J Biomol NMR (2010) 47:55–63
DOI 10.1007/s10858-010-9410-3

ARTICLE

CACA-TOCSY with alternate ^{13}C–^{12}C labeling: a $^{13}C^{\alpha}$ direct detection experiment for mainchain resonance assignment, dihedral angle information, and amino acid type identification

Koh Takeuchi · Dominique P. Frueh · Zhen-Yu J. Sun · Sebastian Hiller · Gerhard Wagner

Received: 19 February 2010 / Accepted: 19 March 2010 / Published online: 10 April 2010

Abstract We present a ^{13}C direct detection CACA-TOCSY experiment for samples with alternate ^{13}C–^{12}C labeling. It provides inter-residue correlations between $^{13}C^{\alpha}$ resonances of residue i and adjacent C^{α}s at positions $i - 1$ and $i + 1$. Furthermore, longer mixing times yield correlations to C^{α} nuclei separated by more than one residue. The experiment also provides C^{α}-to-sidechain correlations, some amino acid type identifications and estimates for ψ dihedral angles. The power of the experiment derives from the alternate ^{13}C–^{12}C labeling with [1,3-^{13}C] glycerol or [2-^{13}C] glycerol, which allows utilizing the small scalar $^{3}J_{CC}$ couplings that are masked by strong $^{1}J_{CC}$ couplings in uniformly ^{13}C labeled samples.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-010-9410-3) contains supplementary material, which is available to authorized users.

K. Takeuchi · D. P. Frueh · Z.-Y. J. Sun · S. Hiller · G. Wagner (✉)
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA
e-mail: gerhard_wagner@hms.harvard.edu

Present Address:
K. Takeuchi
Biomedicinal Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), Aomi 2-4-7, Koto-ku, Tokyo 135-0064, Japan

D. P. Frueh
Department of Biophysics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

S. Hiller
Laboratorium für Physikalische Chemie, ETH Zürich, Zurich 8093, Switzerland

Keywords Alternate ^{13}C labeling · C^{α} direct detection · TOCSY · Nuclear magnetic resonance (NMR) · Resonance assignment · Dihedral angle

Introduction

The motivation for using ^{13}C detection in protein NMR is to take advantage of the slower transverse relaxation rates of ^{13}C magnetization compared to ^{1}H. Carbon detection may extend the limits of NMR in structural and functional analyses of larger, highly dynamic, or paramagnetic proteins. According to this idea, a variety of pulse sequences have recently been developed to perform main-chain resonance assignment and structural analyses of proteins (Serber et al. 2001; Bermel et al. 2003; Arnesano et al. 2005; Lee et al. 2005; Bermel et al. 2006a, b; Takeuchi et al. 2008b; Felli and Brutscher 2009; Hsu et al. 2009; Takeuchi et al. 2010). The majority of these pulse programs use carbonyl carbons (C') as the detected nuclei. This is mainly due to the simple spin–spin coupling pattern of C' that has only one strong scalar coupling to the alpha carbon (C^{α}) (Bermel et al. 2006b). However, in large-molecular-weight systems, particularly at higher magnetic fields, carbonyl carbons suffer from fast transverse relaxation due to their large chemical shift anisotropy (CSA). In contrast, C^{α} nuclei have a small CSA and consequently slow transverse relaxation if samples are deuterated. Thus, deuterated C^{α} would be the better nuclei for ^{13}C detection experiment in large and/or fast relaxing systems. Direct detection of $^{13}C^{\alpha}$ is not straightforward in uniformly ^{13}C-labeled proteins owing to the presence of strong ^{13}C–^{13}C couplings with $^{13}C'$ ($\sim$55 Hz) and $^{13}C^{\beta}$ ($\sim$35 Hz) that split the signals into multiplets and hence reduce the sensitivity. The spectral complexity can be

overcome by computational deconvolution (Shimba et al. 2003) or by spin-state selective schemes such as using IPAP or S3E (Bermel et al. 2007). However, these schemes require more complicated pulse programs and/or processing. Recently, we developed a strategy using ^{13}C enrichment at alternating carbon sites and deuteration at the C^{α} position to avoid these difficulties (Takeuchi et al. 2008b). This can be achieved by using either [2-^{13}C]- or [1,3-^{13}C]-glycerol (or pyruvate) as the carbon source while growing the *E. coli* bacteria expressing the target protein in D_2O/^{15}N containing media (LeMaster and Kushlan 1996; Castellani et al. 2002; Takeuchi et al. 2008b; Guo et al. 2009) (Hereafter the terms 2-^{13}C and 1,3-^{13}C labeling are used for alternate ^{13}C labeling using [2-^{13}C]- or [1,3-^{13}C]-glycerol, respectively). This approach offers advantages for the detection of $^{13}C^{\alpha}$ NMR signals, while maintaining heteronuclear ^{13}C–^{15}N couplings for establishing complete backbone sequential assignments (Takeuchi et al. 2008b, 2010).

Here we show that alternate ^{13}C–^{12}C labeling enables direct correlations between C^{α} nuclei in two neighboring amino acids via weak long range ^{3}J coupling ($\sim$2 Hz) (Hennig et al. 2000). These weak couplings are usually masked by the strong ^{1}J carbon–carbon splittings but are readily observed in proteins labeled with the alternate ^{13}C–^{12}C-labeling scheme.

Materials and methods

All chemicals were purchased from Sigma (St. Louis, MO) unless otherwise noted. All stable-isotope-labeled materials were acquired from Cambridge Isotope Laboratories (Cambridge, MA).

Expression and purification of the B domain of protein G (GB1)

The gene for 6His-tagged GB1, consisting of 64 amino acid residues, was cloned into the pET9d vector (Novagen, San Diego, CA) as previously described (Frueh et al. 2005). GB1 was expressed in commercially available BL21 (DE3) *E. coli* cells (Novagen) at 37°C and protein expression was induced for 6 h at the same temperature. For [2-^{13}C-glycerol] (or [1,3-^{13}C-glycerol]) labeled samples, the cells were cultured in ^{2}H, ^{15}N M9 media containing 8.5 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$, and 1 g/L of $^{15}NH_4Cl$ in D_2O, which was supplemented with 2 g/L [2-^{13}C] (or [1,3-^{13}C]) glycerol and 2 g/L $NaH^{13}CO_3$ (or $NaH^{12}CO_3$). The protein was purified with Ni–NTA affinity chromatography as previously described (Frueh et al. 2005).

NMR experiments

NMR spectra were recorded on a Bruker (Billerica, MA) Avance 500 spectrometer equipped with a triple-resonance carbon-cryogenic probe (TXO) designed for carbon detection experiments. Spectra of the 2-^{13}C labeled GB1 sample (5 mM) were recorded at 15°C in buffer containing 10 mM sodium phosphate (pH 6.8), 100 mM NaCl and 20% w/v deuterated glycerol in D_2O. The molecular tumbling of GB1 under those conditions corresponds to a 90 kDa protein at 25°C (Takeuchi et al. 2008a). For the 1,3-^{13}C labeled GB1 sample (2.5 mM), spectra were recorded at 25°C in buffer containing 10 mM sodium phosphate (pH 6.8), 100 mM NaCl in D_2O, without deuterated glycerol. 2D CACA-TOCSY experiments were recorded with a spectral width of 6,887 Hz (direct) × 6,910 Hz (indirect), centered at 43 ppm. CA(N)CA experiments were recorded with a spectral width of 3,531 Hz (direct) × 3,517 Hz (indirect) centered at 55 ppm. 512 and 128 complexed data points were recorded for direct and indirect directions, respectively, in both experiments. Eight scans were applied for each increment. The wider spectrum width for the 2D CACA-TOCSY was chosen to avoid folding of side chain resonances. The digital resolution in the direct and indirect dimensions after Fourier transformation was 13.5 and 54.0 Hz for the 2D CACA-TOCSY and 6.9 and 27.5 Hz for CA(N)CA. The recycling delay was optimized to be 1.25 × T_1 (longitudinal relaxation time) = 4.5 s. All spectra were analyzed with the program Sparky (Goddard and Kneller 2004).

Results and discussion

We show that direct sequential connectivities between C^{α} spins via $^{3}J(C_i^{\alpha} - C_{i+1}^{\alpha})$ coherence transfer can be established by a CACA-TOCSY experiment. Figure 1 depicts the pulse scheme optimized for deuterated and alternately ^{13}C–^{12}C-labeled proteins. Although not necessary for this experiment, our sample was ^{15}N labeled to compare the sensitivity of the CACA-TOCSY experiment with the previously reported CA(N)CA experiment, as we will discuss below. In short, a spin lock, which follows excitation and chemical shift evolution of C^{α}, correlates C_i^{α} with the adjacent C_{i+1}^{α} via ^{3}J couplings between C^{α} spins. A related pulse scheme has been proposed previously for uniformly ^{13}C labeled proteins and was used for obtaining side-chain assignments (Bermel et al. 2005). However, the authors also predicted that a modified version that uses C^{α}-selective pulses throughout, including those for the spin-lock mixing periods, could provide also sequential connectivities. To eliminate carbon–carbon couplings, the experiment also employs a double IPAP module, which

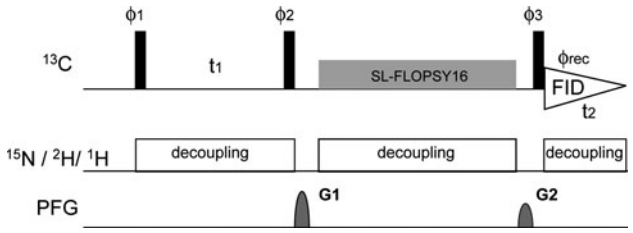

Fig. 1 Pulse scheme of the CACA-TOCSY experiment optimized for uniformly $^{2}H^{15}N$- and alternate ^{13}C–^{12}C -labeled samples. *Black bars* indicate non-selective $\pi/2$ pulses. FLOPSY-16 was applied for spin locking (3.5 kHz). The phase cycle employed was $\phi_1 = (x, -x)$, $\phi_2 = (x, x, -x, -x)$, $\phi_3 = (x, x, x, x, -x, -x, -x, -x)$, and $\phi^{rec} = (x, -x, -x, x, -x, x, x, -x)$. Phase sensitive spectra in the indirect dimension are obtained by incrementing the phases ϕ_1 in a States-TPPI manner (Marion et al. 1989). The recycling delay is optimized based on the estimated longitudinal relaxation time (4.5 s was used in here). The two sine-shaped pulsed field gradients were applied along the z-axis for 1.0 ms with maximum intensities of G1 = 40 G/cm and G2 = 35 G/cm. Broad-band decoupling was achieved by using WALTZ16 (Shaka et al. 1983) for proton and deuterium (3.1 and 1 kHz, respectively) and GARP (Shaka et al. 1985) for nitrogen (1.5 kHz)

adds to the length of the pulse sequence, however. Due to the selective C^{α} pulses and their non-ideal inversion properties, the signals from Gly, Ser, and high-field C^{α} resonances would become significantly weaker. Furthermore, a weak and less homogeneous B_1 field produced by a spin mixing with longer selective pulses would significantly decrease the efficiency of TOCSY transfers. In addition, C^{α} selective pulses do not eliminate evolution under $^{1}J(C^{\alpha}C^{\beta})$ during t_1, which broadens the lines and may even split each signal into a doublet at high resolution. Probably because of these reasons no experimental CACA-TOCSY spectrum has been reported in a literature so far showing sequential correlations. The $^{13}C^{\alpha}$-detected CACA-TOCSY experiment presented here is simple and sensitive. It is designed for proteins expressed to have an alternate ^{13}C–^{12}C labeling pattern using procedures originally described in (LeMaster and Kushlan 1996).

Experiments were recorded on a 5 mM sample of the uniformly $^{2}H^{15}N$ and alternately ^{13}C–^{12}C labeled B1 domain of protein G (GB1) in D_2O. Here, 2-^{13}C glycerol was used as carbon source to maximize the $^{13}C^{\alpha}$ labeling rate (Takeuchi et al. 2008b). To simulate the tumbling of a 90 kDa protein, 20% glycerol was added, and the experiment was recorded at 285 K (for a calibration of the correlation time, see Takeuchi et al. 2008b). All experiments were recorded in D_2O to minimize the long-range proton-carbon dipole interaction. Under these conditions, average C^{α} transverse and longitudinal relaxation times were determined with spin echo and inversion recovery experiments and found to be 90 ms and 3.6 s, respectively. Usually, an effective relaxation rate $1/T_{\mathrm{eff}} = W_t/T_2 + (1 - W_t)/T_1$ is used to estimate the signal decay during the mixing scheme (Felli et al. 2009). The transverse weight W_t depends on the spin-lock scheme, and one wants to achieve a low value of W_t to minimize the decay of coherence during the mixing sequence. Since W_t for the FLOPSY-16 pulse sequence is $\sim$0.65 (Felli et al. 2009), the estimated effective relaxation time at this condition is $\sim$140 ms. The effective relaxation rate during the spin lock can be decreased by using a MOCCA-XY16 mixing sequence which has a smaller W_t value (<0.2) than FLOPSY-16 as shown in a recent publication (Felli et al. 2009). This might be particularly important for large molecular weight systems in order to take advantage of the long T_1 of deuterated $^{13}C^{\alpha}$. However, despite the apparent advantages, the MOCCA-XY16 mixing scheme has the disadvantage that it covers a much narrower bandwidth compared to FLOPSY-16. Thus, the MOCCA mixing scheme may not be an ideal choice, especially for high magnetic fields. Furthermore, as discussed below, the correlations between $^{13}C^{\alpha}$ and side-chain ^{13}C spins would be difficult to observe with the MOCCA-XY16 mixing scheme.

To shorten recycling delays and thus enhancing sensitivity by acquiring more scans one can safely use paramagnetic enhancement reagents, such as Gd(DTPA-BMA) or Ni(DO2A) (Eletsky et al. 2003; Cai et al. 2006; Takeuchi et al. 2010). Fortunately, T_{eff} is rather insensitive to shortening of the longitudinal relaxation time T_1 in particular for FLOPSY-16, which has a larger W_t value but also in the MOCCA-XY16 mixing sequence (Eletsky et al. 2003). Addition of 3 mM Gd(DTPA-BMA) to the "90 kDa" GB1 sample decreased T_1 values from 3.6 to 0.9 s as reported previously (Takeuchi et al. 2010) while the T_2 values are only decreased from 90 to 80 ms. Thus it becomes possible to acquire scans four times faster, with only a 10% decrease in T_{eff} when using FLOPSY-16 (from 140 to 120 ms). When using the MOCCA-XY16 mixing sequence the value of T_{eff} is estimated to decrease about 25% from 400 to 300 ms. Thus, when a moderately narrow mixing band width is acceptable, such as for just alpha carbon correlations, the MOCCA-XY16 sequence may have advantages over FLOPSY-16. Development of more powerful mixing sequences is still desirable.

Figure 2a shows 2D CACA-TOCSY spectra for mixing times ranging from 13 to 514 ms. The measuring time of each spectrum was 5.5 h. At mixing times of 13 and 26 ms, only diagonal signals and a few cross-peaks are observed. Except for several weak C^{α}–C^{α} signals seen in the 26-ms mixing-time spectrum, strong signals corresponding to ^{1}J carbon correlation from natural abundance ^{13}C in the 20% glycerol in the NMR sample solution (indicated by asterisks in the 13-ms mixing time spectrum). The $C^{\alpha}_{i} - C^{\alpha}_{i+1}$ correlations of the protein become obvious in spectra with mixing times $\geq$66 ms. At 132 ms, all expected inter-residue signals are observed, except for three crosspeaks that

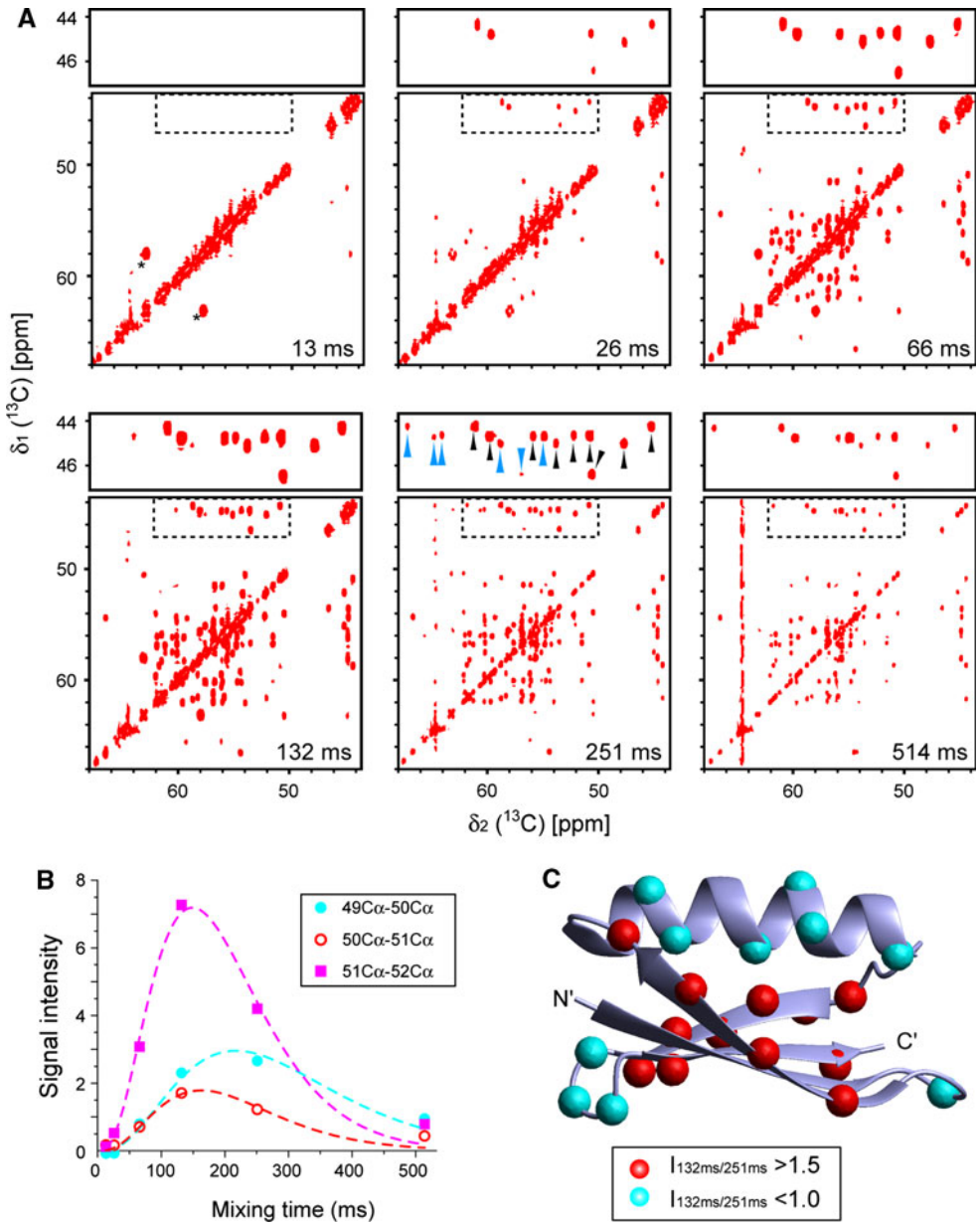

Fig. 2 Mixing efficiency of the CACA-TOCSY experiment with [U-^{2}H, ^{15}N + 2-^{13}C]-GB1 (5 mM) recorded in 20% glycerol at 285 K, which corresponds to a 90 kDa protein in aqueous solution at room temperature. **a** The CACA-TOCSY spectra at various mixing times ranging from 13 to 514 ms. As easily analyzable areas, the regions containing correlations to glycine residues (*dotted square*) are enlarged on top of each spectrum. *Black* and *cyan arrow* heads in the enlarged section at 251-ms mixing time indicate sequential and long range (*i*:*i*−2 or *i*:*i*+2) correlations between C^{α} nuclei, respectively. The measuring time for each spectrum was 5.5 h. The black asterisks in the 13 ms mixing time spectrum denote signals arising from natural abundance ^{13}C-nuclei of glycerol. 2D CACA-TOCSY experiments were recorded with a spectral width of 6,887 Hz (direct) × 6,910 Hz (indirect), on a Bruker (Billerica, MA) Avance 500 spectrometer equipped with a triple-resonance carbon-cryogenic probe (TXO) designed for carbon detection experiments. 512 and 128 complexed data points were recorded for direct and indirect directions, respectively. Eight scans were applied for each increment. The digital resolution in the direct and indirect dimensions after Fourier transformation was 13.5 and 54.0 Hz for the 2D CACA-TOCSY. **b** Examples of build-up curves for three sequential $C^{\alpha}_{i} - C^{\alpha}_{i+1}$ correlations (residues 49–50, 50–51, and 51–52). Curves are drawn to approximately fit the TOCSY time course. **c** Location of residues that have fast and slow build up rates, respectively. On the three-dimensional structure of GB1, $C^{\alpha}_{i} - C^{\alpha}_{i+1}$ correlations with ratios of intensities between 132 and 251 ms, $I_{132}/I_{256} \geq 1.5$ or ≤ 1.0 are indicated by the red and cyan spheres in C^{α}_{i} positions, respectively

overlap with diagonal peaks (Ala23-Ala24, Glu27-Lys28, and Asn35-Adp36) and the peaks of the three Leu residues. The C^{α} carbons of Leu are not ^{13}C-labeled with the 2-^{13}C labeling scheme. However, these correlations can be readily observed with the 1,3-^{13}C labeling (Fig. S1). An alternative would be the addition of ^{2}H, ^{15}N, $^{13}C^{\alpha}$-labeled Leu to the

2-^{13}C labeling culture media, which would increase the cost of labeling, but would allow obtaining all assignments with a single sample.

Interestingly, at a mixing time of 251 ms, "supra-sequential" $C_i^{\alpha} - C_{i+2}^{\alpha}$ correlations are observed for 23 out of 56 possible residue pairs (41%). This represents a unique feature of the CACA-TOCSY spectrum. It is in principle possible to observe correlations farther than $C_i^{\alpha} - C_{i+2}^{\alpha}$ and indeed one weak $C_i^{\alpha} - C_{i+3}^{\alpha}$ correlation between Trp43 and Asp46 was observed in the 2D CACA-TOCSY spectrum at a mixing time of 251 ms.

Most of the $C_i^{\alpha} - C_{i+1}^{\alpha}$ cross peaks showed a maximum intensity at 132-ms mixing time while some of the correlations give strongest intensity at 251 ms (Fig. 2b). We realized that the maxima of the time courses of correlation signals depend on the secondary structure. We use the ratios of the peak heights of $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations at mixing times of 132 and 251 ms, I_{132}/I_{251}, to quantify this dependence. Residues in β-sheets were found to have I_{132}/I_{251} values >1.5, whereas those located in α-helical or in loop regions had I_{132}/I_{251} values <1.0 (Fig. 2c). This observation is consistent with a previous report showing that the $^3J(C_i^{\alpha} - C_{i+1}^{\alpha})$ coupling constants depend on the ψ_i angle, albeit not by a Karplus-type relation (Hennig et al. 2000). The mean $^3J(C_i^{\alpha} - C_{i+1}^{\alpha})$ coupling constants reported for β-sheet secondary structure were 1.7 ± 0.1 Hz, while they were substantially smaller for α-helical secondary structure (Hennig et al. 2000; Peti et al. 2000). To verify these reported coupling constant values, we recorded ^{13}C-detected DQF-COSY spectra with the same protein sample but at higher temperature (308 K). For residues in β-sheets, we observed a ~3 Hz separation between the most upfield and downfield components of the unresolved doublet of doublet (data not shown), which corresponds to the sum of $^3J(C_{i-1}^{\alpha} - C_i^{\alpha})$ and $^3J(C_i^{\alpha} - C_{i+1}^{\alpha})$ couplings and thus confirming the reported values. However, the coupling constants for the α-helical region were too small to be observed in the DQF-COSY experiment and we estimate their value to be less than 1 Hz. Thus, the CACA-TOCSY experiment (or DQF-COSY) with a ^{13}C–^{12}C alternate sample provides qualitative information on the ψ torsion angle by distinguishing larger $^3J(C_i^{\alpha} - C_{i+1}^{\alpha})$ for β-sheet secondary structures from smaller values for helical or loop conformations.

Since the COCO-TOCSY experiment provides information on the ϕ dihedral angle by a Karplus dependence (Balayssac et al. 2006), the complementary use of the CACA- with COCO-TOCSY experiments would particularly be interesting to obtain structural information of perdeuterated proteins. In this context, it is also worth mentioning that most of the $^3J(C_i' - C_{i+1}')$ couplings are smaller than 1.5 Hz (Hu and Bax 1996; Balayssac et al. 2006). Thus, at least for β-sheet regions, the CACA-TOCSY experiment is more sensitive than COCO-TOCSY due to the larger coupling (1.5–2 Hz) and slower relaxation properties of C^{α} atoms compared to C'.

It is of interest to compare the sensitivity of the 2D CACA-TOCSY experiment with the previously reported CA(N)CA experiment (Takeuchi et al. 2010). Figure 3 shows such a comparison of the 2D CACA-TOCSY (132-ms mixing time) and 2D CA(N)CA experiments (Takeuchi et al. 2010) recorded under the same conditions and measuring times. The signal intensities of most of the $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations are substantially higher in the CACA-TOCSY experiment (see slices in Fig. 3a, b). On average, the 2D CACA-TOCSY was found to have a S/N ratio 3.7 times higher than the one for the 2D CA(N)CA experiment. Only eight $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations had smaller S/N ratios than in the 2D CA(N)CA experiment (red bar in Fig. 3b). These eight correlations involve the 4 Val residues, which are labeled in an exceptional way under ^{13}C–^{12}C alternate labeling. For each Val residue in position i, the $C_{i-1}^{\alpha} - C_i^{\alpha}$ and $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations have weak intensity. Unlike all other residues (Fig. 3c), Val residues, as well as Ile, are ^{13}C labeled simultaneously in C^{α} and C^{β} positions with 2-^{13}C glycerol. Therefore the strong ^{1}J carbon–carbon coupling causes detrimental coherence "leaking", which prevents the weak $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations to be observed. However, this effect can be used to identify valine and isoleucine residues as discussed below. A tailored mixing sequence designed to avoid touching the C^{β} resonances of Val and Ile would prevent the "leakage" of coherence.

The CACA-TOCSY experiment with alternate labeling also correlates C^{α} to specific side chain carbons in certain types of amino acid residues. As it is summarized in Fig. 4 for observed and expected C^{α} to side chain carbon correlations, the information can be quite useful for identifying the amino acid type in the same experiment that also provides the sequential assignments. As shown in Fig. 5, intra residue C^{α}–C^{δ} correlations via $^3J(C^{\alpha}–C^{\delta})$ coupling were observed for all Lys residues in the spectra with 132 and 251-ms mixing time (black arrowheads). More interestingly, two sequential correlations between Lys C^{δ} and the C^{α} of the neighboring residues via intra-residual $^3J(C^{\alpha}–C^{\delta})$ and sequential $^3J(C^{\alpha}–C^{\alpha})$ couplings were observed (Fig. 5, cyan arrowheads). Those C^{α}–C^{δ} correlations are also expected for arginine residues (no Arg is present in GB1), as Arg has the same labeling pattern as Lys under 2-^{13}C labeling ($\mathbf{^{13}C^{\alpha}}$–$^{12}C^{\beta}$–$^{12}C^{\gamma}$–$\mathbf{^{13}C^{\delta}}$). Thus, this unique information is helpful in specifying amino acid types as well as in providing additional sequential cross peaks. Such cross peaks can be particularly valuable, if the sequential connections cannot be resolved due to the overlapping in C^{α} signals. The peaks from Lys and Arg can be unambiguously distinguished because of their substantially different C^{δ} chemical shifts. In addition, since proton (or deuterium) in the δ sites of these residues

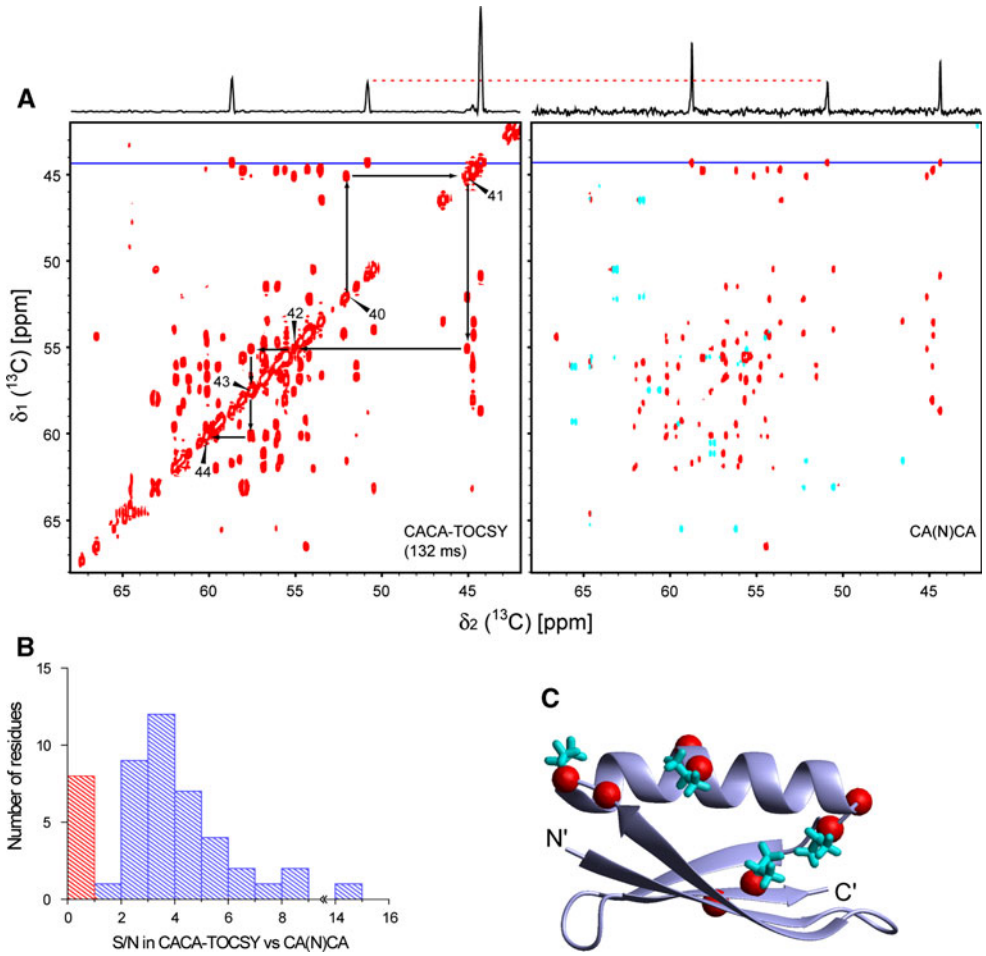

Fig. 3 Comparison between the CACA-TOCSY and CANCA experiments. **a** The CACA-TOCSY (132 ms) and CANCA spectra were recorded at the same condition as in Fig. 2. The measuring time for each experiment was 5.5 h. Slices at the position of the *horizontal blue lines* (corresponding to Gly9) are shown on *top* of each spectrum. The *vertical scales* of the slices are normalized by the peak intensity of the middle resonances originating from C^{α} of residue Asn8. *Black arrows* in the CACA-TOCSY indicate the assignment walk connecting the C^{α} of residues 40–44. Starting from C_i^{α} diagonal at ω (C_i^{α}), one can find the C^{α} frequency of the preceding or succeeding residue ω (C_{i-1}^{α} or C_{i+1}^{α}). **b** Histogram of the ratios of the signal-to-noise values (S/N) in the CACA-TOCSY over the S/N in CANCA. The average of the S/N values for $C_{i+1}^{\alpha} - C_i^{\alpha}$ and $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations is used. *Bars* representing ratios larger and smaller than 1 are colored *blue* and *red*, respectively. **c** Mapping of $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations that have smaller S/N value in the CACA-TOCSY. On the three-dimensional structure of GB1, *red spheres* positioned at the C_i^{α} positions indicate $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations with S/N(CACA-TOCSY)/S/N(CA(N)CA) < 1.0. The side chains of Val residues are shown in *cyan stick* representations

originates from the solvent used for protein expression (H_2O or D_2O); the C^{δ} positions are highly deuterated in our experimental condition. Thus, those correlations are observable in proteins with a high molecular weight. While Glx (Glu and Gln), His, and Trp are also labeled with the same pattern, observing those correlations is not straightforward considering the chemical shift separation between C^{α} and C^{δ} nuclei in those residues (Average C^{δ} chemical shifts in BMRB are 181.5 ppm for Glu, 179.5 ppm for Gln, 119.8 ppm for His, and 126.4 ($C^{\delta 1}$) and 127.5 ($C^{\delta 2}$) ppm for Trp). Leu and Ile sidechain would also have the same ^{13}C-labeling pattern with 1,3-^{13}C labeling. C^{α}–C^{δ} correlations in Leu are rather easy to detect (Figure S2), providing additional amino acid specific information. The C^{α}–C^{δ} correlation of the Ile residue (Ile6) in GB1 was not observed in this study. This might be because the coupling constant happens to be small or the spin mixing was not efficient due to the rather large chemical shift difference between C^{α} and C^{δ} (C^{α}: 61.6 ppm and C^{δ}: 13.5 ppm, in average), in this residue. One can easily detect one bond C^{α}–C^{β} correlation for Ile in 2-^{13}C labeling with a short mixing time (13 ms, Fig. S3). With 2-^{13}C labeling, Val residues are also simultaneously ^{13}C labeled at C^{α} and C^{β} positions as discussed above. However, Val and Ile C^{α}–C^{β} correlation can be easily distinguished from each other by their C^{β} chemical shifts (Fig. S3). Pro residues also exhibit C^{α}–C^{γ} correlations via ^{3}J($^{13}C^{\alpha}$–^{15}N–$^{12}C^{\delta}$–$^{13}C^{\gamma}$) couplings with 1,3-^{13}C

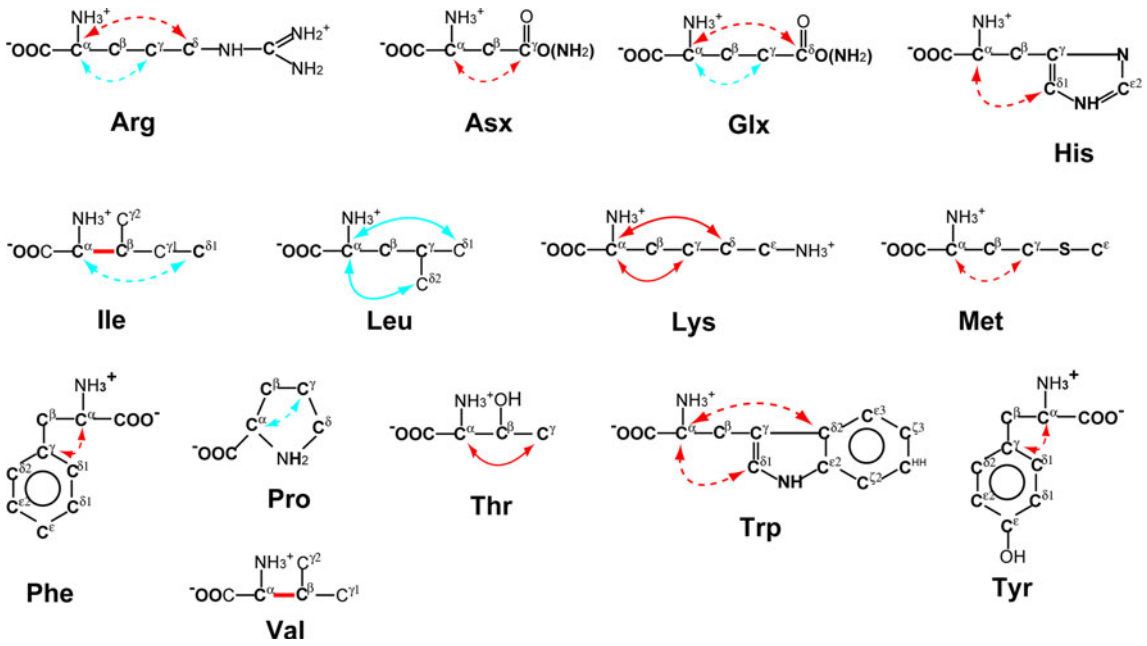

Fig. 4 Summary of the amino acid specific correlations between C^{α} and side chain carbon in CC-TOCSY experiment with ^{13}C–^{12}C alternate labeling strategies. *Red* and *cyan arrows* indicate ^{2}J and ^{3}J carbon–carbon couplings that are expected to be observed with 2-^{13}C and 1,3-^{13}C-labeling, respectively. Red bonds indicate ^{1}J carbon–carbon coupling observed with 2-^{13}C labeling. Note that most amino acids can be ^{13}C labeled in a couple of different patterns as it is exemplified in Arg (see text). The correlations observed in this study are shown in *solid lines*, whereas, the expected but not observed correlations are shown by *broken lines*. The correlations with broken lines are not observed due to the absence of the residues in the GB1 protein or the inefficient TOCSY transfer

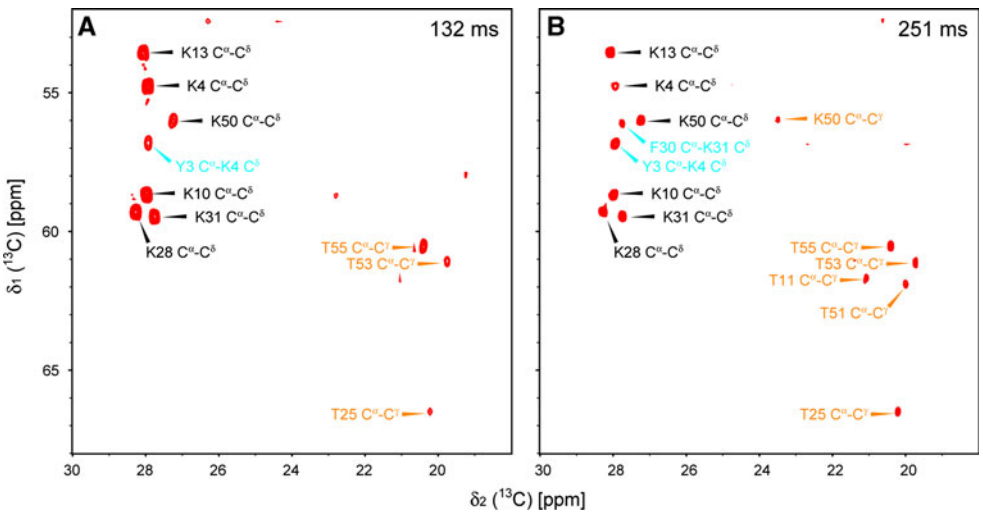

Fig. 5 C^{α} to sidechain carbon correlations in the CACA-TOCSY experiment. The spectra were recorded with [U-^{2}H, ^{15}N + 2-^{13}C]-GB1 (5 mM) under the same conditions as those in Fig. 2, with mixing times of 132 ms (**a**) and 251 ms (**b**). The measuring time for each experiment was 5.5 h. *Black* and *cyan arrowheads* in the spectra indicate intra-residual and sequential C^{α}–C^{δ} correlations, respectively. *Orange arrowheads* are for intra-residual C^{α}–C^{γ} correlations

labeling. However, ^{2}H and ^{13}C labeled media have to be used to achieve complete deuteration of the C^{γ} and C^{δ} positions in 1,3-^{13}C labeling.

In addition to correlations via ^{3}J and ^{1}J couplings mentioned above, intra-residual C^{α}–C^{γ} crosspeaks via $^{2}J(C^{\alpha}$–$C^{\gamma})$ coupling were also observed in one out of six Lys and five out of ten Thr residues (Fig. 5, orange arrow heads). Because the coupling constants for these ^{2}J C–C couplings are smaller than those for ^{3}J couplings, those correlations have a maximum at 256 ms mixing time and are relatively weak in intensity. Those cross peaks would be easier to observe in smaller molecular weight systems. With 2-^{13}C labeling schemes, one can also identify C^{α}–C^{γ} correlations in Met, Asx (Asp and Asn), Tyr and Phe, which have the same labeling pattern as Thr ($^{13}C^{\alpha}$–$^{12}C^{\delta}$–$^{13}C^{\gamma}$). However, observing C^{α}–C^{γ} correlations in Asx, Phe, and Tyr are not straightforward unless broad-band spin-locking schemes are used (Average chemical

shifts of Asp, Asn, Phe, and Tyr C^{δ} in the BMRB are 179.2, 176.8, 137.3, 129.1 ppm, respectively). Note that $^{2}J\ C^{\alpha}$–C^{γ} correlations can additionally be observed for Arg and Glx with the 1,3-^{13}C labeled schemes (see Fig. S3 for Glx correlations). The amino-acid selective information mentioned above can be used complementary with amino-acid spin type information, such as obtained from the CC-TOCSY-DIPAP experiment using a uniformly labeled sample (Bermel et al. 2005).

Conclusion

In summary, we presented the CACA-TOCSY, a novel ^{13}C direct detection experiment that provides sequential and supra-sequential correlations among C^{α} resonances via $^{3}J(C^{\alpha}–C^{\alpha})$ couplings. The experiment becomes possible with the alternate $^{13}C–^{12}C$ labeling strategy, which makes weak scalar couplings to be the dominant routes for coherence transfer pathways, and except for two residue types, there is no need for eliminating large one-bond carbon–carbon couplings during detection. The CACA-TOCSY with alternate ^{13}C labeling has improved sensitivity compared to the established coherence transfer pathways that rely on heteronuclear scalar coupling, such as CA(N)CA experiment. Furthermore, with a longer spin-mixing time, the experiment provides unique supra-sequential correlations between C^{α} atoms separated by $\geq$2 residues. In addition, since $^{3}J(C_i^{\alpha} - C_{i+1}^{\alpha})$ couplings correlate with the ψ_i protein backbone dihedral angles, CACA-TOCSY build-up curves reveal information about protein secondary structure. In the same experiment as used for sequential assignments, the C–C TOCSY mixing scheme with the alternate ^{13}C-^{12}C labeling strategy also provides C^{α} to C^{β}, C^{γ}, and C^{δ} correlations for certain amino acids depending on side chain ^{13}C labeling pattern. These correlations are useful for unambiguous determination of amino acid types. Val (C^{β}), Ile (C^{β}), Met (C^{γ}), Thr (C^{γ}), Arg (C^{δ}), and Lys (C^{δ}) can be identified in 2-^{13}C labeled samples and Arg (C^{γ}), Glx (C^{γ}), Pro (C^{γ}), and Leu (C^{δ}) in 1,3-^{13}C labeled samples (Fig. 4). However, a deuterated ^{13}C source might be needed for 1,3-^{13}C labeling to detect those couplings in large molecular weight proteins, so that side-chain carbons of interest are deuterated. Lys and Arg in 2-^{13}C labeling and Leu in 1,3-^{13}C labeling are of particular interest providing also sequential $C_i^{\alpha} - C_{i+1}^{\alpha}$ correlations. These sequential correlations would be useful to solve problems with the degeneracy of overlapped C^{α} signals. The experimental scheme also has a potential to observe correlation between C^{α} and sidechain carbonyl carbons as well as with aromatic carbons. These correlations are difficult to observe because of the large chemical shift difference between the spins of interests. Thus, implementation of broadband mixing schemes would be particularly interesting and are under development. While the 2D spectra shown here provide quite narrow lines and good dispersion, one could further improve the spectral resolution and sensitivity by applying non-uniform sampling and processing schemes, which will be especially helpful for peaks near the diagonal (Hyberts et al. 2010). In addition, a 3D version of this experiment, which avoids signal overlap in larger proteins, is currently under development. Although the sensitivity of the experiment was largely improved by using the homonuclear TOCSY transfer step rather than C–N heteronuclear coherences, further improvement in sensitivity might be needed in the routine use of this experiment for very large single polypeptide proteins. In this regard, an experiment utilizing the larger ^{1}H polarization as a magnetization source but not for the detection can be considered (Bermel et al. 2009a, b). For systems with less severe relaxation losses, proton-excited/detected 3D experiments that contain CACA-TOCSY transfers can also provide unique sequential information. Experiments of this type have been shown already for a series of 3D pulse sequences that use C′C′-TOCSY transfers. These include 3D (H)N(CO-TOCSY)NH, 3D (H)CA(CO-TOCSY)NH, and 3D (H)CBCA(CO-TOCSY)NH but used uniformly ^{13}C-labeled samples (Liu et al. 2000). Proton detected experiments that use the CACA-TOSY module might provide the method of choice given the increased polarization of proton nuclei.

Acknowledgments Dr. Tsyr-Yan (Dharma) Yu, Dr. Arthanari Haribabu, and Dr. Alexander Koglin are acknowledged for stimulating discussions. This work was supported by the NIH (grants AI37581, GM47467 and EB 002026).

References

Arnesano F, Banci L, Piccioli M (2005) NMR structures of paramagnetic metalloproteins. Q Rev Biophys 38:167–219

Balayssac S, Jimenez B, Piccioli M (2006) ^{13}C direct detected COCO-TOCSY: a tool for sequence specific assignment and structure determination in protonless NMR experiments. J Magn Reson 182:325–329

Bermel W, Bertini I, Felli IC, Kummerle R, Pierattelli R (2003) ^{13}C Direct detection experiments on the paramagnetic oxidized monomeric copper zinc superoxide dismutase. J Am Chem Soc 125:16423–16429

Bermel W, Bertini I, Duma L, Felli IC, Emsley L, Pierattelli R, Vasos PR (2005) Complete assignment of heteronuclear protein resonances by protonless NMR spectroscopy. Angew Chem Int Ed Engl 44:3089–3092

Bermel W, Bertini I, Felli IC, Lee YM, Luchinat C, Pierattelli R (2006a) Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. J Am Chem Soc 128:3918–3919

Bermel W, Bertini I, Felli IC, Piccioli M, Pierattelli R (2006b) ^{13}C-detected protonless NMR spectroscopy of proteins in solution. Prog Nucl Magn Res Spec 48:25–45

Bermel W, Bertini I, Felli IC, Matzapetakis M, Pierattelli R, Theil EC, Turano P (2007) A method for Cα direct-detection in protonless NMR. J Magn Reson 188:301–310

Bermel W, Bertini I, Csizmok V, Felli IC, Pierattelli R, Tompa P (2009a) H-start for exclusively heteronuclear NMR spectroscopy: the case of intrinsically disordered proteins. J Magn Reson 198: 275–281

Bermel W, Bertini I, Felli IC, Pierattelli R (2009b) Speeding up ^{13}C direct detection biomolecular NMR spectroscopy. J Am Chem Soc 131:15339–15345

Cai S, Seu C, Kovacs Z, Sherry AD, Chen Y (2006) Sensitivity enhancement of multidimensional NMR experiments by paramagnetic relaxation effects. J Am Chem Soc 128:13474–13478

Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature 420:98–102

Eletsky A, Moreira O, Kovacs H, Pervushin K (2003) A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using ^{13}C spectroscopy. J Biomol NMR 26:167–179

Felli I, Brutscher B (2009) Recent advances in solution NMR: fast methods and heteronuclear direct detection. ChemPhysChem 10: 1356–1368

Felli I, Pierattelli R, Glaser S, Luy B (2009) Relaxation-optimised Hartmann–Hahn transfer using a specifically Tailored MOCCA-XY16 mixing sequence for carbonyl–carbonyl correlation spectroscopy in ^{13}C direct detection NMR experiments. J Biomol NMR 43:187–196

Frueh DP, Arthanari H, Wagner G (2005) Unambiguous assignment of NMR protein backbone signals with a time-shared triple-resonance experiment. J Biomol NMR 33:187–196

Goddard TD, Kneller DG (2004) SPARKY 3. University of California, San Francisco

Guo C, Geng C, Tugarinov V (2009) Selective backbone labeling of proteins using 1, 2-$^{13}C_2$-pyruvate as carbon source. J Biomol NMR 44:167–173

Hennig M, Bermel W, Schwalbe H, Griesinger C (2000) Determination of ψ torsion angle restraints from $^3J(C\alpha, C\alpha)$ and $^3J(C\alpha, HN)$ coupling constants in proteins. J Am Chem Soc 122:6268–6277

Hsu S-TD, Bertoncini CW, Dobson CM (2009) Use of protonless NMR spectroscopy to alleviate the loss of information resulting from exchange-broadening. J Am Chem Soc 131:7222–7223

Hu J-S, Bax A (1996) Measurement of three-bond ^{13}C–^{13}C J couplings between carbonyl and carbonyl/carboxyl carbons in isotopically enriched proteins. J Am Chem Soc 118:8170–8171

Hyberts SG, Takeuchi K, Wagner G (2010) Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR Data. J Am Chem Soc 132:2145–2147

Lee D, Vögeli B, Pervushin K (2005) Detection of C′, Cα correlations in proteins using a new time- and sensitivity-optimal experiment. J Biomol NMR 31:273–278

LeMaster DM, Kushlan DM (1996) Dynamical mapping of *E. coli* thioredoxin via ^{13}C NMR relaxation analysis. J Am Chem Soc 118:9255–9264

Liu A, Riek R, Wider G, von Schroetter C, Zahn R, Wuthrich K (2000) NMR experiments for resonance assignments of ^{13}C, ^{15}N doubly-labeled flexible polypeptides: application to the human prion protein hPrP(23–230). J Biomol NMR 16:127–138

Marion D, Ikura M, Tschudin R, Bax A (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. J Magn Reson 85:393–399

Peti W, Hennig M, Smith LJ, Schwalbe H (2000) NMR spectroscopic investigation of ψ torsion angle distribution in unfolded ubiquitin from analysis of $^3J(C_{\alpha,}\ C\alpha)$ coupling constants and cross-correlated relaxation rates. J Am Chem Soc 122:12017–12018

Serber Z, Richter C, Dotsch V (2001) Carbon-detected NMR experiments to investigate structure and dynamics of biological macromolecules. Chembiochem 2:247–251

Shaka AJ, Keeler J, Frenkiel T, Freeman R (1983) An improved sequence for broadband decoupling: WALTZ-16. J Magn Reson 52:335–338

Shaka AJ, Barker PB, Freeman R (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. J Magn Reson 64:547–552

Shimba N, Stern AS, Craik CS, Hoch JC, Dotsch V (2003) Elimination of ^{13}Calpha splitting in protein NMR spectra by deconvolution with maximum entropy reconstruction. J Am Chem Soc 125: 2382–2383

Takeuchi K, Sun ZY, Wagner G (2008a) Alternate ^{13}C–^{12}C labeling for complete mainchain resonance assignments using C alpha direct-detection with applicability toward fast relaxing protein systems. J Am Chem Soc 130:17210–17211

Takeuchi K, Sun ZY, Wagner G (2008b) Alternate ^{13}C–^{12}C labeling for complete mainchain resonance assignments using Cα direct-detection with applicability toward fast relaxing protein systems. J Am Chem Soc 130:17210–17211

Takeuchi K, Frueh DP, Hyberts SG, Sun ZJ, Wagner G (2010) High-resolution 3D CANCA NMR experiments for complete main-chain assignments using C_α direct-detection. J Am Chem Soc 132:2945–2951