Published in final edited form as:
J Biomol NMR. 2011 January ; 49(1): 17–26. doi:10.1007/s10858-010-9456-2.

HNCA-TOCSY-CANH experiments with alternate ^{13}C-^{12}C labeling: a set of 3D experiment with unique supra-sequential information for mainchain resonance assignment

Koh Takeuchi,
Department of Biochemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan

Maayan Gal,
Department of Biochemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Hideo Takahashi,
Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan. Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo 113-0033, Japan

Ichio Shimada, and
Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan. Department of Physical Chemistry, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo 113-0033, Japan

Gerhard Wagner
Department of Biochemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Gerhard Wagner: gerhard_wagner@hms.harvard.edu

Abstract

Described here is a set of three-dimensional (3D) NMR experiments that rely on CACA-TOCSY magnetization transfer via the weak $^{3}J_{C\alpha C\alpha}$ coupling. These pulse sequences, which resemble recently described ^{13}C detected CACA-TOCSY (Takeuchi et al. 2010) experiments, are recorded in $^{1}H_2O$, and use ^{1}H excitation and detection. These experiments require alternate ^{13}C-^{12}C labeling together with perdeuteration, which allows utilizing the small $^{3}J_{C\alpha C\alpha}$ scalar coupling that is otherwise masked by the stronger $^{1}J_{CC}$ couplings in uniformly ^{13}C labeled samples. These new experiments provide a unique assignment ladder-mark that yields bidirectional supra-sequential information and can readily straddle proline residues. Unlike the conventional HNCA experiment, which contains only sequential information to the $^{13}C^{\alpha}$ of the preceding residue, the 3D hnCA-TOCSY-caNH experiment can yield sequential correlations to alpha carbons in positions *i−1, i + 1* and *i−2*. Furthermore, the 3D hNca-TOCSY-caNH and Hnca-TOC-SY-caNH experiments, which share the same magnetization pathway but use a different chemical shift encoding, directly couple the ^{15}N-^{1}H spin pair of residue *i* to adjacent amide protons and nitrogens at positions *i−2, i−1, i + 1* and *i + 2*, respectively. These new experimental features make protein backbone assignments more robust by reducing the degeneracy problem associated with the conventional 3D NMR experiments.

Correspondence to: Gerhard Wagner, gerhard_wagner@hms.harvard.edu.

Keywords

Alternate ^{13}C labeling; TOCSY; Nuclear magnetic resonance (NMR); Sequential assignment; Triple resonance; Three dimensional; Supra sequential assignments

Introduction

Establishing sequence-specific assignments is fundamental for structural analysis of biological macromolecules using NMR spectroscopy. Numerous pulse sequences have been developed to assign resonances to the right nuclei in the polypeptide sequence (for a review see (Ferentz and Wagner 2000)). For naming pulse sequences and following common practice, upper case letters throughout the manuscript indicate the frequency labeled nuclei and lower case letters stand for nuclei that are used in coherence transfers but are not frequency encoded. The foremost basic experiments for protein backbone assignment, such as the HNCA and HNcoCA, correlate a ^{15}N-^{1}H spin pair to the preceding $^{13}C^{\alpha}_{i-1}$. Where coincidental overlap of C^{α} signals prevents unambiguous assignments, experiments correlating side chain resonances can be recorded, such as the HNCACB and HNcoCACB pair, or the HccoNH or hCcoNH TOCSY type of experiments (Grzesiek et al. 1993; Lin and Wagner 1999; Logan et al. 1993; Lyons et al. 1993). Typically, these experiments rely on matching the carbon frequencies measured in the set of these 3D experiments. However, often severe cases of overlap still exist, or ambiguities arise because the different pulse sequences may lead to slightly altered sample temperature, which increases the chance of incorrect assignments. In future, the latter problem might be eliminated with the development of the "*T*-lock", which automatically compensates the radio-frequency induced sample heating (Hiller et al. 2009). In cases where carbon signals overlap, especially for Gly, which lacks C^{β}, relying on the better dispersed ^{15}N can be advantageous. This can be done with a set of pulse sequences that correlate a pair of ^{15}N-^{1}H nuclei directly to the adjacent N and/or H (*i−1* and *i + 1*) in a single experiment, such as the HNcocaNH (Sun et al. 2005) and HNcaNH (Frueh et al. 2006). These experiments tend to be valuable when dealing with large molecular weight or unstructured proteins.

Although the basic set of triple resonance experiments proofed to be robust, the exclusive use of ^{13}C-^{15}N and ^{15}N-^{1}H direct scalar coupling for INEPT-based (Morris and Freeman 1979) coherence transfer limits the number of sequential correlations that can be observed in these types of experiments. Indeed, detecting a correlation that couples an ^{15}N-^{1}H spin pair with the succeeding $^{13}C^{\alpha}_{i+1}$, or nuclei separated by more then one residue is not straight forward due to the absence of strong scalar coupling connecting these nuclei. However, obtaining these longer-range through-bond correlations is potentially very useful for preventing incorrect assignments due to accidental peak degeneracy. This could be particularly useful for large unfolded proteins, which often contain phosphorylation sites followed by Pro residues (Ser-Pro, Thr-Pro).

Very recently, we have shown that utilizing TOCSY in ^{13}C detected experiments together with alternate ^{13}C-^{12}C labeling (CACA-TOCSY), enables direct connectivity between $^{13}C^{\alpha}$ nuclei in adjacent amino acids via the weak long-range ^{3}J coupling (~2 Hz) (Takeuchi et al. 2010). In uniformly ^{13}C labeled samples, these weak couplings are masked by the strong ^{1}J carbon–carbon splitting but are readily observed in proteins expressed with the alternate ^{13}C-^{12}C labeling scheme (LeMaster and Kushlan 1996; Takeuchi et al. 2010; Takeuchi et al. 2008). Since the heteronuclear ^{13}C-^{15}N scalar couplings are preserved in the labeling scheme, the CACA-TOCSY experiment can easily be integrated into the conventional NH-detected set of 3D experiments. It is worth mentioning that a 3D

experiment utilizing a long range $^{13}C' - {}^{13}C'$ coupling have already been proposed for uniformly ^{13}C labeled samples (Liu et al. 2000). Although these experiments were shown to be very useful in the analysis of unstructured proteins, they are not always the method of choice. The $^{13}C'$ nuclei suffer from fast transverse relaxation due to their large chemical shift anisotropy (CSA) when dealing with large-molecular-weight proteins, particularly at high magnetic-fields. In contrast, C^{α} nuclei have a small CSA and consequently slow transverse relaxation. We have previously shown that deuteration of C^{α} nuclei is straightforward with the alternate ^{13}C-^{12}C labeling scheme (LeMaster and Kushlan 1996; Takeuchi et al. 2010; Takeuchi et al. 2008). Indeed, if samples are deuterated, experiments utilizing long-range C^{α} couplings may have a wide range of applications.

Here we describe a set of proton-detected 3D NMR experiments utilizing the long-range $^{3}J_{C\alpha C\alpha}$ coupling, namely the hnCA-TOCSY-caNH, hNca-TOCSY-caNH, and Hnca-TOCSY-caNH. All of these experiments share the same coherence pathways and the ability to provide unique bidirectional supra-sequential correlations. The merits of the experiments are two-fold. (1) This set of experiments can be used without or in addition to the conventional 3D experiments to avoid the previously described degeneracy problems, as simultaneous observation of more than two sequential correlation peaks makes assignment easier and more reliable. (2) These experiments can also bridge breaks in sequential connectivity introduced by single proline (Pro) residues. As the Pro residues lack the amide proton, the correlation originating from this residue cannot be detected in a conventional 3D experiments. However, the supra-sequential correlation obtained here can bypass these gaps by directly connecting residues preceding and succeeding prolines. This is important especially for the assignment of Pro-rich proteins, which play an important role in many cellular functions (Enkhmandakh et al. 2006; Marintchev et al. 2007; Rao et al. 1997; Ruaro et al. 1997; Venot et al. 1998). In particular, prolines are often found in phosphorylation sites, which has hampered NMR characterization of such important regulatory elements in the past.

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) and the unfolded regulatory domain of NFAT

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] 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] glycerol and 1 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).

The regulatory domain of NFAT was produced similarly. It was expressed as a cleavable fusion with GB1 and the tag was removed with TEV protease prior to the NMR experiments. The sample concentration used in the experiment was 150 μM.

NMR experiments

NMR spectra were recorded on a Bruker (Billerica, MA) Avance 600 spectrometer equipped with a triple-resonance proton-cryogenic probe (TXI). Spectra of the 2-^{13}C labeled GB1 sample (3.5 mM) were recorded at 25°C in buffer containing 10 mM sodium phosphate (pH 6.8), 100 mM NaCl and 40% w/v deuterated glycerol in H_2O. The molecular tumbling of GB1 under those conditions corresponds to a 40 kDa protein at 25°C (Takeuchi et al. 2008). The 3D hnCA-TOCSY-caNH experiment was recorded with a spectral width of 4,839 Hz (^{13}C, indirect) × 2,189 Hz (^{15}N, indirect) × 9,615 Hz (^{1}H, direct), centered at 55 ppm (^{13}C), 118 ppm (^{15}N), and 4.7 ppm (^{1}H). 64 (^{13}C, indirect) × 28 (^{15}N, indirect) × 512 (^{1}H, direct), complex data points were recorded for indirect and direct dimensions, respectively. The 3D hNca-TOCSY-caNH experiment was recorded with spectral widths of 2,189 Hz (^{15}N, indirect) × 2,189 Hz (^{15}N, indirect) × 9,615 Hz (^{1}H, direct), centered at 118 ppm (^{15}N) and 4.7 ppm (^{1}H). A total of 45 (^{15}N, indirect, t_1) × 28 (^{15}N, indirect, t_2) × 512 (^{1}H, direct) complex data points were recorded for indirect and direct dimensions, respectively. The 3D Hnca-TOCSY-caNH experiment was recorded with a spectral width of 3,600 Hz (^{1}H, indirect) × 2,189 Hz (^{15}N, indirect) × 9,615 Hz (^{1}H, direct), centered at 8 ppm (^{1}H, indirect) 118 ppm (^{15}N) and 4.7 ppm (^{1}H, direct). In total, 64 (^{1}H, indirect) × 28 (^{15}N, indirect) × 512 (^{1}H, direct) complex data points were recorded for indirect and direct dimensions, respectively. Spectra were recorded with 16 scans for each increment, with a 1 s recycle delay. Spectra were analyzed with the program Sparky (Goddard and Kneller 2006).

Results and discussion

Figure 1a shows the pulse sequence used to execute the 3D hnCA-TOCSY-caNH NMR experiment. The pulse sequence correlates each ^{15}N-^{1}H pair of residue *i* to the sequentially adjacent C$^{\alpha}$s at positions *i−2* to *i + 1*, as shown in Fig. 1b. Unlike the conventional HNCA and HNcoCA experiments, which only provide a sequential ^{13}C$^{\alpha}$ correlation to the preceding residue, the 3D hnCA-TOCSY-caNH experiment provides sequential connectivity for ^{13}C$^{\alpha}$s in both directions. Clearly, this unique feature yields more reliable assignments in case that the preceding ^{13}C$^{\alpha}$ resonance of a certain residue *i* is degenerate with the ^{13}C$^{\alpha}$ of a different residue. In previous publications it was shown that applying isotropic mixing on ^{13}C$^{\alpha}$ can transfer magnetization among adjacent backbone ^{13}C$^{\alpha}$ carbons (Bermel et al. 2006; Takeuchi et al. 2010). The alternate ^{13}C-^{12}C labeling scheme is essential for establishing this coherence transfer, because in uniformly ^{13}C-labeled samples the strong ^{1}J carbon–carbon couplings cause detrimental coherence “leaking” and prevent the weak $^{3}J_{C\alpha C\alpha}$ correlations to be observed (Takeuchi et al. 2010). While the same correlation can be established by a series of C$^{\alpha}$ –N INEPT-based transfers, magnetization transfer through isotropic mixing is on average ~4 time more efficient as discussed previously (Takeuchi et al. 2010).

Figure 1b shows the coherence transfer pathway resulting from the 3D hnCA-TOCSY-caNH experiment. This can be summarized as $^{1}H \rightarrow {}^{15}N \rightarrow {}^{13}C^{\alpha}(t_1) \rightarrow$ TOCSY mixing $\rightarrow {}^{13}C^{\alpha} \rightarrow {}^{15}N$ (t_2) $\rightarrow {}^{1}H$ (t_3), where t_1 and t_2 indicate the indirect ^{13}C$^{\alpha}$ and ^{15}N evolution periods, respectively. Finally, t_3 is the detection period of proton coherence. Initial amide proton polarization of residue *i* is transferred to the directly attached amide nitrogen through an INEPT step at time point *b*. During the following INEPT (*c*−*d*) $^{15}N_{Hi}$ coherence becomes in-phase with respect to $^{1}H_i$ and antiphase with respect to $^{13}C^{\alpha}_{i-1}$ and $^{13}C^{\alpha}_{i}$. The C′ pulse during nitrogen transverse evolution is applied in case that the preceding C′ is ^{13}C labeled in the alternate labeling scheme. A ^{13}C$^{\alpha}$ constant time (CT) evolution period, t_1, follows through time points *e* and *f* (Santoro and King 1992; Vuister and Bax 1992). This block also refocuses ^{13}C$^{\alpha}$ coherences with respect to $^{15}N_{Hi}$. There is no need to decouple C′ during the C$^{\alpha}$ evolution, since carbonyls are not ^{13}C labeled in the same residue. Except for Ile and

Val, there is no $^{1}J_{C\alpha C\beta}$ coupling in the ^{13}C-^{12}C alternate labeling scheme, and the optimum delay to refocus $^{15}N_H$ and $^{13}C^{\alpha}$, can be set to 22 ms. This delay is more efficient then the commonly used 28.5 ms, which corresponds to $1/^{1}J_{C\alpha C\beta}$. It is worth noting that, although this delay is not optimized for Ile and Val, the presence of the $^{1}J_{C\alpha C\beta}$ coupling allows an easy identification of these residues. The refocusing delay of 22 ms will result in an opposite cross-peak sign due to the remaining $^{1}J_{C\alpha C\beta}$ coupling. The following step plays the key part in this new pulse sequence (*f–g*). The in-phase $^{13}C^{\alpha}$ magnetization at time point "*f*" is transferred to the sequentially adjacent $^{13}C^{\alpha}$s. Then, a homonuclear isotropic mixing block, such as the FLOPSY-16 sequence (Kadkhodaie et al. 1991) is used to accomplish this transfer step. The same mixing procedure simultaneously produces supra-sequential connectivities to $^{13}C^{\alpha}_{i-2}$. In addition, coherence transfer from $^{13}C^{\alpha}_{i}$ to $^{13}C^{\alpha}_{i+1}$ ensures bidirectional sequence connectivity along the protein backbone. After a 22 ms evolution period under the $^{1}J_{C\alpha N}$ and $^{2}J_{C\alpha N}$ coupling, the $^{13}C^{\alpha}$ coherences are then transferred to the directly coupled $^{15}N_H$ nuclei ($^{15}N_{Hi-2}$ to $^{15}N_{Hi+2}$) by the two 90° pulses at points *h* and *i*. This is followed by the $^{15}N_H$ CT evolution period from time points *i* to *j*, which is optimized for refocusing $^{15}N_H$ – $^{13}C^{\alpha}$ couplings at the same time. During this period the $^{15}N_H$ coherences become in-phase with respect to $^{13}C^{\alpha}$ and antiphase with respect to $^{1}H_N$. Finally, the $^{15}N_H$ coherences are transferred back to $^{1}H_N$ by a sensitivity-enhanced gradient-selection scheme (Muhandiram and Kay 1994), and the proton magnetization is detected during acquisition period t_3.

Similar to the HNCA experiment, where the intra-residue cross peak is stronger than the sequential peak, the cross peaks in the hnCA-TOCSY-caNH experiment also differ in intensity. It is expected that the peak intensities have the following orders: $^{13}C^{\alpha}_{i} > ^{13}C^{\alpha}_{i-1} > ^{13}C^{\alpha}_{i+1} > ^{13}C^{\alpha}_{i-2}$. While the latter two correlations need to go through both the TOCSY and INEPT coherence transfers, the former two correlations proceed through an additional transfer pathway via simple INEPT blocks in addition to the TOCSY transfers. The $^{13}C_{i-2}$ cross-peak is the least sensitive since this correlation can be observed only via coherence transfer by TOCSY through $^{2}J_{NC\alpha}$ couplings, while the $^{13}C^{\alpha}_{i+1}$ correlation comes from a TOCSY N to Cα transfer, which relies on the larger $^{1}J_{NC\alpha}$ couplings (Fig. 1b, c). The delays can be optimized to favor the weaker *i* to *i−2* transfer by elongating the mixing time as well as by optimizing the N-CA evolution to 70 ms as we discussed in the previous paper (Takeuchi et al. 2010). However, this strategy would work only for relatively small proteins.

In addition to the previous argument, the intensity of each correlation also depends on the $^{13}C^{\alpha}$ labeling efficiency for each residue involved in the coherence transfer pathway. The *E. coli* metabolic pathways result in different labeling probabilities for each residue when using alternately labeled glycerol (or pyruvate) as the carbon source. The 2-C atoms of glycerol are inserted exclusively into the C^{α} position in some residues, but are directed to C′ or C^{β} for others. If the $^{13}C^{\alpha}$ labeling rate at position *i* is αR_i, the $^{15}N^{1}H_i$ to $^{13}C^{\alpha}_{i}$ and $^{13}C^{\alpha}_{i-1}$ cross-peak intensity depends on αR_i and αR_{i-1}, respectively. On the other hand, $^{13}C^{\alpha}_{i+1}$ and $^{13}C^{\alpha}_{i-2}$ resonances depend on the product of the labeling rates of two $^{13}C^{\alpha}$s involved in the TOCSY transfer pathway, such as $\alpha R_i \times \alpha R_{i+1}$ for $^{13}C^{\alpha}_{i+1}$ and $\alpha R_{i-1} \times \alpha R_{i-2}$ for $^{13}C^{\alpha}_{i-2}$ transfers. The expected $^{13}C^{\alpha}$ labeling rates for each residue have been discussed in detail in previous studies (LeMaster and Kushlan 1996; Takeuchi et al. 2010; Takeuchi et al. 2008). The current alternate labeling protocol that uses 2-^{13}C glycerol, leads to ~100% $^{13}C^{\alpha}$ labeling for Ala, Cys, Gly, His, Lys, Phe, Ser, Trp, Tyr, and Val. A labeling rate of ~60% is observed for Asn, Asp, Ile, Met, and Thr. The rest of the amino acids (Arg, Glu, Gln, and Pro) are labeled at ~20%. Only Leu is not ^{13}C labeled in the α-position when using 2-^{13}C glycerol as precursor.

To test this new experimental scheme we have recorded the 3D hnCA-TOCSY-caNH experiment on uniformly $^{2}H^{15}N$ and alternately ^{13}C-^{12}C labeled B1 domain of protein G (GB1) produced with 2-^{13}C glycerol as the only carbon source to maximize the $^{13}C^{\alpha}$ labeling rate (Takeuchi et al. 2008). The protein was dissolved in H_2O-based buffer, and the data were recorded on a 600 MHz magnet, equipped with a cryogenic probe. To simulate correlation times of a 40 kDa protein, 40% glycerol was added to the buffer while the temperature was kept at 298 K (for a calibration of the correlation time see (Takeuchi et al. 2008)). The rotational correlation time estimated by the [^{15}N,^{1}H]-TRACT experiment (Lee et al. 2006) was 18 ns at this condition. In order to transfer magnetization between different C^{α}s by isotropic mixing, a mixing time of 132 ms was experimentally calibrated for optimal transfer (Takeuchi et al. 2010). A longer mixing time can be used to enhance the coherence transfer in weaker $^{3}J_{C\alpha C\alpha}$ coupling regions, such as in α-helices or tight turns (Takeuchi et al. 2010).

Figure 2a shows $\omega_1(^{13}C^{\alpha})/\omega_3(^{1}H_N)$ strips from the 3D hnCA-TOCSY-caNH spectrum of GB1 with the above experimental settings. For simplicity, only strips showing the sequential connectivity of residues Thr49–Thr53 are depicted. As shown in the figure, both $^{13}C_{i-1}$ and $^{13}C_{i+1}$ correlations are observed. These bidirectional correlations are essential in avoiding degeneracy in the preceding C^{α} correlation ($^{13}C_{i-1}$). This is clearly exemplified in Fig. 2b. When trying to determine the strip originating from Thr44 based on the $^{13}C_{i-1}$ signal in the Tyr45 strip (signal at 57.4 ppm), the HNCA spectrum provides three different candidates (#1 ~ 3). However, if the additional data from the hnCA-TOCSY-caNH experiment are used, one can unambiguously identify strip #1 as representing the correlations to Thr44 by checking the strip that contains the correlation to the $^{13}C^{\alpha}$ signal of Tyr45. The experiment also provides a unique connectivity between the ^{15}N-^{1}H pair of residue *i* to the C^{α} at positions *i−2*. Since the transfer pathway from the $^{15}N_i$-$^{1}H_i$ spin pair to C^{α}_{i-2} does not involve the amide moiety of residue *i−1*, the correlation can be observed even if the residue *i−1* is a proline. This suggests that the experiment can skip single Pro residue gaps, which usually cause breaks in sequential assignments in conventional HNCA type of experiments. Although, a better incorporation of $^{13}C^{\alpha}$-Pro may be needed for a sensitive detection because of the relatively low $^{13}C^{\alpha}$ labeling rate of prolines (~20%).

The sequential connectivities observed in the hnCA-TOCSY-caNH experiment are summarized in Fig. 2c. Without any exception, all of the ^{15}N-^{1}H pairs have correlations to $^{13}C^{\alpha}_{i}$ and $^{13}C^{\alpha}_{i-1}$. In addition, 44 residues (81%) have a bidirectional sequential correlation to the $^{13}C^{\alpha}_{i+1}$. The missing resonances are those involving residues with less than 20% $^{13}C^{\alpha}$ labeling rate. In addition, valuable supra-sequential correlations for $^{13}C^{\alpha}_{i-2}$ were observed for a significant number of residues (23/53 residues, 43%).

In addition to the concentrated GB1 standard sample, we have tested the hnCA-TOCSY-caNH pulse sequence on the deuterated 17 kDa unfolded regulatory domain of nuclear factor of activated T cells (NFAT) that contains many phosphorylation sites and prolines. As shown in Fig. 3a, an hnCA-TOCSY-caNH spectrum was obtained with reasonable sensitivity in ~ 4.5 days. Considering the low concentration of the polypeptide (150 μM) and its length (130 a.a), the inherent sensitivity of this experiment is good enough for practical use. Although the weaker *i−2* transfers were not always detected, most of the NH correlations still have unique bidirectional connectivities with both *i−1* and *i + 1* carbon chemical shift information. Figure 3b shows the counterpart strips from a conventional 3D HNCA experiment obtained with 0.35 mM uniformly labeled sample in one day. Comparing both experiments shows that, although the inherent sensitivity is weaker as expected, the additional sequential information is valuable. The hnCA-TOCSY-caNH spectrum shows three additional *i + 1* (two) and *i − 2* (one) correlations.

The hnCA-TOCSY-caNH experiment can be combined with a different type of chemical shift encoding. For example, in the hNca-TOCSY-caNH experiment, a nitrogen CT evolution is used instead of the $^{13}C^{\alpha}$ t_1 CT evolution which was shown in Fig. 1a. For the Hnca-TOCSY-caNH experiment, a conventional evolution was used for used for ^{1}H t_1 to achieve sufficient resolution along this dimension. In the resulting spectra, a ^{15}N-^{1}H pair of residue i is directly coupled to the adjacent H_N or N_H at positions $i-2$ to $i + 2$, respectively (Fig. 1c, d). Figure 4a, b shows the ω_2 ($^{15}N_H$)/ω_3($^{1}H_N$) and ω_1($^{1}H_N$)/ω_3($^{1}H_N$) strips, respectively from the 3D hNca-TOCSY-caNH and Hnca-TOCSY-caNH spectra. These experiments were recorded with the same experimental conditions as used in Fig. 2a.

The sequential connectivities observed in the hNca-TOCSY-caNH experiment are summarized in Fig. 4c. In the hNca-TOCSY-caNH experiment, 50/54 (92%) and 47/54 (87%) of ^{15}N-^{1}H pairs have correlations to N_{i-1} and N_{i+1}, respectively. As for Supra-sequential correlations, 20/53(38%) of N_{i-2}, and 13/53 (25%) of N_{i+2} correlations were observed. The number of supra-sequential N_{i-2} correlation is lower in comparison to the hnCA-TOCSY-NH experiments (C_{i-2}) presumably due to the fact that coherence exclusively originate from H_{i-2} in the nitrogen-coded experiment while an additional coherence from H_{i-1} can also contribute to C_{i-2} coherence in the carbon-coded experiment via a $^{2}J_{Ni-1Ci-2}$ coupling (Fig. 1b). As shown in Fig. 4d, the analysis of the 3D experiments with ^{15}N-^{1}H HSQC can easily identify the position of adjacent residues in the HSQC spectrum, making the assignment procedure straightforward. Although bidirectional sequential connectivities obtained from these experiments are useful to resolve resonance overlap, the information does not contain the direction in the amino acid sequence. Thus, complementary use with unidirectional experiments such as hNcocaNH and HncocaNH (Sun et al. 2005) would be beneficial.

Conclusion

We presented here a set of 3D experiments utilizing the long-range $^{3}J_{C\alpha C\alpha}$ couplings for magnetization transfer, in addition to the conventionally used one-bond ^{1}H-^{15}N and ^{15}N-^{13}C heteronuclear couplings. The experiments become feasible by using alternate ^{13}C-^{12}C labeling together with uniform $^{2}H^{15}N$ labeling. These set of experiments provides unique bidirectional correlations to both preceding ($i-1$) and succeeding ($i + 1$) $^{13}C^{\alpha}$ spins, which is useful to solve degeneracy problems associated with conventional 3D triple resonance experiments. As expected, the sensitivity for each coherence transfer in this experiment largely depends on the ^{13}C labeling pattern. Thus, for this experiment to have higher sensitivity, it would be of primary importance to improve labeling efficiency for those residue which are not 100% labeled with $^{13}C^{\alpha}$ as well as to avoid $^{1}J_{C\alpha C\beta}$ coupling. This can be achieved by a more elaborate labeling strategy using chemically synthesized $^{2}H^{\alpha}$, $^{13}C^{\alpha}$, $^{15}N_H$ labeled amino acids or corresponding precursors. To achieve this is currently pursued in our lab. The supra-sequential connections were observed for a significant number of residues. These explicit correlation signatures enable the correct association of longer backbone fragments with high fidelity. This feature is not only useful to make assignments more reliable avoiding degeneracy problems, but can also assist in automated assignment protocols. In case that the increased number of correlations cause severe overlapping of cross peaks in 3D experiments, the pulse sequences can easily be extended to a 4D hNCA-TOCSY-caNH experiment. The sensitivity of the 4D experiment would be only 30% less than 3D version as constant time evolution is used in both indirect dimensions. Furthermore, the experiments also can bypass single Pro residues that cause gaps in conventional sequential assignment procedures. This last point is important when dealing with regulatory domains of many large proteins, which often contain phosphorylation sites that typically include prolines (Ser-Pro, Thr-Pro) in addition to multiple glycines. Thus, these sites are difficult to assign in conventional ^{1}H-detected

experiments. The three sets of experiments proposed here would be particularly beneficial for these Pro-rich unstructured regulatory domains, which include functionally important proteins such as 4EBP, p53, NFAT, Ssdp1, etc. (Enkhmandakh et al. 2006; Marintchev et al. 2007; Rao et al. 1997; Ruaro et al. 1997; Venot et al. 1998).

The proton excided/detected experiments based on CACA-TOSY module described here would be the method of choice given the increased polarization of proton nuclei. However, in large molecular weight proteins, long transverse periods for alpha carbon and nitrogen (44 ms for each) might attenuate signal intensity. In addition, the original carbon-excited/carbon-detected version of the experiment has value for the assignment of side-chain resonances as it also provides C^{α} to C^{β}, C^{γ}, and C^{δ} correlations for certain amino acids (Takeuchi et al. 2010). The side-chain information would not be obtained from the HN excited/detected experiments shown here. However, one can also design a pulse scheme that excites side-chain carbons or carbon-attached protons and transfer the coherence to main-chain carbons by TOCSY mixing where the coherences are eventually recorded through the H_N coherences. This type of experiments would be particularly beneficial for assigning Ile and Leu resonances in large molecular weight proteins as their methyl resonances can be directly correlated to main chain resonances. In combination with NOESY transfer elements, the side chain experiment might be able to provide assignment and structure information on a single sample in a single experiment. These kinds of experiments, which demand a wider carbon excitation profile, are currently being explored in our laboratory.

Acknowledgments

This work was supported by the NIH (grants AI37581, GM47467 and EB 002026). M.G would like to thank the Human Frontier science Program (HFSP) for a postdoctoral fellowship.

References

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

Emsley L, Bodenhausen G. Optimization of shaped selective pulses for NMR using a quaternion description of their overall propagators. J Magn Reson. 1992; 97:135–148.

Enkhmandakh B, Makeyev AV, Bayarsaihan D. The role of the proline-rich domain of Ssdp1 in the modular architecture of the vertebrate head organizer. Proc Natl Acad Sci USA. 2006; 103:11631–11636. [PubMed: 16864769]

Ferentz AE, Wagner G. NMR spectroscopy: a multifaceted approach to macromolecular structure. Q Rev Biophys. 2000; 33:29–65. [PubMed: 11075388]

Frueh DP, Arthanari H, Wagner G. Unambiguous assignment of NMR protein backbone signals with a time-shared triple-resonance experiment. J Biomol NMR. 2005; 33:187–196. [PubMed: 16331423]

Frueh DP, Sun ZY, Vosburg DA, Walsh CT, Hoch JC, Wagner G. Non-uniformly sampled double-TROSY hNcaNH experiments for NMR sequential assignments of large proteins. J Am Chem Soc. 2006; 128:5757–5763. [PubMed: 16637644]

Goddard, TD.; Kneller, DG. SPARKY 3—NMR Assignment and Integration Software. University of California; San Francisco: 2006.

Grzesiek S, Bax A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. J Am Chem Soc. 1993; 115:12593–12594.

Grzesiek S, Anglister J, Bax A. Correlation of backbone amide and aliphatic side-chain resonances in ^{13}C/^{15}N-enriched proteins by isotropic mixing of ^{13}C Magnetization. J Magn Reson B. 1993; 101:114–119.

Hiller S, Arthanari H, Wagner G. The T-lock: automated compensation of radio-frequency induced sample heating. J Biomol NMR. 2009; 44:69–76. [PubMed: 19434373]

Kadkhodaie M, Rivas O, Tan M, Mohebbi A, Shaka AJ. Broadband homonuclear cross polarization using flip-flop spectroscopy. J Magn Reson. 1991; 91:437–443.

Lee D, Hilty C, Wider G, Wüthrich K. Effective rotational correlation times of proteins from NMR relaxation interference. J Magn Reson. 2006; 178:72–76. [PubMed: 16188473]

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

Lin Y, Wagner G. Efficient side-chain and backbone assignment in large proteins: application to tGCN5. J Biomol NMR. 1999; 15:227–239. [PubMed: 10677826]

Liu A, Riek R, Wider G, von Schroetter C, Zahn R, Wüthrich K. 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. 2000; 16:127–138. [PubMed: 10723992]

Logan TM, Olejniczak ET, Xu RX, Fesik SW. A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. J Biomol NMR. 1993; 3:225–231. [PubMed: 8477187]

Lyons BA, Tashiro M, Cedergren L, Nilsson B, Montelione GT. An improved strategy for determining resonance assignments for isotopically enriched proteins and its application to an engineered domain of staphylococcal protein A. Biochemistry. 1993; 32:7839–7845. [PubMed: 8394117]

Marintchev A, Frueh D, Wagner G. NMR methods for studying protein-protein interactions involved in translation initiation. Meth Enzymol. 2007; 430:283–331. [PubMed: 17913643]

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

Morris GA, Freeman R. Enhancement of nuclear magnetic resonance signals by polarization transfer. J Am Chem Soc. 1979; 101:760–762.

Muhandiram DR, Kay LE. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson B. 1994; 103:203–216.

Rao A, Luo C, Hogan PG. Transcription factors of the NFAT family: regulation and function. Annu Rev Immunol. 1997; 15:707–747. [PubMed: 9143705]

Ruaro EM, Collavin L, Del Sal G, Haffner R, Oren M, Levine AJ, Schneider C. A proline-rich motif in p53 is required for transactivation-independent growth arrest as induced by Gas1. Proc Natl Acad Sci USA. 1997; 94:4675–4680. [PubMed: 9114050]

Santoro J, King GC. A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. J Magn Reson. 1992; 97:202–207.

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

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

Sun ZY, Frueh DP, Selenko P, Hoch JC, Wagner G. Fast assignment of ^{15}N-HSQC peaks using high-resolution 3D HNcocaNH experiments with non-uniform sampling. J Biomol NMR. 2005; 33:43–50. [PubMed: 16222556]

Takeuchi K, Sun ZY, Wagner G. 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. 2008; 130:17210–17211. [PubMed: 19049287]

Takeuchi K, Frueh D, Sun Z, Hiller S, Wagner G. 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. J Biomol NMR. 2010; 47:55–63. [PubMed: 20383561]

Venot C, Maratrat M, Dureuil C, Conseiller E, Bracco L, Debussche L. The requirement for the p53 proline-rich functional domain for mediation of apoptosis is correlated with specific PIG3 gene transactivation and with transcriptional repression. EMBO J. 1998; 17:4668–4679. [PubMed: 9707426]

Vuister GW, Bax A. Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. J Magn Reson. 1992; 98:428–435.

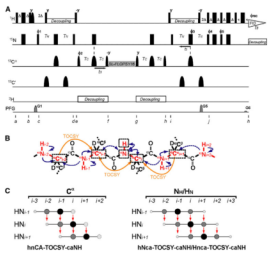

Fig. 1.

Experimental design for the 3D experiments utilizing the CA-CA TOCSY transfer. **a** Pulse scheme of the 3D hnCA-TOCSY-caNH experiment optimized for uniformly $^{2}H^{15}N$- and alternate ^{13}C-^{12}C -labeled samples. Narrow and wide rectangular black bars indicate non-selective $\pi/2$ and π pulses, respectively. Narrow and wide semi elliptical shapes on the carbon channel represent $\pi/2$ and π Gaussian cascades pulses selective for the frequencies of aliphatic carbon nuclei (Q5/256 μs and Q3/205 μs, respectively) (Emsley and Bodenhausen 1992). A selective ^{1}H $\pi/2$ pulse was used for water flip-back during the first INEPT transfer (Grzesiek and Bax, 1993). All pulses are applied along the x-axis unless otherwise indicated. The delays are $\Delta = 2.7$ ms, $T_N = T_C = 11$ ms, and $\tau = 1.2$ ms The phase cycle employed was $Ø_1$ = (x, x, x, x, x, x, x, x, −x, −x, −x, −x, −x, −x, −x, −x), $Ø_2$ = (x, −x), $Ø_3$ = (x, x, x, x, −x, −x, −x, −x), $Ø_4$ = (x, x, −x, −x), $Ø_5$ = (−y,−y, y, y), and $Ø^{rec}$ = (x, −x, −x, x, x, −x, −x, x, −x, x, x, −x, −x, x, x, −x). It is possible to run this experiment with a shorter phase cycle, however, the larger number would generally be better to avoid artifacts. The experimental time for the hnCA-TOCSY-caNH experiment with the 16-step phase cycling was 44 h, which does not exceed the practical experimental time for 3D experiments. Phase sensitive spectra in the indirect C^{α} dimension (t_1) are obtained by incrementing the phases Ø1 in a States-TPPI manner (Marion et al. 1989). For the indirect N_H dimension (t_2), echo and anti-echo coherences are obtained by recording two data sets, whereby the sign of the gradient G5 and phase $Ø_5$ are inverted for the second data set. At the same time, non-^{15}N coupled coherence, such as water, was suppressed by PFG gradient selection. The recycling delay was set to 1 s. The two sine-shaped pulsed field gradients were applied along the *z*-axis for 1.0 ms with maximum intensities of G1 = 15, G2 = 40 and, G3 = 4.05 G/cm. FLOPSY-16 was applied for spin locking (3.2 kHz)(Kadkhodaie et al. 1991). Proton, deuterium and nitrogen decoupling are achieved by using WALTZ16 (Shaka et al. 1983) for proton and deuterium (3.3 and 1 kHz, respectively) and GARP (Shaka et al. 1985) for nitrogen (1 kHz). **b** Illustration of the coherences utilized in the 3D hnCA-TOCSY-caNH experiment. The nuclei involved in this experiment are colored in red. Arrows indicate the coherence transfer pathways in the hnCA-TOCSY-caNH experiment, which are eventually detected as H_{Ni}. Blue solid and dotted arrows indicate coherence transfers via $^{1}J_{NC\alpha}$ and $^{2}J_{NC\alpha}$ heteronuclear scalar couplings. On the other hand, yellow arrows indicate coherence transfers via $^{3}J_{C\alpha C\alpha}$ TOCSY. The nuclei that are correlated to H_{Ni} are boxed with black broken lines. (C) Schematic representation of the observed coherences in the 3D hnCA-TOCSY-caNH experiment. Larger and darker peak sizes indicate higher intensities

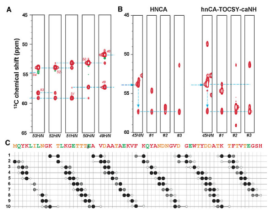

Fig. 2.

Sequential backbone assignments with the 3D hnCA-TOCSY-caNH experiment. **a** Representative $\omega_1(^{13}C^{\alpha})/\omega_3(^{1}H_N)$ strips from the 3D hnCA-TOCSY-caNH spectrum of GB1 connecting Thr49–Thr53. The experiment was recorded on a Bruker AVANCE 600 spectrometer equipped with a cryogenic probe, using a 3.5 mM sample of the uniformly $^{2}H^{15}N$ and alternately ^{13}C-^{12}C labeled B1 domain of protein G (GB1) in H_2O. To simulate the tumbling of a 40 kDa protein, 40% glycerol was added, and the experiment was recorded at 298 K. Sequential correlation peaks are shown by broken lines. The 3D hnCA-TOCSY-caNH experiment was recorded with 4,839 Hz (^{13}C, indirect) × 2,189 Hz (^{15}N, indirect) × 9,615 Hz (^{1}H, direct), centered at 55 ppm (^{13}C), 118 ppm (^{15}N), and 4.7 ppm (^{1}H). A total of 64 (^{13}C, indirect) × 28 (^{15}N, indirect) × 512 (^{1}H, direct), complex data points were recorded for indirect and direct directions, respectively, with $T_C = T_N = 11$ ms. For each increment, 16 scans were acquired. A cosine apodization function was applied to each FID, and the direct dimensions were zero-filled up to 2,048 data points before Fourier transform. **b** Assignment procedure by the complementary use of HNCA and hnCA-TOCSY-caNH experiments. The HNCA $\omega_1(^{13}C^{\alpha})/\omega_3(^{1}H_N)$ strips at the ^{15}N position of Tyr45 are shown in the left panel. The horizontal arrow indicates the chemical shift of Tyr45 $^{13}C^{\alpha}$, whereas the vertical arrow connects Tyr45 $^{13}C^{\alpha}$ to Thr44 $^{13}C^{\alpha}$. By chemical shift matching, three strips (#1 ~ #3) containing a signal corresponding to the Thr44 $^{13}C^{\alpha}$ chemical shift (57.4 ppm) are identified as sequential candidates with ^{15}N chemical shifts of 113.8, 111.6 and 115.5 ppm. The chemical shift of the $^{13}C_{i-1}$ signal is completely overlapped with three candidate strips in the HNCA spectrum, thus the correct sequential neighbor cannot be determined. In the hnCA-TOCSY-caNH experiment (right panel), only the #1 strip has the $^{13}C^{\alpha}$ signal corresponding to Tyr45 $^{13}C^{\alpha}$, providing the unambiguous assignment of strip #1 to Thr44. **c** Summary of observed cross peaks in the 3D hnCA-TOCSY-caNH spectrum. The vertical positions correspond to sequence numbers of HN pairs while horizontal axis shows the presence of $^{13}C^{\alpha}$ signals. The size and darkness of the circles indicate the signal to noise ratio (S/N) of each cross peak estimated by SPARKY. Four levels of size and darkness of circles are employed. Big black, middle size gray, small light gray and small white circles indicate S/N ranges of >100, >50,>10 and>5, respectively. The measurement time for the hnCA-TOCSY-caNH experiments was 44 h

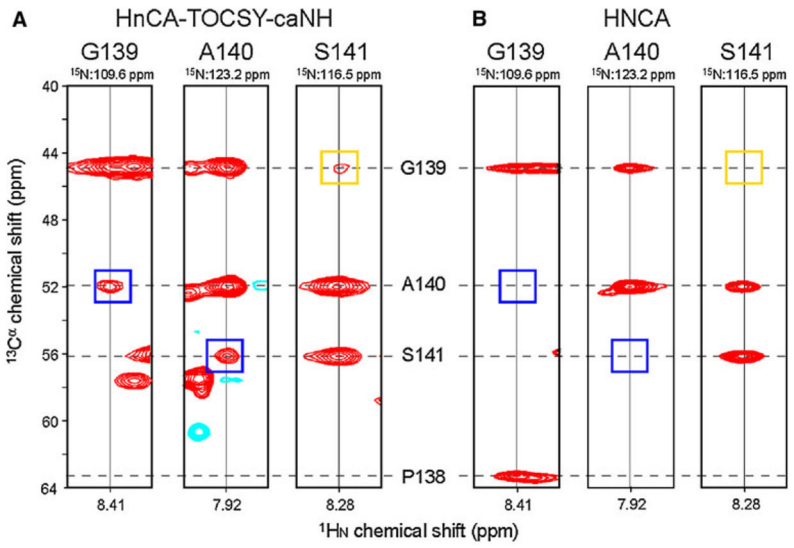

Fig. 3.

Representative strips from the 3D hnCA-TOCSY-caNH experiment acquired on the 17 kDa unfolded regulatory domain of nuclear factor of activated T cells (NFAT). **a** strips originated from 3D hnCA-TOCSY-caNH experiment and **b** from conventional 3D HNCA experiment. The strips connect G139–Ser141. The hnCA-TOCSY-caNH experiment was recorded on a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenic probe, using a 0.15 mM sample of uniformly $^{2}H^{15}N$ and alternately ^{13}C-^{12}C labeled protein in H_2O-based buffer. The 3D HNCA experiment was recorded on 0.35 mM sample of uniformly ^{15}N ^{13}C labeled protein in the same buffer. Chemical shifts for P138–Ser141 C^{α} carbons are shown by broken lines. The 3D hnCA-TOCSY-caNH experiment was recorded with 4,829 Hz (^{13}C, indirect) × 1,642 Hz (^{15}N, indirect) × 9,615 Hz (^{1}H, direct), centered at 55 ppm (^{13}C), 118 ppm (^{15}N), and 4.7 ppm (^{1}H). 40 (^{13}C, indirect) × 55 (^{15}N, indirect) × 1,152 (^{1}H, direct), complex data points were recorded for indirect and direct directions, respectively. For each increment, 32 scans were acquired. The 3D HNCA experiment was recorded with the same SW's. 60 (^{13}C, indirect) × 52 (^{15}N, indirect) × 2,048 (^{1}H, direct), complex data points were recorded for indirect and direct directions, respectively. For each increment, 4 scans were acquired. A cosine apodization function was applied to each FID, and the direct dimensions were zero-filled before Fourier transform. The measurement time for the hnCA-TOCSY-caNH and HNCA experiments was 111 and 26 h, respectively. *Blue* and *orange* squares in the spectrum indicate the positions of $i + 1$ and $i - 2$ resonances that are only observed in the hnCA-TOCSY-caNH strips

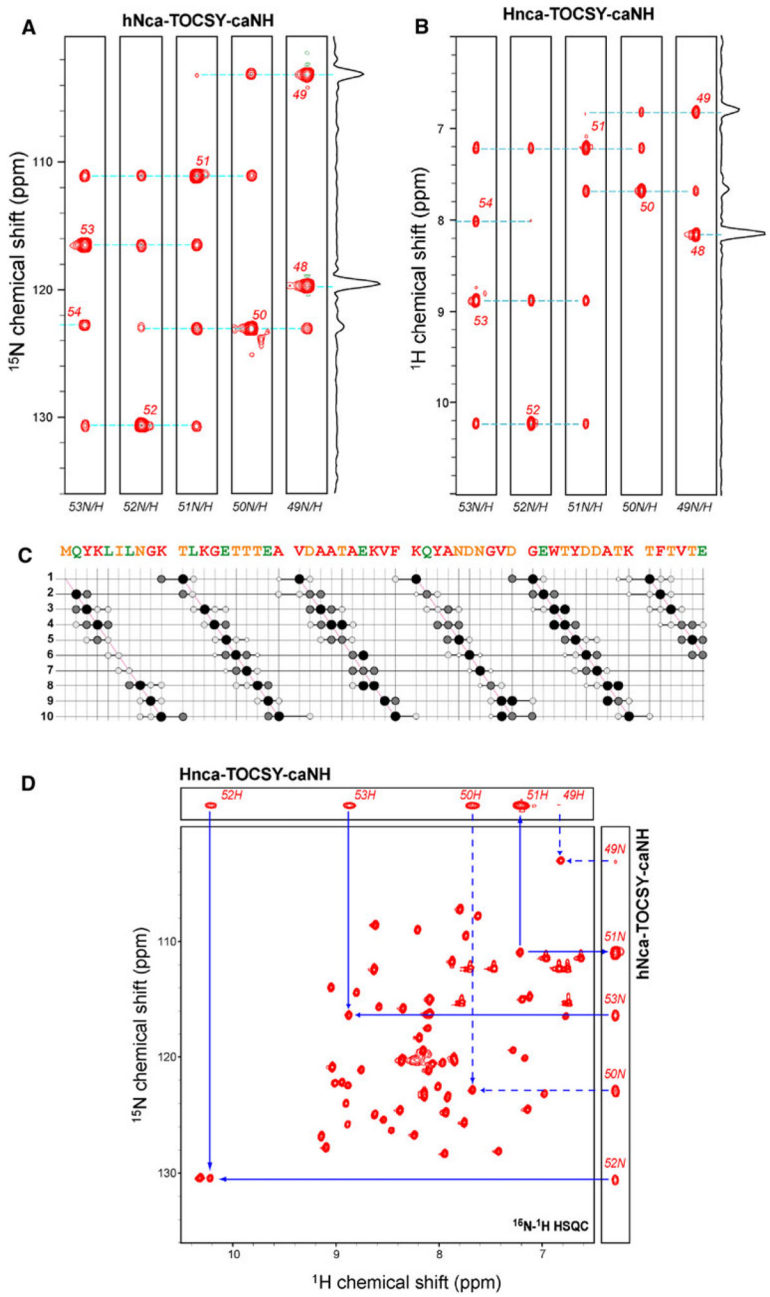

Fig. 4.

Correlations in 3D hNca-TOCSY-caNH and Hnca-TOCSY-caNH spectra: (**a** and **b**) Representative ω_1 ($^{15}N_H$)/ω_3(^1H_N) and ω_1(^1H_N)/ω_3(^1H_N) strips from **a** a 3D hNca-TOCSY-caNH and **b** Hnca-TOCSY-caNH spectra of GB1, respectively. The strips of Thr49–Thr53 are shown. The experiment was performed on a Bruker AVANCE 600 spectrometer equipped with a cryogenic probe, using a 3.5 mM sample of the uniformly $^2H^{15}N$ and alternately ^{13}C-^{12}C labeled B1 domain of protein G (GB1) in H_2O. To simulate the tumbling of a 40 kDa protein, 40% glycerol was added, and the experiment was recorded at 298 K. Sequential connections are shown by broken lines. Both experiments were recorded with a spectral width of 9,615 Hz (^1H, direct), 2,189 Hz (^{15}N, indirect), and/or 3,600 Hz (^1H, indirect). The ^1H and ^{15}N frequencies were centered at 8.0 (^1H, indirect), and 118 (^{15}N)

ppm, and 4.7(^{1}H, direct), respectively. The 3D hNca-TOCSY-caNH experiment was recorded with 45 (^{15}N, indirect, t_1) × 28 (^{15}N, indirect, t_2) × 512 (^{1}H, direct), complex data points were recorded. For the 3D Hnca-TOCSY-caNH experiment, 64 (^{1}H, indirect) × 28 (^{15}N, indirect) × 512 (^{1}H, direct), complex data points were recorded. For each increment, 16 scans were acquired. A cosine apodization function was applied to each FID, and the direct dimensions were zero-filled up to 2,048 data points before Fourier transform. **c** Summary of observed cross peaks in the 3D hNca-TOCSY-caNH spectrum. The vertical positions correspond to sequence numbers of HN pairs while the horizontal axis shows the presence of ^{15}N signals. The size and darkness of the circles are the same as in Fig. 2c d Identification of residues adjacent to Thr51. The $i - 1$ and $i + 1$ correlations are shown with solid arrows pointing to ^{1}H-^{15}N cross peaks in the HSQC spectrum. The supra-sequential $i - 2$ and $i + 2$ correlations were shown in broken arrows. Measurement times for the hNca-TOCSY-caNH and Hnca-TOCSY-caNH experiments were 34 and 44 h, respectively