COMMUNICATION

Identification of C-terminal neighbours of amino acid residues without an aliphatic $^{13}C^{\gamma}$ as an aid to NMR assignments in proteins

Ravi Pratap Barnwal · Ashok K. Rout · Hanudatta S. Atreya · Kandala V. R. Chary

Abstract We propose a methodology that uses GFT (3,2)D CB(CACO)NNH experiment to rapidly collect the data and readily identify six amino acid residue types (Ala, Asn, Asp, Cys, Gly and Ser) in any given protein. Further, the experiment can distinguish the redox state of Cys residues. The proposed experiment in its two forms will have wide range of applications in resonance assignment strategies and structure determination of proteins.

Keywords Isotope labeling · NMR · Automated assignments · Sequence specific resonance assignments

Introduction

Three dimensional (3D) structure determination of medium-sized proteins (up to 30 kDa) using triple resonance NMR techniques has been greatly aided by the ability to label these macromolecules with ^2H, ^{13}C and ^{15}N. Sequence specific resonance assignments remain an important and essential step towards their complete 3D structural characterization (Cavanagh et al. 2006; Chary

R. P. Barnwal · A. K. Rout · K. V. R. Chary (✉)
Department of Chemical Sciences, Tata Institute of Fundamental Research, Colaba, Mumbai 400005, India
e-mail: chary@tifr.res.in

H. S. Atreya (✉)
NMR Research Centre, Indian Institute of Science, Malleshwaram, Bangalore 560012, India
e-mail: hsatreya@sif.iisc.ernet.in

and Girjesh 2008). During the last two decades, several double and triple resonance experiments have been proposed to carry out sequence specific ^1H, ^{13}C, and ^{15}N NMR assignments in isotope labeled proteins (Bax and Grzesiek 1993; Sattler et al. 1999). Despite the demonstrated utility of such techniques for the structural characterization of proteins, one encounters several problems in the resonance assignment procedure, when these techniques are applied to large proteins with molecular weights in excess of 10 kDa. Besides, the most common problem of spectral overlap and rapid relaxation rates of the nuclei result in the broadening of several cross peaks thus hampering the resonance assignments. Pro residues which lack $^1H^N$ further aggravate the assignment problem. This prompts one to have as many good starting points as possible along the polypeptide chain of a given protein. Ala, Gly, Ser and Thr residues have been the most easily identifiable amino acid residues, because of their characteristic $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts (Atreya et al. 2000). The $^{13}C^{\alpha}$ chemical shift of Gly residue (Gly($^{13}C^{\alpha}$)) always resonates around 45.3 ± 5.4 ppm in a region well separated from the $^{13}C^{\alpha}$ chemical shifts of all other residues and thus helps in their identification. On the other hand, Ala($^{13}C^{\beta}$), Ser($^{13}C^{\beta}$) and Thr($^{13}C^{\beta}$) with their characteristic chemical shifts around 18.9 ± 7.5, 63.8 ± 6.4 and 69.6 ± 8.1 ppm, respectively, in regions well seperated from the $^{13}C^{\beta}$ chemical shifts of the rest of the residues, aid in their unambiguous identification. This characterization is based on the complete chemical shift data of proteins available with BioMagResBank (BMRB; http://www.bmrb.wisc.edu). Further, the fact that on an average, the percentage composition of Ala, Gly, Ser and Thr residues together in any given protein amounts to as much as $25.8 \pm 5.8\%$, it should be possible to complete the sequence specific resonance assignments in the protein

with these residues as starting points. In practice, however, even with four triple resonance spectra such as HNCACB (Wittekind and Mueller 1993), CBCA(CO)NH (Grzesiek and Bax 1992a, b), HNCO (Kay et al. 1990) and HN(CA)CO (Clubb et al. 1992), which provide information about H^N, ^{15}N, $^{13}C^\alpha$, $^{13}C^\beta$ and $^{13}C'$ chemical shifts, the success rate in many automated resonance assignment turns out to be low. Thus, there has been a need for unambiguous identification of as many peaks as those belonging to specific amino acid residues, other than those which are easily identifiable in various triple resonance spectra.

Towards this end, two different approaches have evolved. In one approach, selective identification of amino acid type is achived by specific labeling (Muchmore et al. 1989) or "unlabeling" (Atreya and Chary 2001). However, this requires preparation of more than one sample.

In the second approach, identification of amino-acid type is achieved by several amino-acid type selective triple resonance experiments (Dotsch et al. 1996a, b; Dotsch and Wagner 1996; Schubert 2001), with the need for only one uniformly ^{13}C and ^{15}N doubly labeled sample. These experiments select self and sequential connectivities of a select amino-acid(s) in any given protein. For example, a selective CBCA(CO)NH pulse sequence was designed to select connectivities to the C-terminal sequential neighbours of amino acid residues without an aliphatic $^{13}C^\gamma$ (Ala, Asp, Asn, Cys, Gly, and Ser) (Dotsch et al. 1996a, b; Dotsch and Wagner 1996). This experiment was subsequently modified to additionally suppress the sequential connectivities to the C-terminal neighbours of Asp and Asn by an application of a selective π pulse on the carbonyl spins during a delay in the pulse sequence used for transferring polarization from $^{13}C^\beta$ to $^{13}C^\alpha$. This resulted in further simplification of the spectrum by observing sequential connectivities to the C-terminal neighbours of Ala, Cys, Gly and Ser residues alone. In such attempts, though one could simplify the spectrum, it was not straightforward to identify peaks belonging to a specific amino acid type within the given set of selected residues. To overcome this problem, a nonselective 3D CBCA(CO)NH was used to distinguish the peaks. For example, while the sequential connectivities arising from Ala and Ser residues can be identified from their unique $^{13}C^\beta$ chemical shifts in 3D CBCA(CO)NH, the sequential connectivities for Gly residues can be identified from their unique $^{13}C^\alpha$ chemical shift and simultaneously the absence of the peak due to $^{13}C^\beta$ chemical shift.

In this backdrop, we propose an alternate methodology that uses the principle of G-matrix Fourier transform NMR (Atreya and Szyperski 2005; Kim and Szyperski 2003, 2004; Szyperski and Atreya 2006) to rapidly collect the data and identify peaks arising from six different amino

acid types in a given protein: (i) ACS-(3,2)D CB(CACO)NNH for identification of Ala, Cys and Ser and (ii) ACSDNG-(3,2)D CB(CACO)NNH for identification of Ala, Cys, Ser, Asp, Asn and Gly amino acid residues. The idea is built upon the β-carbon-edited form of CBCA(CO)NH experiment described above (Dotsch and Wagner 1996) which provides sequential connectivities for amino acids lacking $^1J(^{13}C^\beta–^{13}C^\gamma)$. In its new form as proposed here, the (3,2)D CB(CACO)NNH experiment rapidly provides an additional information about the $^{13}C^\beta$ chemical shifts. This information aids in straight forward identification of the observed correlations as belonging to specific amino acid residues and can be incorporated readily into resonance assignment procedures.

Figure 1 shows radio-frequency (RF) pulse scheme for the proposed (3,2)D CB(CACO)NNH experiment in its two forms described above. For the nuclei shown underlined ($^{13}C^\beta$ and ^{15}N), chemical shifts are jointly sampled, thereby providing 3D spectral information in the form of a 2D spectrum rapidly with high precision (Atreya and Szyperski 2004, 2005; Kim and Szyperski 2003; Szyperski and Atreya 2006). Phase-sensitive joint-sampling of $^{13}C^\beta$ and ^{15}N chemical shifts is achieved by co-incrementing their respective chemical-shift evolution periods with the $^{13}C^\beta$ shifts scaled by a factor 'κ' relative to ^{15}N (Atreya and Szyperski 2005; Szyperski and Atreya 2006). This results, after G-matrix transformation, in two sub-spectra each comprising of peaks at a given linear combination of chemical shifts along the indirect dimension (t_1): ω_1: $\Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{13}C^\beta_i)$. Peaks corresponding to $^{13}C^\alpha$ of all residues are suppressed by setting the $^{13}C^\alpha$–$^{13}C^\beta$ delay to $1/2J_{CC}$. In ACS-(3,2)D CB(CACO)NNH, the delays are adjusted such that in addition to $^{13}C^\alpha$–$^{13}C^\beta$ polarization transfer period of $1/2J_{CC}$, $^{13}C^\alpha$–$^{13}C'$ coupling is kept active for a total duration of $1/2J_{CC'}$ (~ 10 ms). This results in absence of peaks belonging to ^{13}C spins directly attached to the $^{13}C'$ such as $^{13}C^\alpha$ of all residues and $^{13}C^\beta$ of Asp and Asn residues. In ACSDNG-(3,2)D CB(CACO)NNH, the selective pulse on $^{13}C'$ is removed and the pulses on $^{13}C^{\alpha/\beta}$ are adjusted to minimally effect the $^{13}C'$ and $^{13}C^\gamma$ spins of aromatic residues (Fig. 1). Thus, the resultant (3,2)D spectrum shows connectivities to the C-terminal sequential neighbours of all amino acid residues without an aliphatic $^{13}C^\gamma$ (i.e., Ala, Asp, Asn, Cys, Gly, and Ser). As discussed earlier, the scaling factor, κ, allows one to increase the dispersion of peaks or to restrict the chemical shift evolution of $^{13}C^\beta$ to avoid loss in sensitivity due to transverse relaxation during the t_1 (Atreya and Szyperski 2004; Atreya and Szyperski 2005). An additional 2D [^{15}N, ^1H] HSQC recorded with the same sample provides central peak information (ω_1: $\Omega(^{15}N)$) needed to analyse the data (Atreya and Szyperski 2005; Kim and Szyperski 2003). The observed cross-peaks are characterized by the chemical

shift information of $^1H^N_{i+1}$, $^{13}C^{\alpha/\beta}_i$ and $^{15}N_{i+1}$. The extra information about the $^{13}C^\beta_i$ chemical shift thus derived aids in identification of the edited correlations as belonging to two of the six amino acid residues (Ala and Ser), without resorting to record a non-selective 3D CBCA(CO)NH. To illustrate this point, we have simulated two sub-spectra of ACS-(3,2)D CB(CACO)NNH (Fig. 2b) each comprising of peaks at $(\omega_1, \omega_2) = [\Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{13}C^\beta_i), \Omega(^1H^N_{i+1})]$ for only the Ala, Cys$^{oxy/red}$ and Ser residues (Fig. 2a) and two sub-spectra of (3,2)D CB(CACO)NNH (Fig. 2b) each comprising of peaks at $(\omega_1, \omega_2) = [\Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{13}C^\beta_i), \Omega(^1H^N_{i+1})]$ for all the Ala, Asp, Asn, Cys$^{oxy/red}$, Gly $[(\omega_1, \omega_2) = \Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{13}C^\alpha_i)]$, and Ser residues. This characterization is based on the statistical analysis of chemical shift data of proteins available in the BioMagResBank (BMRB, http://www.bmrb.wisc.edu) which includes a database of 235 proteins. A total number of 5355 $^1H^N_{i+1}$, 1205 $^{13}C^\alpha_i$/4150 $^{13}C^\beta_i$ and 5355 $^{15}N_{i+1}$ chemical shifts has been used in this analysis. As is evident from these simulations, the sequential connectivities arising from two of the six filtered residues (Ala and Ser) can be readily identified from their unique $^{15}N \pm {}^{13}C^\beta$ chemical-shift signatures. The correlations arising from the Asn, Asp, Cys$^{oxy/red}$ and Gly, however, show significant overlap, hampering their individual identification. To circumvent this problem, we adopted the method of Dotsch et al. (Dotsch et al. 1996a; Dotsch and Wagner 1996) as described above to additionally suppress the sequential connectivities to the C-terminal neighbours of Asp and Asn by an application of a selective π pulse on the carbonyl spins during the delay $2T_{AB}$ (i.e., $\tau_3 + \tau_4 + \tau_5$ in Fig. 1; ACS-(3,2)D CB(CACO)NNH). The peaks corresponding to $^{13}C^\alpha$ spins are suppressed due to the fact that they are coupled to backbone $^{13}C'$. This results in further simplification of the spectrum by observing sequential connectivities to the C-terminal neighbours of Ala, Cys$^{oxy/red}$, and Ser residues alone. As is evident from these simulations, the sequential connectivities arising from these three residues can be readily identified from their unique $^{15}N_{i+1} \pm {}^{13}C^\beta_i$ chemical shift signatures without resorting to record a non-selective CBCA(CO)NH experiment. Further, this spectrum aids to discern the redox state of the Cys residues based upon their $^{13}C^\beta$ chemical shifts (Atreya et al. 2000). Once the sequential connectivities arising from Cys residues are discerned, one can distinguish the signatures arising from Asn, Asp and Gly residues. Out of these, the sequential connectivities for Gly residues can be identified from their unique $^{13}C^\alpha$ chemical shift and also the absence of the peak due to $^{13}C^\beta$ chemical shifts.

Thus, the (3,2)D CB(CACO)NNH spectrum in its two forms described here can be used to identify connectivities to the C-terminal sequential neighbours of amino acid residues without an aliphatic $^{13}C^\gamma$ (Ala, Asp/Asn, Cys, Gly, and Ser). This (3,2)D GFT NMR experiment further provides several other advantages. (i) The spectral information can be obtained rapidly with good sensitivity (Supplementary Table S1). The measurement time can be further reduced in combination with other methods for fast data collection (Atreya and Szyperski 2005). (ii) Data are acquired in the form of a two-dimensional spectrum facilitating high spectral/digital resolution (Supplementary Table S1). (iii) Spectra have high dispersion due to joint sampling of ^{15}N and $^{13}C^{\alpha/\beta}$ shifts. And, (iv) most importantly, there is no need to collect non-selective 3D CBCA(CO)NH to distinguish the peaks arising from Ala, Cys and Ser residues.

To demonstrate the utility of the proposed (3,2)D CB(CACO)NNH in its two forms, we recorded the spectra for the following samples: (i) an u-$^{13}C/^{15}N$-doubly labeled calbindin sample (calbindin; 9 kDa; 1.5 mM) without and with a paramagnetic landthanide ion (Bertini et al. 2001). For this purpose, two uniformly ^{13}C and ^{15}N doubly labeled calbindin samples, one with Ca^{2+}-bound to both N- and C-terminal sites ([Ca^{2+}]$_2$Cb) and the other Yb^{3+}-substituted in the C-terminal site ([Ca^{2+}][Yb^{3+}]Cb) were purchased from ProtEra SRL (http://www.protera.it). Yb^{3+} was chosen as the paramagnetic metal ion due to its favorable PCS/line-broadening ratio compared to other Ln^{3+} (Allegrozzi et al. 2000; Atreya et al. 2003; Mustafi et al. 2004); (ii) u-$^{13}C/^{15}N$-doubly labeled dsmt3 in its completely unfolded form (10 kDa; 1.0 mM) (Kumar et al. 2007); and (iii) an u-$^{13}C/^{15}N$-doubly labeled (predominantly) α-helical protein (Eh-CaBP; 16 kDa; 1.0 mM) (Atreya et al. 2001; Sahu et al. 1999). These systems were chosen to demonstrate the feasibility of the proposed experiment with different types and sizes of proteins in their different states.

NMR experiments were performed at 25°C on a Varian Inova 600 MHz spectrometer equipped with a cryogenic probe with four different kinds of samples mentioned above. The RF pulse scheme for (3,2)D CB(CACO)NNH in its two forms is shown in Fig. 1. The scaling factor, κ, was set to 0.5 for all the experiments discussed here. The total measurement time for the spectra ranged from 0.3 to 1.3 h (acquisition parameters are provided in Supplementary Table S1). The data were pre-processed with the G-matrix using in-house written scripts (using the methodology outlined in refs. Atreya and Szyperski 2004, 2005; Kim and Szyperski 2003) and subsequently processed with NMRPipe (Delaglio et al. 1995) and analyzed using XEASY (Bartels et al. 1995) and CARA (Keller 2004).

Figure 3 shows ACS-(3,2)D CB(CACO)NNH (Fig. 3a) and ACSDNG-(3,2)D CB(CACO)NNH (Fig. 3b) spectra recorded for [Ca^{2+}]$_2$-calbindin. The 75 amino-acid residue long Cys-free calbindin has 2 Ala, 2 Asn, 4 Asp, 5 Gly and 6 Ser residues. As is evident from Fig. 3a, we could readily
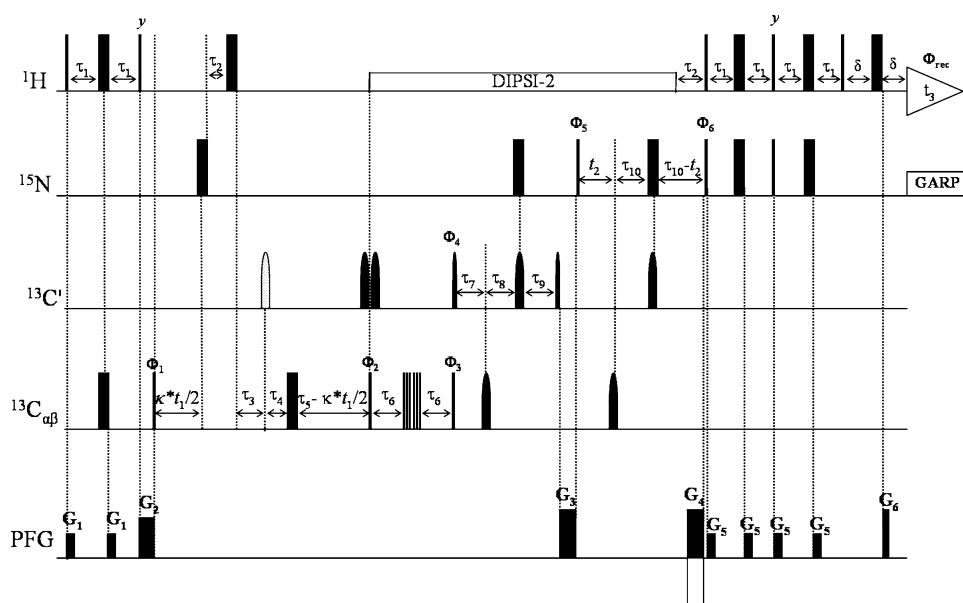
Fig. 1 R.f. pulse scheme of GFT (3,2)D CB(CACO)NHN for selective detection of Ser, Ala, Cys, Asp, Asn and Gly residues. Rectangular 90° and 180° pulses are indicated by thin and thick vertical bars, respectively, and phases are indicated above the pulses. Where no r.f. phase is marked, the pulse is applied along x. High-power 90° pulse lengths are: 5.7 μs for ^{1}H, 33 μs for ^{15}N and 15.8 μs for 13C. $\kappa = 0.5$ (see text). Pulses on ^{13}C prior to $t_1(^{13}$C) are applied at high power. Subsequently, the 90° and 180° pulse lengths applied on ^{13}C$^{\alpha\beta}$ (at a ^{1}H resonance frequency of 600 MHz) are adjusted to 48 and 43 μs, respectively, during the first t_1 delay [$t_1(^{13}$C$^{\alpha\beta})$]. All pulses applied on ^{13}C′ are of Sinc shape. Off-resonance 180° pulses with a length of 80.8 μs is used for decoupling and during polarization transfer periods involving ^{13}C′. For selective detection of Asn, Asp and Gly residues in addition to observing Ser, Ala and Cys, the pulse on ^{13}C′ shown as hashed is removed. The duration of 90° pulse applied on-resonance at ^{13}C′ is 89.2 μs. Off-resonance sinc-180° pulses of duration 48.6 μs is used for decoupling and during the polarization transfer periods. DIPSI (Shaka et al. 1988) is employed to decouple ^{1}H (r.f. field strength = 6.0 kHz) during the heteronuclear magnetization transfers. A six-pulse composite sequence (Shaka

1985) is used for ^{13}C$^\alpha$–^{13}C′ polarization transfer. The ^{1}H r.f. carrier is placed at the position of the solvent line at 4.7 ppm. The ^{15}N carrier position is set to 118.5 ppm. The ^{13}C r.f. carrier is kept at 43 ppm throughout the sequence. GARP (Shaka et al. 1985) is employed to decouple ^{15}N (r.f. = 1.50 kHz) during acquisition. All pulsed z-field gradients (PFGs) are rectangular shaped with gradient recovery delay of 100 μs. The duration and strengths of the PFGs are: G1 (0.5 ms, 8 G/cm); G2 (1.0 ms, 22 G/cm); G3 (1.0 ms, 22 G/cm); G4 (1.25 ms, 30 G/cm); G5 (0.5 ms, 8 G/cm) and G6 (0.125 ms, 30 G/cm). The delays are: $\tau_1 = 2.3$ ms, $\tau_2 = 850$ μs, $\tau_3 = 4.1$ ms, $\tau_4 = 2.2$ ms, $\tau_5 = 7.2$ ms, $\tau_6 = 4.4$ ms, $\tau_7 = 4.4$ ms, $\tau_8 = 8.1$ ms, $\tau_9 = 12.5$ ms, $\tau_{10} = 12.5$ ms and $\delta = 300$ μs. Phase cycling: $\varphi_1 = x$; $\varphi_2 = x$; $\varphi_3 = x, -x$; $\varphi_4 = x$; $\varphi_5 = x$; $\varphi_6 = x$; φ_8(receiver) = $x, -x$. Quadrature detection in $t_1(^{15}$N) is accomplished using the sensitivity enhanced scheme (Kay et al. 1992), that is, by inverting the sign of gradient G4 in concert with phases φ_5 (i.e., $\varphi_5 = x, -x$). GFT NMR phase-cycle: $\varphi_1 = x$, y yields, in conjunction with quadrature detection in $t_1(^{15}$N), two data sets which are linearly combined employing a G-matrix transformation with the G-matrix (Atreya and Szyperski 2005; Kim and Szyperski 2003)

identify all the expected connectivities to the C-terminal sequential neighbours of Ala and Ser residues present in the Cys-free protein. It may be noted here that, for the ACS case, a simple selective 2D CB(CACON)H or 2D (CBCACO)NH experiments without the combination of ^{13}C$^\beta$ and ^{15}N chemical shift evolution periods cannot replace a 3D experiment (such as the (3,2)D GFT experiments described here) wherein these two chemical shifts are correlated. On the other hand, acquiring a full 3D spectrum will require long minimal measurement time.

On the other hand, Fig. 3b shows the expected extra signatures arising from Asn, Asp and Gly residues. The peaks arising from all the ten Gly residues could be readily distinguished from that of Asn/Asp peaks from their unique ^{13}C$^\alpha$ chemical shifts. It is worth mentioning here that one can not deconvolute the signatures of Asp and Asn because of their degenerate ^{13}C$^\beta$ chemical shifts. However, they

could be distinguished from the signatures arising from Ala, Ser and Cys, which itself helps to a large extent in resonance assignments.

In order to further assess the sensitivity of these experiments, spectra were recorded with a larger molecular-weight protein (a calcium binding protein from *Entamoeba Histolytica*; predominantly α-helical protein; abbreviated as *Eh*-CaBP; 16.2 kDa with u-^{13}C/^{15}N labeling; 134 amino acid residues). All the expected connectivities to the C-terminal sequential neighbours of all Ala, Asn, Asp, Gly and Ser residues were observed in the protein (See Supplementary Figure S1).

Figure 3c and d show the (3,2)D CB(CACO)NNH spectra recorded in its two forms, for dsmt3 ($\kappa = 0.5$) in its completely urea denatured state (unfolded form), where the ^{1}H^N chemical shift dispersion is very narrow (0.6–0.8 ppm) (Dyson and Wright 2004). This 90 amino-acid residue long

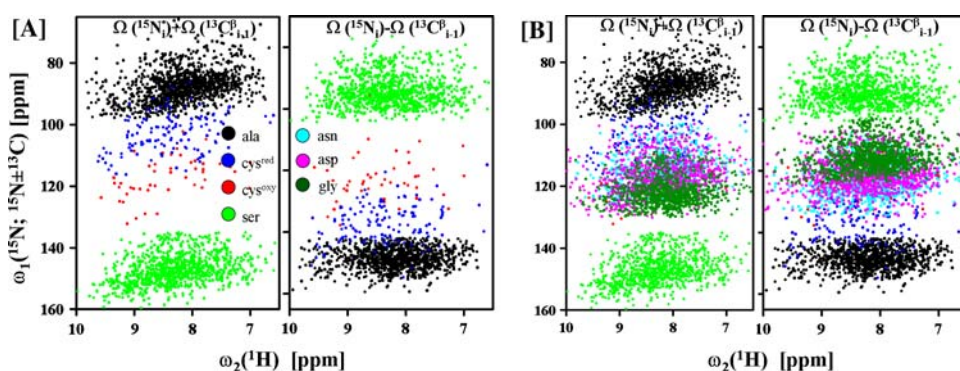**Fig. 2** Simulation of two sub-spectra of (3,2)D CB(CACO)NNH. (**a**) each comprising of peaks at $(\omega_1, \omega_2) = [\Omega(^{15}N_{i+1}) \pm \kappa*\Omega(^{13}C^\beta_i)]$ for only the Ala, Cys$^{oxy/red}$ and Ser residues. The scaling factor κ is set to 0.5. This characterization is based on the statistical analysis of chemical shift data of proteins available in the BioMagResBank (BMRB; http://www.bmrb.wisc.edu) which includes a database of 235 proteins. (**b**) each comprising of peaks at (ω_1, ω_2) $[\Omega(^{15}N_{i+1}) \pm \kappa*\Omega(^{13}C^\beta_i)]$ for all the Ala, Asp, Asn, Cys$^{oxy/red}$, Gly $[(\omega_1, \omega_2) = \Omega(^{15}N_{i+1}) \pm \kappa*\Omega(^{13}C^\alpha_i)]$ and Ser residues

Fig. 3 Plots of (**a**) ACS-(3,2)D CB(CACO)NNH and (**b**) ACSDNG-(3,2)D CB(CACO)NNH spectra of Calbindin sample ($\kappa = 0.5$) and (**c**) ACS-(3,2)D CB(CACO)NNH and (**d**) ACSDNG-(3,2)D CB(CACO)NNH spectra of dsmt3 sample ($\kappa = 0.5$)
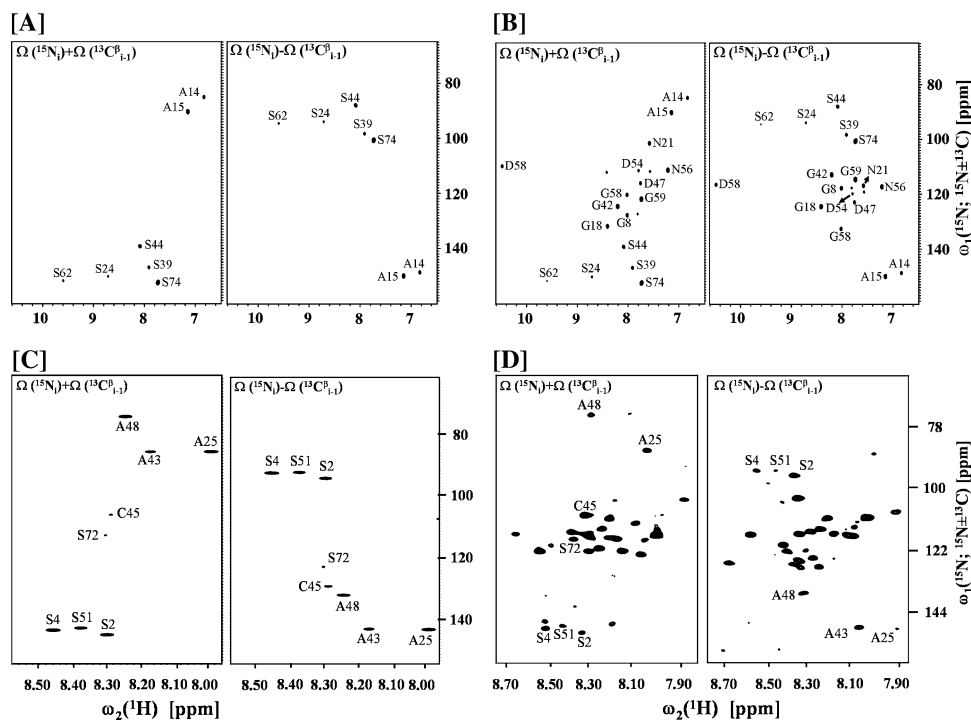
protein has 3 Ala, 5 Asn, 6 Asp, 1 Cys, 9 Gly and 4 Ser residues. As is evident from Fig. 3c, we could readily identify all the expected connectivities to the C-terminal sequential neighbours of Ala, Cys and Ser residues present in the protein. Further, based on its C^β chemical shift signature, we could comment on the redox state of the lone Cys residue (reduced) present in the protein. On the other hand, Fig. 3d shows the expected extra signatures arising from Asn, Asp and Gly residues. The peaks arising from all the ten Gly residues could be readily distinguished from that of Asn/Asp peaks as discussed earlier.

The utility of the proposed experiment is further demonstrated with [Ca^{2+}][Yb^{3+}]-calbindin. The paramagnetic lanthanide ion (Yb^{3+}) is introduced in the protein by selective displacement of the metal ion in the C-terminal site. Yb^{3+} was chosen as the paramagnetic metal ion due to its favorable PCS/line-broadening ratio compared to other Ln^{3+} (Allegrozzi et al. 2000; Atreya et al. 2003; Mustafi et al. 2004). The (3,2)D CB(CACO)NNH spectra are shown in Supplementary Figure S2, for the paramagnetic [Ca^{2+}][Yb^{3+}]Cb, wherein all expected connectivities are observed.

In conclusion, we propose a methodology that uses (3,2)D CB(CACO)NNH experiment to rapidly collect the data and readily identify peaks arising from six (Ala, Asn, Asp, Cys, Gly and Ser) residues in any given protein,

whether it is in its native, completely unfolded, diamagnetic or paramagnetic state. Further, the experiment can distinguish the redox state of Cys residues. In the case of large proteins, out-and-back versions of the experiments presented here can be used in combination with deuteration (Gardner and Kay 1998) and TROSY (Pervushin 2000) for increased resolution/sensitivity. The measuremental times for these experiments can be further reduced in combination with other fast data collection methods, such as longitudinal ^{1}H relaxation optimization (Atreya and Szyperski 2004). Taken together, the proposed experiment in its two forms will have wide range of applications in resonance assignment strategies like TATAPRO (Atreya et al. 2000) and structure determination of proteins.

References

Allegrozzi M, Bertini I, Janik MB, Lee YM, Liu G, Luchinat C (2000) Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. J Am Chem Soc 122:4154–4161

Atreya HS, Chary KV (2001) Selective 'unlabeling' of amino acids in fractionally 13C labeled proteins: an approach for stereospecific NMR assignments of CH3 groups in Val and Leu residues. J Biomol NMR 19:267–272

Atreya HS, Szyperski T (2004) G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. Proc Natl Acad Sci USA 101:9642–9647

Atreya HS, Szyperski T (2005) Rapid NMR data collection. Methods Enzymol 394:78–108

Atreya HS, Sahu SC, Chary KV, Govil G (2000) A tracked approach for automated NMR assignments in proteins (TATAPRO). J Biomol NMR 17:125–136

Atreya HS, Sahu SC, Bhattacharya A, Chary KV, Govil G (2001) NMR derived solution structure of an EF-hand calcium-binding protein from Entamoeba Histolytica. Biochemistry 40:14392–14403

Atreya HS, Mukherjee S, Chary KV, Lee YM, Luchinat C (2003) Structural basis for sequential displacement of Ca(2 +) by Yb(3+) in a protozoan EF-hand calcium binding protein. Protein Sci 12:412–425

Bartels C, Xia TH, Billeter M, Guntert P, Wüthrich K (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J Biomol NMR 6:1–10

Bax A, Grzesiek S (1993) Methodological advances in protein NMR. Acc Chem Res 26:131–138

Bertini I, Janik MB, Lee YM, Luchinat C, Rosato A (2001) Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix. J Am Chem Soc 123:4181–4188

Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ, Rance M (2006) Protein NMR spectroscopy: principles and practice, 2nd edn. Elsevier Academic Press, Boston

Chary KVR, Girjesh G (2008) NMR in biological systems: from molecules to humans. Springer, The Netherlands, pp 1–521

Clubb RT, Thanabal V, Wagner G (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C? chemical shifts in 15N–13C-labelled proteins. J Magn Reson 97:213–217

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Dotsch V, Wagner G (1996) Editing for amino-acid type in CBCACONH experiments based on the 13C beta–13C gamma coupling. J Magn Reson B 111:310–313

Dotsch V, Matsuo H, Wagner G (1996a) Amino-acid-type identification for deuterated proteins with a beta-carbon-edited HNCOCACB experiment. J Magn Reson B 112:95–100

Dotsch V, Oswald RE, Wagner G (1996b) Amino-acid-type-selective triple-resonance experiments. J Magn Reson B 110:107–111

Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. Chem Rev 104:3607–3622

Gardner KH, Kay LE (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct 27:357–406

Grzesiek S, Bax A (1992a) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. J Magn Reson 99:201–207

Grzesiek S, Bax A (1992b) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. J Am Chem Soc 114:6291–6293

Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. J Magn Reson 89:496–514

Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665

Keller RLJ (2004) The computer aided resonance assignment tutorial. Cantina Verlag, Goldau, Switzerland

Kim S, Szyperski T (2003) GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. J Am Chem Soc 125:1385–1393

Kim S, Szyperski T (2004) GFT NMR experiments for polypeptide backbone and 13Cbeta chemical shift assignment. J Biomol NMR 28:117–130

Kumar D, Kumar A, Misra JR, Chugh J, Sharma S, Hosur RV (2007) 1H, 15N, 13C resonance assignment of folded and 8 M urea-denatured state of SUMO from Drosophila melanogaster. Biomol NMR Assign

Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol 177:44–73

Mustafi SM, Mukherjee S, Chary KV, Del BC, Luchinat C (2004) Energetics and mechanism of Ca^{2+} displacement by lanthanides in a calcium binding protein. Biochemistry 43:9320–9331

Pervushin K (2000) Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. Q Rev Biophys 33:161–197

Sahu SC, Atreya HS, Chauhan S, Bhattacharya A, Chary KV, Govil G (1999) Sequence-specific 1H, 13C and 15N assignments of a calcium binding protein from Entamoeba histolytica. J Biomol NMR 14:93–94

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for structure determination of

proteins in solution employing pulsed field gradients. Prog NMR Spectroscopy 34:93–158

Schubert M (2001) Development and application of novel triple-resonance experiments for the assignment of protein NMR spectra. PhD Dissertation, pp 1–170

Shaka AJ (1985) Composite pulses for ultraa-broadband spin inversion. Chem Phys Lett 120:201–205

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

Shaka AJ, Lee CJ, Pines A (1988) Iterative schemes for bilinear operators: application to spin decoupling. J Magn Reson 77:274–293

Szyperski T, Atreya HS (2006) Principles and applications of GFT projection NMR spectroscopy. Magn Reson Chem 44 Spec No:S51–S60

Wittekind M, Mueller L (1993) 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 101:201–205