Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 165 (2003) 315–319
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Communication

High-throughput backbone resonance assignment of small ^{13}C, ^{15}N-labeled proteins by a triple-resonance experiment with four sequential connectivity pathways using chemical shift-dependent, apparent $^{1}J(^{1}H,^{13}C)$: HNCACB^{coded}HAHB

Scott Pegan, Witek Kwiatkowski, Senyon Choe, and Roland Riek*

Structural Biology Laboratory, The Salk Institute, La Jolla, CA 92037, USA

Received 28 March 2003; revised 25 July 2003

Abstract

The proposed three-dimensional triple-resonance experiment HNCACB^{coded}HAHB correlates sequential ^{15}N, ^{1}H moieties via the chemical shifts of $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$, and $^{1}H^{\beta}$. The four sequential correlation pathways are achieved by the incorporation of the concept of chemical shift-coding [J. Biomol. NMR 25 (2003) 281] to the TROSY-HNCACB experiment. The monitored $^{1}H^{\alpha}$ and $^{1}H^{\beta}$ chemical shifts are then coded in the line shape of the cross-peaks of $^{13}C^{\alpha}$, $^{13}C^{\beta}$ along the ^{13}C dimension through an apparent residual scalar coupling, the size of which depends on the attached hydrogen chemical shift. The information of four sequential correlation pathways enables a rapid backbone assignment. The HNCACB^{coded}HAHB experiment was applied to ~85% labeled ^{13}C,^{15}N-labeled amino-terminal fragment of Vaccinia virus DNA topoisomerase I comprising residues 1–77. After one day of measurement on a Bruker Avance 700 MHz spectrometer and 8 h of manual analysis of the spectrum 93% of the backbone assignment was achieved.

Keywords: Multi-dimensional experiment; NMR; Protein; Chemical shift-coded experiment; High-throughput resonance assignment

Rapid chemical shift assignment of proteins combined with a reduction of NMR measuring time are instrumental for high-throughput structure activity relationship by NMR [1] and structural genomics [2]. In the so-called ''non-sensitivity-limited'' data collection regime [3],[1] the combination of both requirements can be fulfilled by limiting the number of dimensions to less than 4, but increasing the number of (sequential) correlations per data set. One approach to achieve this is the concept of reduced dimensionality introduced by Szyperski et al. [4] and successively expanded to a variety of experiments [3,5] including G-matrix Fourier transform NMR spectroscopy [6]. Reduced dimensionality NMR is based on simultaneous evolutions of two or more different chemical shifts reducing a $n+1$-dimensional triple-resonance experiment to n dimensions. Indeed, the measurement of five 3D reduced dimensionality triple-resonance experiments and one 3D HNCACB experiment with a total measuring time of 42.5 h enabled an automatic backbone assignment (including β-carbon and β-hydrogens) of a 71 residue ^{13}C,^{15}N-labeled protein (concentration 1 mM, pH 6.5, 298 K; [3]). An additional approach to increase the quality of correlation in triple-resonance experiments was proposed by Zweckstetter and Bax [7] using splitting of cross-peaks by residual dipolar coupling.

We proposed an alternative approach to increase the number of correlations without the increase of the number of dimensions [8]. The strength of the so-called concept of chemical shift-coding is the ease of implementation and the addition of chemical shift information without major losses of signal, since no additional

* Corresponding author. Fax: 1-858-452-3683.
E-mail address: riek@sbl.salk.edu (R. Riek).

[1] In the ''non-sensitivity-limited'' data collection regime the sampling of the multi-dimensional NMR experiment determines the minimally achievable measurement time and not the sensitivity of the experiment.

1090-7807/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/j.jmr.2003.08.012

chemical shift evolution and polarization transfers are required, whereas both are required in the reduced dimensionality experiments [3]. In return, the resolution of the additional chemical shifts is limited [8]. The minimal sensitivity loss enables the incorporation of the concept of chemical shift-coding to the TROSY-HNCACB experiment [9–13], which is low in sensitivity when compared with the corresponding HNCA[2] experiment [14]. As a result, this implementation enhances the sequential correlations between ^{15}N,^{1}H moieties from two ($^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts) to four ($^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$, $^{1}H^{\beta}$ chemical shifts). By employing this method, we demonstrate four sequential correlation pathways enables a high-throughput manual backbone assignment for a small ^{13}C,^{15}N-labeled protein.

Technically, the proposed HNCACB^{coded}HAHB experiment is essentially based on the 3D TROSY-HNCACB experiment, but obtains additional information of the chemical shift of $^{1}H^{\alpha}$ and $^{1}H^{\beta}$ through the concept of chemical shift-coding [8]. This pulse sequence is depicted in the experimental scheme of Fig. 1. Between time points *a* and *c*, magnetization is transferred from ^{1}H via ^{15}N to $^{13}C^{\alpha/\beta}$, using three successive INEPT steps [15,16]. During the frequency labeling period t_2 the $^{13}C^{\alpha/\beta}$ chemical shifts evolve. Simultaneously, the $^{1}J(^{13}C,^{1}H)$ scalar coupling of ~140–155 Hz begins to evolve. However, the apparent extent of the evolution of the $^{1}J(^{13}C,^{1}H)$ coupling is under the control of the 180° soft pulse on $^{1}H^{\alpha/\beta}$. The soft pulse is on-resonance at 3.5 ppm and contains a Gaussian inversion profile, which decreases to ~5% at 2 and 5 ppm. Thus, the multiplet splitting of the cross-peak along the ^{13}C dimension and concomitantly the extent of the apparent $^{1}J(^{13}C,^{1}H)$ coupling depend on the position of the attached $^{1}H^{\alpha/\beta}$ chemical shifts. In other words, the $^{1}H^{\alpha/\beta}$ chemical shifts are encoded in the pattern of the cross-peak along ^{13}C dimension through an apparent residual scalar coupling $^{residual}J(^{13}C,^{1}H)$ [8]. After chemical shift evolution modulated by the chemical shift-dependent scalar coupling, magnetization is transferred back to ^{15}N at time point *d*. Between *d* and *e* the ^{15}N nuclei evolve with their chemical shifts during a constant time period. From time point *e* onward magnetization is transferred back to ^{1}H. The flow coherence can thus be described as follows:

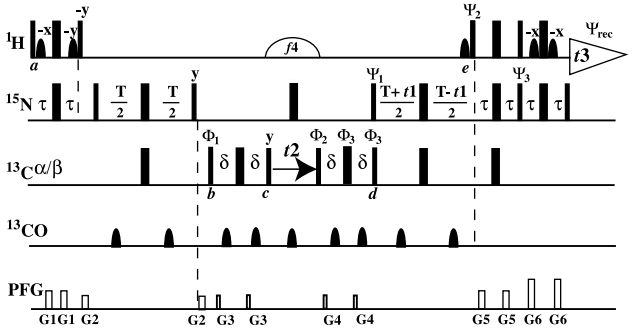
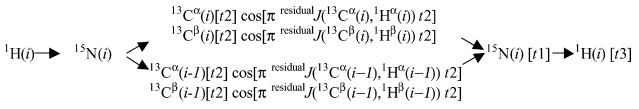


Fig. 1. Experimental scheme for the HNCACB^{coded}HAHB experiment for ^{13}C,^{15}N-labeled proteins. The radiofrequency pulses on ^{1}H, ^{15}N, $^{13}C^{\alpha}$ or $^{13}C^{\beta}$ were applied at 4.8, 119, 45, and 174 ppm, respectively. The narrow and wide black bars indicate non-selective 90° and 180° pulses, respectively. On the line marked ^{1}H a large sine bell shaped pulse labeled with *f4* indicates a selective 180° pulse with a Gaussian shape truncated at 5% and the duration of 1.3 ms. The soft pulse is on-resonance at 3.5 ppm and its length has been optimized for the magnetic field strength of 700 MHz ^{1}H frequency. The inversion profile is described in Kwiatkowski and Riek [8]. On the line marked ^{13}C, black sine bell shapes indicate selective 180° pulses with the duration of 0.08 ms and a Gaussian shape truncated at 5%. The line marked PFG indicates durations and amplitudes of sine shaped pulse magnetic field gradients applied along the *z*-axis: G1: 0.8 ms, 15 G/cm; G2: 0.8 ms, 9 G/cm; G3: 0.5 ms, 18 G/cm; G4: 0.5 ms, 18 G/cm; G5: 0.8 ms, 12 G/cm; and G4: 0.8 ms, 22 G/cm. The delays τ, T, δ are 2.7, 24, and 3.6 ms, respectively. The phase cycle is $\Psi_1 = \{y,-y,x,-x\}$, $\Psi_2 = \{-y\}$, $\Psi_3 = \{-y\}$, $\Phi_1 = \{x,x,x,x,-x,-x,-x,-x\}$, $\Phi_2 = \{-y\}$, $\Phi_3 = \{x\}$, and $\Phi_{rec} = \{y,-y,-x,x,-y,y,x,-x\}$. All other radiofrequency pulses are applied either with phase *x* or as indicated above the pulses. In the ^{15}N($t1$) dimension a phase-sensitive spectrum is obtained by recording a second FID for each increment of $t1$, with $\Psi_1 = \{y,-y,-x,x\}$, $\Psi_2 = \{y\}$, and $\Psi_3 = \{y\}$ and the data are processed as described by Kay et al. [20]. Quadrature detection in the $^{13}C^{\alpha}(t2)$ dimension is achieved by the States-TPPI method [21] applied to phases Φ_1, Φ_2, and Φ_3. Between time points *c* and *d* the timing of the radiofrequency pulses of the different nuclei have been implemented in a parallel manner to achieve an initial $t_2 = 0$. On the line marked ^{1}H, black sine bell shapes indicate water-selective 90° pulses with the duration of 1 ms and a Gaussian shape truncated at 5%. With these pulses, the water magnetization stays aligned along the $+z$-axis throughout the experiment [22]. Alternatively, we replaced the water-selective pulses with a superposition of two Gaussian pulses (truncated at 5% and length 1 ms) on resonance at the water-frequency and at 1.7 ppm, respectively. This procedure aligns the water magnetization and part of the aliphatic hydrogens along the $+z$-axis throughout the experiment yielding longitudinal relaxation-optimized spectroscopy [23]. The measured gain in signal-to-noise was about 10% (data not shown). In our hands, the performance of the superposition of two Gaussian pulses is superior to the E-BURP pulse proposed by Pervushin et al. [23], since the extent of the water-flip back can be controlled better and an incomplete alignment of the water magnetization along the $+z$-axis results in severe losses in signal-to-noise when applied to our high pH sample (pH 7). The soft pulse *f4* is chosen to perturb the water resonance frequency minimally.



where $t1$ and $t2$ are the ^{15}N and ^{13}C evolution times and $t3$ is the ^{1}H acquisition time. Indices i and $i-1$ indicate

[2] *Abbreviations used:* HNCA, amide proton-to-nitrogen-to-α-carbon correlation; HNCACB^{coded}HAHB, HNCACB with the additional frequency of α- and β-hydrogens coded in the multiplet pattern of the cross-peaks along the carbon frequency; INEPT, insensitive nuclei enhanced by polarization transfer; NMR, nuclear magnetic resonance; PFG, pulsed field gradient; TOPO(1–77), residues 1–77 of Vaccinia virus DNA topoisomerase I; TROSY, transverse relaxation-optimized spectroscopy; 2D, two-dimensional; 3D, three-dimensional; 4D, four-dimensional.

that coherence is transferred from $^{15}N(i)$–$^{1}H(i)$ to both the sequentially and the intra-residually adjoining carbons. The coded ^{1}H chemical shift is designated with the cosine modulation of its apparent residual scalar coupling, $^{residual}J(^{13}C^{\alpha},^{1}H^{\alpha})$ or $^{residual}J(^{13}C^{\beta},^{1}H^{\beta})$ active during $t2$.

Fig. 2 shows strips from a 3D HNCACB^{coded}HAHB spectrum of uniformly 85% ^{13}C,^{15}N-labeled N-terminal domain of Vaccinia virus DNA topoisomerase I comprising residues 1–77 [17]. In addition, the protein samples contains a six residues-long residual histidine-tag with the sequence GSHGGS. The relatively low incorporation of stable isotopes is due to the cost-effective, efficient, and rapid approach of protein preparation that was used [18]. As in the conventional HNCACB experiment, two positive cross-peaks for the sequential and intra-residual $^{13}C^{\alpha}$s and two negative cross-peaks for the sequential and intra-residual $^{13}C^{\beta}$s are observed for each spin-system. Usually, the stronger cross-peaks correspond to intra-residue correlations between the carbons and the ^{15}N–^{1}H moiety, and the weaker cross-peak represents the sequential correlations. These cross-peaks are split due to the active $^{residual}J(^{13}C,^{1}H)$-coupling, the size of which is related to the attached hydrogen chemical shift. The cross-peak is not split if the attached $^{1}H^{\alpha}$ chemical shift is 3.5 ppm, whereas the cross-peak is a doublet of 144 Hz if the corresponding $^{1}H^{\alpha}$ chemical shift is $\geqslant$ 5.3 ppm. An apparent residual scalar coupling of 0–144 Hz corresponds to a $^{1}H^{\alpha}$ chemical shift between 3.5 and 5.3 ppm with a resolution of ~0.2 ppm [8]. Similarly, the $^{13}C^{\beta}$-cross-peaks are split due to the $^{1}H^{\beta}$ chemical shifts. Thus, sequential correlations can be traced with four probes: the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$, and $^{1}H^{\beta}$ chemical shifts. Although the resolution of the coded chemical shifts in the line shape of the cross-peak is low, it enhances the sequential correlation quality by more than one order in magnitude [8]. Indeed, with the HNCACB^{coded}HAHB experiment a straightforward sequential walk can be traced through the multiplets of the ^{13}C cross-peaks including the overall position of the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ cross-peaks as a point of departure (Fig. 2). Only when the multiplet pattern of the sequential $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ cross-peaks coincides with the multiplet of the intra-residual cross-peaks, the sequential connectivity can be traced. When the patterns do not fit, the two involved ^{15}N–^{1}H moieties are not neighbors. With this procedure the sequential assignment is ensured by four independent correlations, the α-carbon and β-carbon chemical shifts and indirectly the α-hydrogen and β-hydrogen chemical

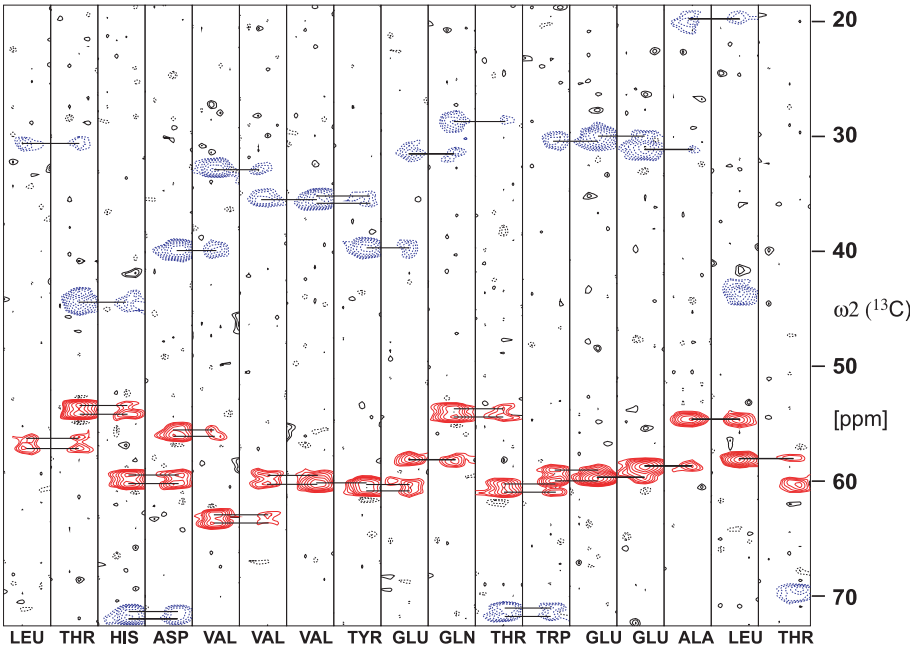


Fig. 2. Strips along the $\omega_2(^{13}C)$ dimension of the 3D HNCACB^{coded}HAHB spectrum of ^{13}C,^{15}N-labeled TOPO(1–77). The sequence-specific backbone assignment of residues 39–55, which corresponds to 20% of the protein, are shown. The strips are centered about the corresponding $^{1}H^{N}$ chemical shifts. The sequential connectivities are indicated by horizontal lines. Positive contour lines are red, negative contour lines blue, respectively. The experiment was recorded at 20 °C with a 0.8 mM sample of ^{13}C,^{15}N-labeled TOPO(1–77) in a mixed solvent of 95% H_2O/5% D_2O at pH 7, using a Bruker Avance 700 MHz spectrometer equipped with five radiofrequency channels, a pulsed field gradient unit, and a triple-resonance probe with an actively shielded z-gradient coil. The following parameter settings were used: data size = $50(t1) \times 85(t2) \times 1024(t3)$ complex points; $t1_{max}(^{15}N) = 23$ ms, $t2_{max}(^{13}C^{\alpha}) = 7.5$ ms, and $t3_{max}(^{1}H) = 100$ ms. The data set was zero-filled to $128 \times 1024 \times 2048$ complex points; eight scans per increment were acquired, resulting in 22 h of measuring time. Prior to Fourier transformation the data were multiplied with a 75° shifted sine bell window in all dimensions.

shifts by the residual scalar couplings. Importantly, both the α-carbon and the β-carbon chemical shifts are measured, which is relevant for the mapping of connected spin-systems onto the amino acid sequence. Since this high quality of information generates practically no ambiguities, sequential assignment is also possible in rare cases where the sequential and intra-residual cross-peaks have similar intensities and cannot be classified straightforwardly. In the conventional experiments additional spectra, i.e., HN(CO)CA or CBCA(CO)NH, are collected to classify the cross-peak into a sequential or intra-residual correlation.

The HNCACB^{coded}HAHB experiment of Fig. 2 was measured in 24 h on our Bruker Avance 700 MHz spectrometer, which corresponds to about 3 h measurement time with a cryoprobe. Manual analysis of the spectrum during one 8 h working day yielded 93% of the backbone assignment of 85%-labeled ^{13}C,^{15}N-labeled TOPO(1–77). The remaining five ^{15}N–^{1}H moieties were not assigned due to the lack of sequential correlations in the HNCACB^{coded}HAHB experiment. The absence of sequential peaks is probably because of local relaxation properties attributed to slow conformational exchange. In combination with a ^{13}C,^{15}N-resolved [^{1}H,^{1}H]NOESY experiment, most of these residual moieties were assigned. Alternatively, a HN(CO)CA experiment [9] with inherently higher sensitivity for sequential cross-peaks than the HNCACB^{coded}HAHB would give complementary information.

As demonstrated with TOPO(1–77) (Fig. 2 shows 20% of the sequential backbone assignment) the proposed HNCACB^{coded}HAHB is a very powerful experiment in the "non-sensitivity-limited" data collection regime and therefore highly recommended for small ^{13}C,^{15}N-labeled proteins. We believe that the HNCACB^{coded}HAHB experiment is superior to the published reduced dimensionality experiments due to the following reasons: (i) The lack of additional polarization transfers and evolution times makes the coded experiment more sensitive. We find a factor of three signal loss due to $^{13}C \rightarrow {}^{1}H$ polarization transfer, ^{1}H evolution, and back transfer for a protein with a rotational correlation time τ_c of 4 ns. (ii) The presence of four sequential correlations reduces ambiguities dramatically. (iii) Due to the small splitting of the cross-peaks, the peak pattern recognition for manual assignment is simple. (iv) The experiment is easily set up. (v) The simultaneous detection of α-carbon and β-carbon are important for mapping connected spin-systems onto the amino acid sequence. (vi) Only one experiment is measured.

For a 20 kDa ^{13}C,^{15}N-labeled protein the HNCACB^{coded}HAHB experiment is a factor of 1.5–2 less sensitive than a conventional TROSY-HNCACB experiment, which itself is a factor of 10 less sensitive than a TROSY-HNCA experiment [14]. Therefore, for large target systems we propose the use of the highly sensitive HNCA^{coded}CB and HNCA^{coded}CO experiments combined with ~70% deuteration and uniform ^{13}C and ^{15}N-labeling [19]. In conclusion, the proposed HNCACB^{coded}HAHB reduces dramatically ambiguities in linking the spin-systems of adjacent residues in the protein sequence during the sequential assignment process and, therefore, allows the fast manual backbone assignment of small ^{13}C,^{15}N-labeled proteins.

Acknowledgments

We thank the H. and J. Weinberg Foundation, the H.N. and F. C. Berger Foundation, the Auen Foundation (R.R.), and American Heart Association (S.C.) for financial support.

References

[1] E.R. Zuiderweg, Mapping protein–protein interactions in solution by NMR spectroscopy, Biochemistry 41 (2002) 1–7.
[2] G.T. Montelione, D. Zheng, Y.J. Huang, K.C. Gunsalus, T. Szyperski, Protein NMR spectroscopy in structural genomics, Nat. Struct. Biol. 7 (2000) 982–985.
[3] T. Szyperski, D.C. Yeh, D.K. Sukumaran, H.N.B. Moseley, G.T. Montelione, Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment, Proc. Natl. Acad. Sci. USA 99 (2002) 8009–8014.
[4] T. Szyperski, G. Wider, J. Bushweller, K. Wüthrich, Reduced dimensionality in triple-resonance NMR experiments, J. Am. Chem. Soc. 115 (1993) 9307–9308.
[5] J. Ding, A. Gronenborn, Novel 2D triple-resonance NMR experiments for sequential resonance assignments of proteins, J. Magn. Reson. 156 (2002) 262–268.
[6] S. Kim, T. Szperski, GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information, J. Am. Chem. Soc. 125 (2003) 1385–1393.
[7] M. Zweckstetter, A. Bax, Single-step determination of protein substructures using dipolar couplings: aid to structural genomics, J. Am. Chem. Soc. 123 (2001) 9490–9491.
[8] W. Kwiatkowski, R. Riek, Chemical shift-dependent apparent scalar couplings: an alternative concept of chemical shift monitoring in multi-dimensional NMR experiments, J. Biomol. NMR 25 (2003) 281–290.
[9] A. Bax, S. Grzesiek, Methodological advances in protein NMR, Acc. Chem. Res. 26 (1993) 131–138.
[10] M. Witekind, L. Muller, HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha and beta carbon resonances in proteins, J. Magn. Reson. B 101 (1993) 201–205.
[11] R. Riek, K. Pervushin, K. Wüthrich, TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution, Trends Biochem. Sci. 25 (2000) 462–468.
[12] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, Proc. Natl. Acad. Sci. USA 94 (1997) 12366–12371.
[13] M. Salzmann, G. Wider, K. Pervushin, H. Senn, K. Wuthrich, [^{13}C]-constant-time [^{15}N,^{1}H]-TROSY-HNCA for sequential

assignments of large proteins, J. Am. Chem. Soc. 121 (1999) 844–848.
[14] 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.
[15] G.A. Morris, R. Freeman, Enhancement of NMR signals by polarization transfer, J. Am. Chem. Soc. 101 (1979) 760–762.
[16] G. Wider, Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution, Prog. NMR Spectrosc. 32 (1998) 193–275.
[17] A. Sharma, R. Hanai, A. Mondragon, Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution, Structure 2 (1994) 767–777.
[18] J. Marley, M. Lu, C. Bracken, A method for efficient isotopic labeling of recombinant proteins, J. Biomol. NMR 20 (2001) 70–75.
[19] C. Ritter, T. Luhrs, W. Kwiatkowski, R. Riek, 3D TROSY-HNCA^{coded}CB and TROSY-HNCA^{coded}CO experiments: two triple resonance NMR experiments with two sequential connectivity pathways and high sensitivity, J. Biomol. NMR, 2003, accepted.
[20] L.E. Kay, P. Keifer, T. Saarinen, Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, J. Am. Chem. Soc. 114 (1992) 10663–10665.
[21] 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.
[22] S. Grzesiek, A. Bax, The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements, J. Am. Chem. Soc. 115 (1993) 12593–12594.
[23] K. Pervushin, B. Vogeli, A. Eletsky, Longitudinal (1)H relaxation in TROSY NMR spectroscopy, J. Am. Chem. Soc. 124 (2002) 12898–12902.