}H, 176.03 MHz for ^{13}C) equipped with a ^{13}C-{^{1}H,^{15}N}TXO probe head. Chemical shifts were referenced to external 2,2-dimethyl-2-silapentane-5-sulfonic acid. The protein sample (2 mM) was prepared as previously reported[38] in phosphate buffer (20 mM, pH 5.0). The direct links between the different correlations in the spectra of panels a), c), and d) are reported for four different amino acid types. For comparison, panel b) shows the DIPAP–COCA spectrum rotated by 90°; the connectivities shown in panel a) are circled as well. The inset in panel c) shows an expansion of the four cross-peaks present in the IPAP–CBCACO spectrum for Glu78. For the spectra in panel a) and c), matrices of 1024×128 and 1024×192 complex points, respectively, in the ^{13}CO and $^{13}C^{\alpha/\beta}$ dimensions were acquired with 256 transients per increment and a recycle delay of 1.25 s; ^{1}H, ^{2}H, and ^{15}N were decoupled during the whole sequence. For each time increment in the indirect dimension, two FIDs were stored separately, one each for the anti-phase and in-phase components. The FIDs were then added and subtracted to separate the two multiplet components. These were then shifted to the center of the original multiplet by ($J_{CO-C^{\alpha}}$/2) Hz and again added to obtain a singlet.[22] Prior to Fourier transformation, an expansion to 2048×512 by linear prediction in the indirect dimension, together with zero filling and squared-sine-multiplication functions shifted by $\pi/4$ were applied in the t_2 and t_1 dimensions. For the spectrum in panel b) a matrix of 1024×96 complex points in the $^{13}C^{\alpha}$ and ^{13}CO dimensions was acquired, with 256 transients per increment and a recycle delay of 1.25 s; ^{1}H, ^{2}H, and ^{15}N were decoupled during the whole sequence. For each time increment in the indirect dimension, four FIDs were stored separately, and the resulting submatrices were added and subtracted in pairs to separate the four multiplet components, and then shifted by ($\pm J_{CO-C^{\alpha}}$/2) Hz and ($\pm J_{C^{\alpha}-C^{\beta}}$/2) Hz and added again to obtain a singlet.[22] Prior to Fourier transformation, an expansion to 2048×512 by linear prediction in the indirect dimension, together with zero filling and squared-sine-multiplication functions shifted by $\pi/4$ were applied in the t_2 and t_1 dimensions. For spectrum in panel d) a matrix of 1024×256 complex points was acquired, with 256 transients per increment and a recycle delay of 1.25 s; ^{1}H, ^{2}H, and ^{15}N were decoupled during the whole sequence, except during the DIPSI-3[39] spin lock of 23 ms. The processing of data for the four DIPAP submatrices was performed as described for panel b). Prior to Fourier transformation, an expansion of 2048×512 with zero filling and squared-sine-multiplication functions shifted by $\pi/4$ and $\pi/2$ were applied in the t_2 and t_1 dimensions, respectively.

into a 3D version by evolving the C^{α} chemical shift in the third dimension.

The IPAP filter can be implemented in a CBCACO experiment without introducing additional delays and it offers a gain in sensitivity with respect to other methods of *J*-coupling suppression.[*] The IPAP filter can, of course, be easily included in the CACO experiment.[**]

The extension of heteronuclear assignments through a ^{13}C–^{13}C TOCSY experiment[11,15] is conveniently performed by detection of the C^{α} resonances, which experience splitting from the coupling of the C^{α} nuclei to both the CO and the C^{β} nuclei. Therefore, a double IPAP (DIPAP) scheme[28] was implemented to remove the double splitting present on the C^{α} signal. As nuclear relaxation is a dominant limitation in complex pulse sequences, we succeeded in merging the two IPAP blocks into a single short building block. An example of C^{α} detection is provided by the DIPAP–COCA spectrum recorded on the same SOD sample (Figure 2b) at 16.4 T. It is nearly identical to the IPAP–CACO spectrum (Figure 2a) in terms of resolution and number of signals, except for signals that do not show the secondary coupling such as those of Gly residues. Furthermore, C^{α} detection is expected to be better than the detection of CO at high fields, as CO relaxation is dominated by the chemical-shift anisotropy (CSA) interaction that increases with the square of the magnetic field.[29,30] The detailed pulse sequence is reported in the Supporting Information.

Figure 2b shows that the DIPAP building block is quite effective. It allows one to successfully record highly resolved ^{13}C–^{13}C TOCSY spectra, as shown in Figure 2d. In the present case, the DIPAP–TOCSY experiment allowed a significant extension of the assignments for SOD[24] (expected C' and C^{α}, 100%; C^{β}, 98%; C^{γ}, 93%; C^{δ}, 89%; C^{ε}, 82%; side-chain CO, 100%).[***]

The same set of experiments was performed on a sample of ^{13}C,^{15}N-enriched oncomodulin (2.5 mM), a small Ca-binding protein of $M_W = 11\,500$. This enabled the assignment of all the resonances already obtained by standard ^{1}H-based NMR spectroscopy.[31] With this method of spin-state-selective detection, the assignments were expanded to almost all complete spin systems that were expected on the basis of the protein primary sequence (expected C', C^{α} and C^{β}, 100%; C^{γ} and C^{δ}, 98%; C^{ε}, 73%; side-chain CO, 100%;).[***]

In conclusion, the data presented herein show that removal of ^{13}C–^{13}C splitting in the acquisition dimension creates a significant step forward in direct-detection-^{13}C NMR spectroscopy and opens new prospects in the resolution of biomolecular NMR spectroscopic data. Indeed, there are a large number of experiments that could benefit from the present approach, including 3D experiments that are important to reduce spectral overlap typical of large molecules. Notably, a 3D-IPAP–CBCACO experiment on a sample of oncomodulin (2.5 mM) requires 16 scans, which is not an unreasonably long duration in comparison with standard proton-based 3D experiments at the same field. Similarly, the DIPAP–TOCSY experiment is highly sensitive and an optimal data set can be obtained in less than 24 h. The ^{13}C–^{13}C experiments presented herein can thus be included in any standard biomolecular NMR spectroscopy protocol for structure determination.

The measurement of ^{13}C–^{13}C couplings to determine residual dipolar couplings for the purposes of structure determination[32] can be performed by analysis of the individual submatrices obtained in the IPAP and DIPAP experiments, even if the resolution is still not ideal. Simultaneous decoupling of the different nuclei limits the effective RF field strength available for each nucleus as well as the duration of the acquisition time. Care should be taken to minimize overheating the sample during experiments. ^{13}C homodecoupling[10] is no longer necessary, making these experiments possible with any standard NMR spectrometer equipped for triple resonance experiments. The increase in sensitivity expected with the use of cryogenic technology in probe design, which is not yet fully exploited in ^{13}C NMR spectroscopy, will further assist in the establishment of direct-detection ^{13}C NMR as a routine technique for protein investigations and also for samples that are more dilute than those used in the experiments reported herein.

These experiments are also advantageous to the study of unfolded, highly mobile, and exchangeable systems in which the broadening of the NH signals cannot be recovered through the TROSY approach. In general, direct-detection ^{13}C NMR experiments are the only source of information for cases in which the ^{1}H resonances are broadened beyond detection through relaxation or exchange effects. Finally, ^{13}C NMR spectroscopy is the appropriate solution to paramagnetic relaxation as has been previously shown;[10,16] indeed this has been proven in studies of lanthanide-substituted oncomodulin, as a continuation of this research. In this case, a number of structural restraints based on paramagnetism are also available for solution-structure determination.[12,33–37]

[*] Several methods have been proposed to remove *J* splitting from the acquisition dimension, such as band-selective homodecoupling,[10] maximum entropy reconstruction,[25] and other post-acquisition methods.[26] We propose the method based on spin-state-selective schemes with the IPAP sequence[20,21] as an example, because it provides the best signal-to-noise ratio (S/N) in comparison with other methods (not shown). The gain in S/N compared with the most intense component of the coupled spectrum is close to that theoretically expected of SQRT(2), owing to the uniform *J*-coupling values found in ^{13}C signals. For band-selective homodecoupling, the lower S/N obtained results mainly from the higher intrinsic level of noise introduced by the acquisition method, rather than from unsatisfactory decoupling efficiency. Maximum entropy reconstruction apparently gives higher signal than the homodecoupled and nondecoupled spectra which comes at the expense of a nonuniform noise that surrounds the reconstructed peaks if the *J* coupling is not exactly matched.[25]

[**] Similarly, a S^3E filter[27] can be used, which is an even more compact way to obtain spin-state selection. The corresponding sequence and spectrum are reported in the Supporting Information.

[***] In the case of aromatic amino acids, an additional experiment is necessary to conveniently assign the aromatic ring carbon nuclei that resonate around 130 ppm.[23]

Received: August 26, 2004
Revised: December 16, 2005
Published online: April 14, 2005

Keywords: carbon NMR spectroscopy · NMR spectroscopy · protein structures · proteins · spin-state selection

[1] G. Wagner, *Prog. Nucl. Magn. Reson. Spectrosc.* **1990**, *22*, 101 – 139.
[2] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93 – 158.
[3] M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wüthrich, *J. Am. Chem. Soc.* **2000**, *122*, 7543 – 7548.
[4] B.-H. Oh, W. M. Westler, P. Darba, J. L. Markley, *Science* **1988**, *240*, 908 – 911.
[5] U. Kolczak, J. Salgado, G. Siegal, M. Saraste, G. W. Canters, *Biospectroscopy* **1999**, *5*, S19-S32.
[6] I. Bertini, Y.-M. Lee, C. Luchinat, M. Piccioli, L. Poggi, *ChemBioChem* **2001**, *2*, 550 – 558.
[7] Z. Serber, C. Richter, V. Dötsch, *ChemBioChem* **2001**, *2*, 247 – 251.
[8] M. Kostic, S. S. Pochapsky, T. C. Pochapsky, *J. Am. Chem. Soc.* **2002**, *124*, 9054 – 9055.
[9] T. E. Machonkin, W. M. Westler, J. L. Markley, *J. Am. Chem. Soc.* **2002**, *124*, 3204 – 3205.
[10] W. Bermel, I. Bertini, I. C. Felli, R. Kümmerle, R. Pierattelli, *J. Am. Chem. Soc.* **2003**, *125*, 16423 – 16429.
[11] A. Eletsky, O. Moreira, H. Kovacs, K. Pervushin, *J. Biomol. NMR* **2003**, *26*, 167 – 179.
[12] F. Arnesano, L. Banci, I. Bertini, I. C. Felli, C. Luchinat, A. R. Thompsett, *J. Am. Chem. Soc.* **2003**, *125*, 7200 – 7208.
[13] I. Bertini, I. C. Felli, R. Kümmerle, D. Moskau, R. Pierattelli, *J. Am. Chem. Soc.* **2004**, *126*, 464 – 465.
[14] I. Bertini, I. C. Felli, R. Kümmerle, C. Luchinat, R. J. Pierattelli, *J. Biomol. NMR* **2004**, *30*, 245 – 251.
[15] B. Vögeli, H. Kovacs, K. Pervushin, *J. Am. Chem. Soc.* **2004**, *126*, 2414 – 2420.
[16] T. E. Machonkin, W. M. Westler, J. L. Markley, *J. Am. Chem. Soc.* **2004**, *126*, 5413 – 5426.
[17] E. Babini, I. Bertini, F. Capozzi, I. C. Felli, M. Lelli, C. Luchinat, *J. Am. Chem. Soc.* **2004**, *126*, 10496 – 10497.
[18] I. Bertini, L. Duma, I. C. Felli, M. Fey, C. Luchinat, R. Pierattelli, P. R. Vasos, *Angew. Chem.* **2004**, *116*, 2307 – 2309; *Angew. Chem. Int. Ed.* **2004**, *43*, 2257 – 2259.
[19] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, A. Rosato, M. S. Viezzoli, *Biochemistry* **2000**, *39*, 9108 – 9118.
[20] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, *131*, 373 – 378.
[21] P. Andersson, J. Weigelt, G. Otting, *J. Biomol. NMR* **1998**, *12*, 435 – 441.
[22] N. C. Nielsen, H. Thøgersen, O. W. Sørensen, *J. Am. Chem. Soc.* **1995**, *117*, 11365 – 11366.
[23] H. Zhang, S. Neal, D. S. Wishart, *J. Biomol. NMR* **2003**, *25*, 173 – 195.
[24] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, M. S. Viezzoli, *Eur. J. Biochem.* **2002**, *269*, 1905 – 1915.
[25] N. Shimba, A. S. Stern, C. S. Craik, J. C. Hoch, V. Dötsch, *J. Am. Chem. Soc.* **2003**, *125*, 2382 – 2383.
[26] K. Pervushin, A. Eletsky, *J. Biomol. NMR* **2003**, *25*, 147 – 152.
[27] A. Meissner, J. O. Duus, O. W. Sørensen, *J. Magn. Reson.* **1997**, *128*, 92 – 97.
[28] L. Duma, S. Hediger, B. Brutscher, A. Bockmann, L. Emsley, *J. Am. Chem. Soc.* **2003**, *125*, 11816 – 11817.
[29] L. Zeng, M. W. F. Fischer, E. R. P. Zuiderweg, *J. Biomol. NMR* **1996**, *7*, 157 – 162.
[30] K. T. Dayie, G. Wagner, *J. Am. Chem. Soc.* **1997**, *119*, 7797 – 7806.
[31] E. Babini, I. Bertini, F. Capozzi, C. Del Bianco, D. Holleder, T. Kiss, C. Luchinat, A. Quattrone, *Biochemistry* **2004**, *43*, 16076 – 16085.
[32] A. Bax, G. Kontaxis, N. Tjandra, *Methods Enzymol.* **2001**, *339*, 127 – 174.
[33] I. Bertini, C. Luchinat, G. Parigi, *Concepts Magn. Reson.* **2002**, *14*, 259 – 286.
[34] I. Bertini, A. Donaire, B. Jimenez, C. Luchinat, G. Parigi, M. Piccioli, L. Poggi, *J. Biomol. NMR* **2001**, *21*, 85 – 98.
[35] G. Pintacuda, A. Moshref, A. Leonchiks, A. Sharipo, G. Otting, *J. Biomol. NMR* **2004**, *29*, 351 – 361.
[36] R. Barbieri, C. Luchinat, G. Parigi, *ChemPhysChem* **2004**, *5*, 797 – 806.
[37] T. Ikegami, L. Verdier, P. Sakhaii, S. Grimme, P. Pescatore, K. Saxena, K. M. Fiebig, C. Griesinger, *J. Biomol. NMR* **2004**, *29*, 339 – 349.
[38] E. D. Getzoff, D. E. Cabelli, C. L. Fisher, H. E. Parge, M. S. Viezzoli, L. Banci, R. A. Hallewell, *Nature* **1992**, *358*, 347 – 351.
[39] A. J. Shaka, C. J. Lee, A. Pines, *J. Magn. Reson.* **1988**, *77*, 274 – 293.

