



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

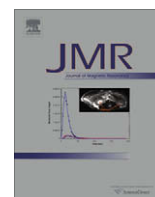
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Communication

A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment

Mehdi Mobli^{a,*}, Alan S. Stern^b, Wolfgang Bermel^c, Glenn F. King^a, Jeffrey C. Hoch^d^a Institute for Molecular Bioscience, The University of Queensland, St. Lucia, 4072 QLD, Australia^b Rowland Institute at Harvard, Cambridge, MA 02142, USA^c Bruker BioSpin GmbH, 76287 Rheinstetten, Germany^d University of Connecticut Health Center, Farmington, CT 06030-3305, USA

ARTICLE INFO

Article history:

Received 14 December 2009

Revised 27 January 2010

Available online 1 March 2010

Keywords:

Multidimensional NMR

Heteronuclear NMR

HCC(CO)NH-TOCSY

Maximum entropy

Non-uniform sampling

Protein assignment

ABSTRACT

One of the stiffest challenges in structural studies of proteins using NMR is the assignment of sidechain resonances. Typically, a panel of lengthy 3D experiments are acquired in order to establish connectivities and resolve ambiguities due to overlap. We demonstrate that these experiments can be replaced by a single 4D experiment that is time-efficient, yields excellent resolution, and captures unique carbon–proton connectivity information. The approach is made practical by the use of non-uniform sampling in the three indirect time dimensions and maximum entropy reconstruction of the corresponding 3D frequency spectrum. This 4D method will facilitate automated resonance assignment procedures and it should be particularly beneficial for increasing throughput in NMR-based structural genomics initiatives.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The development of non-uniform sampling (NUS) approaches to data collection in multidimensional NMR experiments represents the most fundamental change in the way that NMR data are collected since the development of Fourier transform (FT) NMR by Ernst and Anderson some 50 years ago [1]. Uniform sampling that is required by the discrete Fourier transform (DFT) prevents full realization of the benefits from costly ultra-high-field magnets because it is impractical to sample long evolution times, required for high resolution, in the indirect dimensions of multidimensional experiments. The use of non-uniform sampling is essential for making high-resolution 3D and 4D NMR experiments practical, especially for marginally soluble or transiently stable systems.

High-resolution multidimensional spectra are of particular interest in studies of complex biomolecules that are isotopically enriched with ¹³C and ¹⁵N [2]. These experiments are used both for resonance assignment of all ¹H, ¹³C and ¹⁵N atoms in the molecule and to extract interproton distance restraints from dipolar interactions. For proteins, this generally involves sequential assignment of backbone amide resonances followed by assignment of sidechain nuclei and, finally, extraction of distance and dihedral-angle restraints.

One of the most challenging steps in this process is the assignment of sidechain resonances. This process commonly involves the acquisition of two complementary 3D experiments, the H(CC)(CO)NH-TOCSY and (H)CC(CO)NH-TOCSY, which produce proton and carbon correlations, respectively [3]. These experiments reveal correlations between the amide ¹H–¹⁵N pair of residue *i* and the aliphatic sidechain atoms of the previous residue, *i* – 1. Thus, each spectrum contains roughly one resonance per aliphatic sidechain nucleus. These experiments, however, do not provide information about carbon–hydrogen connectivities, often leading to ambiguous assignments. This problem can be ameliorated using a 4D version of this experiment [4], but conventional uniform sampling of this experiment requires a prohibitively long acquisition time. In principle, it should be possible to reduce the experiment time through NUS.

A variety of methods have recently been introduced for processing NUS data with more than three dimensions [5–10]. These methods are generally restricted to situations of high signal-to-noise ratio (S/N). The relatively low sensitivity of HCC(CO)NH-TOCSY type experiments (compared to HNCOC, CBCA(CO)NH, etc.) makes them poor candidates for these reconstruction methods. An alternative approach that works well in situations of low S/N is maximum entropy reconstruction (MaxEnt) [11,12]. However, an implementation capable of reconstructing three indirect dimensions simultaneously (i.e., 4D experiments) has not been available (note that very recently the FM approach was applied to a 4D-NOESY dataset [13]: this method attempts to recover signal ampli-

* Corresponding author. Fax: +61 7 3346 2101.

E-mail address: m.mobli@uq.edu.au (M. Mobli).

Fig. 1. For ^1H and ^{15}N , narrow and wide rectangles represent 90° and 180° pulses (for hard pulses at 28.3 and 6.25 kHz), respectively. For ^{13}C , solid narrow and wide rectangles denote on-resonance Q5 and Q3 pulses [27], while open rectangles represent pulses that are applied off-resonance by means of a phaseramp [28]. The hatched pulse represents a C_α -selective Q3 pulse. For ^{13}C , every second 90° pulse is a time-reversed Q5. The pulse lengths are 274 μs for Q5/time-reversed Q5 pulses at a RF field strength of 16.7 kHz, 170 μs at 19.4 kHz for 180° pulses applied to all aliphatic carbons, and 667 μs at 4.95 kHz for the C_α -selective Q3 pulse. Solid shapes are selective pulses applied to the water magnetization (Sinc1 shape of 1 ms length at 425 kHz). The ^1H decoupling (DIPS12 [29] at 6.25 kHz) is flanked by 90° pulses given at the same power level. A DIPS12 [29] spinlock is applied for 11.5 ms at 10 kHz to achieve the CC TOCSY transfer. A garp4 supercycle is used for ^{15}N decoupling during acquisition (1.39 kHz) [30]. The delays are: $\tau_1 = 1/(4^3J_{\text{CH}})$ (1.7 ms), $\tau_2 = 1/(6^3J_{\text{CH}})$ (1.1 ms), $\tau_3 = 1/(4^3J_{\text{CCOCCO}})$ (4.5 ms), $\tau_4 = 1/(4^3J_{\text{CCOCCO}})$ (4.4 ms), $\tau_5 = 1/(4^3J_{\text{CON}})$ (12.4 ms), $\tau_6 = \tau_5$, $\tau_7 = (1/2^3J_{\text{NH}})$ (5.5 ms), $\tau_8 = 2.3$ ms $1/(4^3J_{\text{NH}})$ (2.3 ms; reduced to partially compensate for relaxation losses), $\epsilon_1 = \tau_p$ (^{13}C 180°) + initial setting of (1 - κ_3) $t_1/2$, $\epsilon_2 = \tau_p$ (^1H 180°) + initial setting of t_2 , $\epsilon_3 = \tau_p$ (^{13}C 180°) + initial setting of (1 - κ_3) $t_3/2$, $\delta_1 = \tau_1 + t_1/2$, $\delta_2 = (1 - \kappa_1)t_1/2$, $\delta_3 = \tau_1 - \kappa_1 t_1/2 + \epsilon_1$, $\delta_4 = \tau_2 + \epsilon_2$, $\delta_5 = \tau_6 - \kappa_3 t_3/2 + \epsilon_3$, $\delta_6 = (1 - \kappa_1)t_3/2$, $\delta_7 = \tau_6 + t_3/2$. All pulses are along x except otherwise noted. The phase cycles are $\phi_1 = x, x, x, x, x, x, x, x, -x, -x, -x, -x, -x, -x, -x, -x$; $\phi_2 = x, x, -x, -x$; $\phi_3 = x, x, x, x, -x, -x, -x, -x$; $\phi_4 = x, -x$; $\phi_{\text{receive}} = x, -x, -x, x, -x, x, -x, -x, x, -x, x, -x, -x, x, -x, x$. Gradients (1 ms) are sine-shaped with an amplitude of 25, 20, 30 and 15 G/cm, respectively. Labeling of chemical shifts is performed in a semi-constant time fashion [31–33]. Phase sensitive spectra are obtained using the States-TPPI method [34] by either incrementing ϕ_1 , ϕ_2 or decrementing ϕ_4 respectively in different experiments. Water suppression was achieved by using a soft pulse WATERGATE scheme [35,36].

where lw_k and sw_k are the linewidth and sweep width in the k^{th} dimension, respectively, dim is the number of dimensions and the equation may be further altered to include, for example, J modulation. The algorithm then calculates a rank for each point $r(i)$ according to:

$$r(i) = x(i)^{1/w(i)}$$

where $x(i)$ is a random number between 0 and 1 generated for each point. A subset of M points having the highest values of $r(i)$ are then selected from N total points. This method is fast and it avoids sampling the same point multiple times.

All data were processed using the Rowland NMR Toolkit (<http://nmrtoolkit.uchc.edu>). The 4D experiment was processed using the new 3D MaxEnt program (msa3d). The MaxEnt reconstruction parameters were determined using a previously reported approach ($def = 0.01$, $\lambda = 0.02$) [22,23]. Fig. 3 shows a 2D cross-section of the indirect ^1H and ^{13}C dimensions, at the direct ^1H and indirect ^{15}N frequencies corresponding to residue Ser30, revealing correlations from the sidechain protons of Arg29. This example illustrates: (i) the superb resolution of the 4D experiment, which is particularly useful for assigning complex sidechains; (ii) the excellent sensitivity of the 4D spectrum (note that Ser30 is one of the weakest peaks in the ^1H – ^{15}N HSQC spectrum; see Fig. 4); (iii) the ability to unambiguously determine which protons are attached to which carbons using the 4D experiment.

3. Discussion

The number of sample points required to faithfully reconstruct a spectrum depends on the complexity of the underlying data (e.g., the number of signals, dynamic range, etc.), with the dimensional-

ity of the experiment being of little consequence [20]. Since the number of unknowns (i.e., frequencies and phases of signals) in each of the two 3D experiments is roughly the same as in the 4D experiment, the number of data points required to obtain spectra with sufficient information content should be the same in all three experiments. The time required to carry out a 4D experiment would be about the same as the time required for two 3D experiments if all the hypercomplex components are collected (8 elements for 3D and 16 for 4D).

Crucially, however, in the two 3D experiments the signal acquired in one experiment does not contribute to the other whilst in the 4D all acquired data contribute to the sensitivity of the final reconstruction. This suggests that the 4D experiment will have the same overall sensitivity as both 3D experiments combined (neglecting losses due to differential relaxation), without incurring additional experiment time.

The 4D HCC(CO)NH-TOCSY experiment described here lends itself particularly well to rapid sidechain assignment in the context of structural genomics initiatives. The ability to unequivocally elucidate C–H correlations should facilitate automated assignment approaches, which currently suffer from the ambiguity present when the conventional experiments are used for aliphatic sidechain assignment [16,24]. Furthermore, dispersion of resonances into an additional dimension significantly reduces the possibility of peak overlap; for example, in the conventional 3D H(CC)(CO)NH-TOCSY experiment it would not be possible to resolve two overlapping peaks from protons attached to carbon atoms resonating at different frequencies, whereas in the 4D experiment these would be resolved and unambiguously assigned.

4. Conclusions

In conclusion, we have described an approach for aliphatic sidechain assignment employing NUS and non-Fourier spectrum analysis that offers significant advantages over existing methods. This is an attractive alternative to existing fast methods as it is practical, simple to implement and applicable to most proteins routinely analysed by NMR, and is likely to benefit all proteins where sampling to an optimal evolution time of $1.6 \times T_2$ in all dimensions is unfeasible [25,26]. Furthermore, the experiment yields results that can be readily analysed using either manual or automated approaches.

5. Experimental

The direct dimension was Fourier transformed. At each frequency point in the resulting interferogram, a 3D hyperplane (cube) was reconstructed using the MaxEnt approach. The time required to compute each cube using MaxEnt reconstruction varied and depends on the data in that cube. The example 2D plane of the 4D-NUS experiment shown in Fig. 3 was taken from a cube processed on a laptop computer and converged after a total of 4 h; the average processing time per cube was ~ 2 h/CPU. Thus to reconstruct 512 pts would take ~ 1000 h on a single CPU or 128 h (5 days) on an 8 core computer. The full dataset presented here was processed on an Apple X-grid cluster consisting of 104 CPUs and converged overnight.

The 3D H(CC)(CO)NH-TOCSY experiment was acquired using 32 and 64 complex samples in the nitrogen and proton dimensions, respectively, while the 3D (H)CC(CO)NH-TOCSY experiment was acquired using 28 and 56 complex samples in the nitrogen and carbon dimensions, respectively. Both 3D experiments were acquired using 16 transients per time increment. The 3D data was processed using linear prediction extrapolation, shifted sine-bell apodization, zero-filling and discrete Fourier transform (DFT) in the indirect

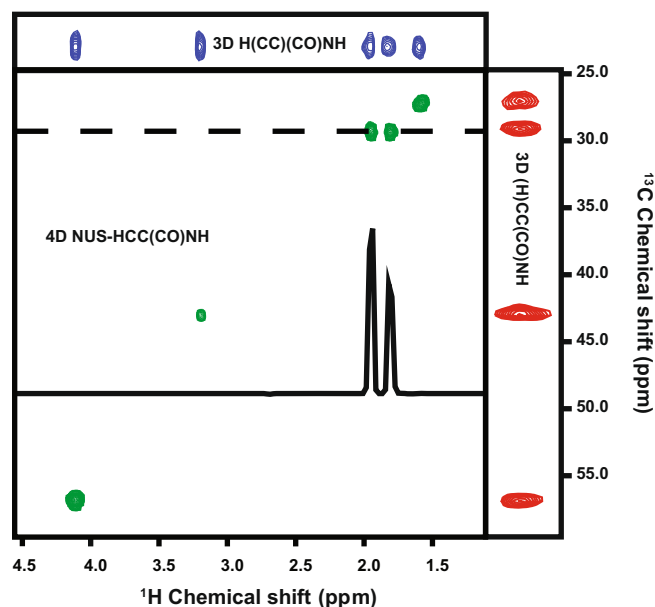


Fig. 3. Comparison of 2D planes from 3D (H)CC(CO)NH-TOCSY and H(CC)(CO)NH-TOCSY experiments and the 4D HCC(CO)NH-TOCSY experiment. Experiments were acquired using a cryogenically cooled probe on a Bruker Avance 900 MHz spectrometer using a 300 μM sample of the 4.5-kDa toxin PcTx1. The 3D experiments were sampled and processed using traditional methods. The 4D experiment was acquired using non-uniform sampling and processed using MaxEnt reconstruction; each experiment was acquired in ~ 1.5 days. All planes depict the same ^1H and ^{15}N backbone-amide frequencies. The axes of the 3D experiments are the directly-detected ^1H dimension and either the ^1H or ^{13}C indirectly-detected dimension. In the 4D experiment the axes are the indirect ^1H and ^{13}C dimensions. The 1D trace shown in the 4D spectral plane is taken at the row indicated by the dashed line. All panels show sidechain resonances from residue Arg29.

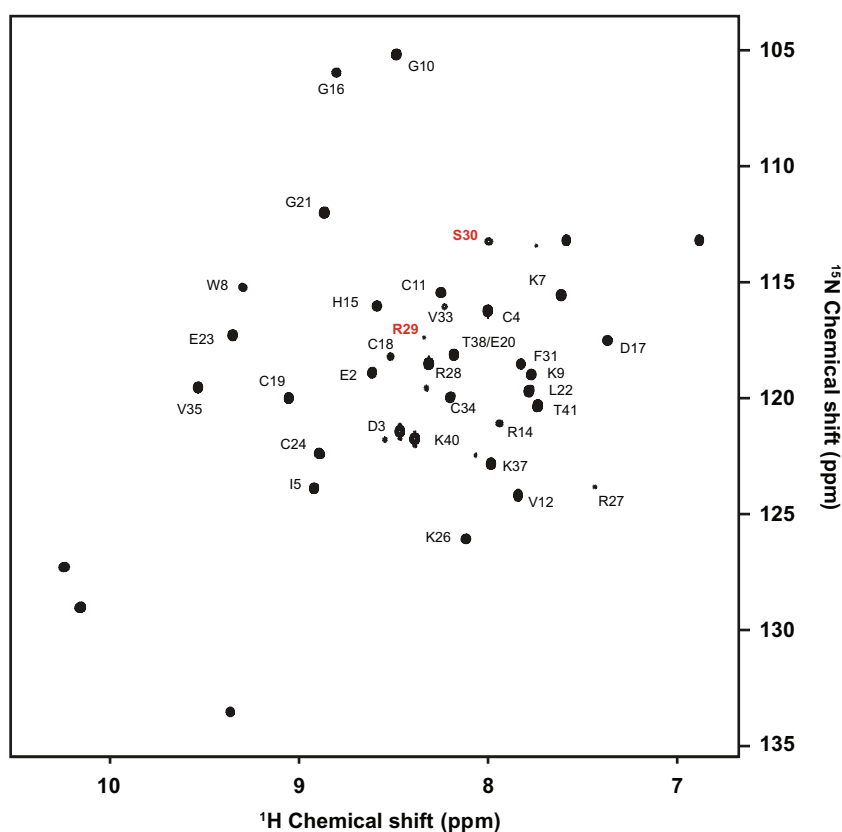


Fig. 4. Annotated 2D ^1H – ^{15}N HSQC spectrum of PcTx1. Note the low S/N of the Arg29/Ser30 resonance (^1H and ^{15}N chemical shifts of 7.99 and 113.0 ppm, respectively).

dimensions. The 4D experiment was collected using 8 scans per increment and processed using the new 3D MaxEnt algorithm (msa3d) to yield a digital resolution of $64 \times 256 \times 256$ in the ^{15}N , ^{13}C , and ^1H dimensions, respectively. Each experiment (3Ds and 4D) was acquired in ~ 1.5 days.

Complete ^1H , ^{15}N , and ^{13}C chemical shifts for PcTx1 have been deposited in BioMagResBank (Accession No. 16468), while coordinates for both low-resolution and high-resolution structures are available from the Protein Data Bank (coordinate files 1LMM and 2KNI, respectively).

Acknowledgments

Supported by grants from the Australian Research Council (DP0774245 to G.F.K.), the US National Institutes of Health (GM-47467 and RR-020125 to J.C.H.), and the Queensland Smart State Research Facilities Fund. The authors thank Natalie Saez for the sample of PcTx1.

References

- [1] R.R. Ernst, W.A. Anderson, Application of Fourier transform spectroscopy to magnetic resonance, *Rev. Sci. Instrum.* 37 (1966) 93–102.
- [2] M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, *Prog. Nucl. Magn. Reson. Spectrosc.* 34 (1999) 93–158.
- [3] G.T. Montelione, B.A. Lyons, S.D. Emerson, M. Tashiro, An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins, *J. Am. Chem. Soc.* 114 (1992) 10974–10975.
- [4] R.T. Clowes, W. Boucher, C.H. Hardman, P.J. Domaille, E.D. Laue, A 4D HCC(CO)NNH experiment for the correlation of aliphatic side-chain and backbone resonances in $^{13}\text{C}/^{15}\text{N}$ -labelled proteins, *J. Biomol. NMR* 3 (1993) 349–354.
- [5] V.A. Jaravine, A.V. Zhuravleva, P. Permi, I. Ibraghimov, V.Y. Orekhov, Hyperdimensional NMR spectroscopy with nonlinear sampling, *J. Am. Chem. Soc.* 130 (2008) 3927–3936.
- [6] K. Kazimierczuk, A. Zawadzka, W. Kozminski, I. Zhukov, Determination of spin–spin couplings from ultrahigh resolution 3D NMR spectra obtained by optimized random sampling and multidimensional fourier transformation, *J. Am. Chem. Soc.* 130 (2008) 5404–5405.
- [7] D.A. Snyder, Y. Xu, D. Yang, R. Bruschweiler, Resolution-enhanced 4D $^{15}\text{N}/^{13}\text{C}$ NOESY protein NMR spectroscopy by application of the covariance transform, *J. Am. Chem. Soc.* 129 (2007) 14126–14127.
- [8] B. Coggins, P. Zhou, High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN, *J. Biomol. NMR* 42 (2008) 225–239.
- [9] A. Zawadzka-Kazimierczuk, K. Kazimierczuk, W. Kozminski, A set of 4D NMR experiments of enhanced resolution for easy resonance assignment in proteins, *J. Magn. Reson.* 202 (2010) 109–116.
- [10] S. Hiller, I. Ibraghimov, G. Wagner, V.Y. Orekhov, Coupled decomposition of four-dimensional NOESY spectra, *J. Am. Chem. Soc.* 131 (2009) 12970–12978.
- [11] S. Sibisi, Two-dimensional reconstructions from one-dimensional data by maximum entropy, *Nature* 301 (1983) 134–136.
- [12] J.C.J. Barna, S.M. Tan, E.D. Laue, Use of CLEAN in conjunction with selective data sampling for 2D NMR experiments, *J. Magn. Reson.* 78 (1988) 327–332.
- [13] S. Hyberts, D. Frueh, H. Arthanari, G. Wagner, FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation, *J. Biomol. NMR* 45 (2009) 283–294.
- [14] J.C. Hoch, A.S. Stern, and V.D.U.S. Thomas L. James, Maximum entropy reconstruction, spectrum analysis and deconvolution in multidimensional nuclear magnetic resonance, *Methods in Enzymology*, Academic Press, 2002, pp. 159–178.
- [15] J.C. Hoch, A.S. Stern, *NMR Data Processing*, Wiley-Liss, New York, 1996.
- [16] S. Hiller, R. Joss, G. Wider, Automated NMR assignment of protein side chain resonances using automated projection spectroscopy (APSY), *J. Am. Chem. Soc.* 130 (2008) 12073–12079.
- [17] M. Mobli, A.S. Stern, J.C. Hoch, Spectral reconstruction methods in fast NMR: reduced dimensionality, random sampling and maximum entropy, *J. Magn. Reson.* 182 (2006) 96–105.
- [18] D. Malmodin, M. Billeter, Multiway decomposition of NMR spectra with coupled evolution periods, *J. Am. Chem. Soc.* 127 (2005) 13486–13487.
- [19] K. Kazimierczuk, A. Zawadzka, W. Komiski, Narrow peaks and high dimensionalities: exploiting the advantages of random sampling, *J. Magn. Reson.* 197 (2009) 219–228.
- [20] M. Mobli, J.C. Hoch, Maximum entropy spectral reconstruction of nonuniformly sampled data, *Concepts Magn. Reson. A* 32 (2008) 436–448.

- [21] P.S. Efrimidis, P.G. Spirakis, Weighted random sampling with a reservoir, *Inform. Process. Lett.* 97 (2006) 181–185.
- [22] M. Mobli, M.W. Maciejewski, M.R. Gryk, J.C. Hoch, An automated tool for maximum entropy reconstruction of biomolecular NMR spectra, *Nature Methods* (2007).
- [23] M. Mobli, M. Maciejewski, M. Gryk, J. Hoch, Automatic maximum entropy spectral reconstruction in NMR, *J. Biomol. NMR* 39 (2007) 133–139.
- [24] A. Bahrami, A.H. Assadi, J.L. Markley, H.R. Eghbalnia, Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy, *PLoS Comp. Biol.* 5 (2009) e1000307.
- [25] A.S. Stern, K.B. Li, J.C. Hoch, Modern spectrum analysis in multidimensional NMR spectroscopy: comparison of linear-prediction extrapolation and maximum-entropy reconstruction, *J. Am. Chem. Soc.* 124 (2002) 1982–1993.
- [26] D. Rovnyak, D.P. Frueh, M. Sastry, Z.J. Sun, A.S. Stern, J.C. Hoch, G. Wagner, Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy, *J. Magn. Reson.* 170 (2004) 15–21.
- [27] L. Emsley, G. Bodenhausen, Optimization of shaped selective pulses for NMR using a quaternion description of their overall propagators, *J. Magn. Reson.* 97 (1992) 135–148.
- [28] J. Boyd, N. Soffe, Selective excitation by pulse shaping combined with phase modulation, *J. Magn. Reson.* 85 (1989) 406–413.
- [29] A.J. Shaka, C.J. Lee, A. Pines, Iterative schemes for bilinear operators; application to spin decoupling, *J. Magn. Reson.* 77 (1988) 274–293.
- [30] R. Freeman, E. Kupce, Decoupling: theory and practice I. Current methods and recent concepts, *NMR Biomed.* 10 (1997) 372–380.
- [31] T.M. Logan, E.T. Olejniczak, R.X. Xu, S.W. Fesik, A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments, *J. Biomol. NMR* 3 (1993) 225–231.
- [32] S. Grzesiek, A. Bax, Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins, *J. Biomol. NMR* 3 (1993) 185–204.
- [33] S. Grzesiek, J. Anglister, A. Bax, Correlation of backbone amide and aliphatic side-chain resonances in $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of ^{13}C magnetization, *J. Magn. Reson. B* 101 (1993) 114–119.
- [34] D. Marion, M. Ikura, R. Tschudin, A. Bax, Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins, *J. Magn. Reson.* 85 (1989) 393–399.
- [35] V. Sklenar, M. Piotto, R. Leppik, V. Saudek, Gradient-tailored water suppression for ^1H – ^{15}N HSQC experiments optimized to retain full sensitivity, *J. Magn. Reson. A* 102 (1993) 241–245.
- [36] M. Piotto, V. Saudek, V. Sklenář, Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR* 2 (1992) 661–665.