J Biomol NMR (2012) 52:115–126
DOI 10.1007/s10858-011-9598-x

ARTICLE

Efficient sequential assignments in proteins with reduced dimensionality 3D HN(CA)NH

Kousik Chandra · Garima Jaipuria · Divya Shet · Hanudatta S. Atreya

Received: 25 October 2011 / Accepted: 6 November 2011 / Published online: 7 January 2012

Abstract We present reduced dimensionality (RD) 3D HN(CA)NH for efficient sequential assignment in proteins. The experiment correlates the ^{15}N and ^{1}H chemical shift of a residue ('*i*') with those of its immediate N-terminal ($i-1$) and C-terminal ($i+1$) neighbors and provides four-dimensional chemical shift correlations rapidly with high resolution. An assignment strategy is presented which combines the correlations observed in this experiment with amino acid type information obtained from 3D CBCA (CO)NH. By classifying the 20 amino acid types into seven distinct categories based on $^{13}C^{\beta}$ chemical shifts, it is observed that a stretch of five sequentially connected residues is sufficient to map uniquely on to the polypeptide for sequence specific resonance assignments. This method is exemplified by application to three different systems: maltose binding protein (42 kDa), intrinsically disordered domain of insulin-like growth factor binding protein-2 and Ubiquitin. Fast data acquisition is demonstrated using longitudinal ^{1}H relaxation optimization. Overall, 3D HN(CA)NH is a powerful tool for high throughput resonance assignment, in particular for unfolded or intrinsically disordered polypeptides.

Keywords Sequence specific resonance assignment · Reduced dimensionality NMR · Protein structure · GFT NMR

Introduction

The process of three-dimensional (3D) structure determination of proteins by NMR begins with sequence specific resonance assignments of its backbone and side-chain ^{1}H, ^{13}C and ^{15}N nuclei. This is accomplished for folded/structured proteins with molecular mass >10 kDa using a set of 3D triple resonance spectra acquired with a doubly or triply labeled (^{13}C, ^{15}N or ^{2}H, ^{13}C, ^{15}N) protein sample (Cavanagh et al. 2007). The method essentially consists of establishing sequential connectivities between neighbouring amino acid residues along the polypeptide chain. Once a reasonably long stretch of contiguous residues are sequentially connected, the segment is mapped on to the primary sequence for assigning the resonances sequence specifically. This also forms the basis of automated assignment approaches which exploit the high sensitivity and resolution of triple resonance experiments (Baran et al. 2004). In the case of unstructured/unfolded and/or larger molecular weight proteins, sequential assignments become a challenging task. The conventional procedure, as alluded to above, has to be combined with alternative methods. For instance, it becomes necessary to employ experiments which provide a large number of different chemical shift correlations to establish sequential connectivities (Tugarinov et al. 2004). This is further supplemented in certain cases with selective spin system identification to aid resonance assignments (Ohki and Kainosho 2008).

Among the different methods proposed for sequential assignments, some are based on 3D/4D HNCANH, which

Electronic supplementary material The online version of this article (doi:10.1007/s10858-011-9598-x) contains supplementary material, which is available to authorized users.

K. Chandra · G. Jaipuria · D. Shet · H. S. Atreya (✉)
NMR Research Centre, Indian Institute of Science, Bangalore, 560012, India
e-mail: hsatreya@sif.iisc.ernet.in

G. Jaipuria
Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560012, India

correlate the chemical shifts of ^{15}N and $^{1}H^{N}$ of a residue with those of its N-terminal and/or C-terminal neighbours. These are more efficient in that they directly correlate $^{15}N/^{1}H^{N}$ shifts of a residue with those of its neighbouring residues. Different versions of this experiment have been proposed (Frueh et al. 2006, 2009; Ikegami et al. 1997; Panchal et al. 2001; Shirakawa et al. 1995; Weisemann et al. 1993). In the method proposed by Weisemann et al. the chemical shifts of ^{15}N and $^{1}H^{N}$ of a given residue i are correlated with those of $^{15}N_{i+1}/^{1}H^{N}_{i+1}$ and $^{15}N_{i-1}/^{1}H^{N}_{i-1}$ using two 3D experiments: one correlating $^{15}N_i/^{1}H^{N}_i$ with $^{15}N_{i+1}/^{15}N_{i-1}$ and another correlating $^{15}N_i/^{1}H^{N}_i$ with $^{1}H^{N}_{i+1}/^{1}H^{N}_{i-1}$. Using these two experiments in concert the sequential connectivities between neighboring residues are established. Subsequently, a single 4D HNCANH experiment providing these correlations was proposed (Ikegami et al. 1997). The HNCANH based experiments described above are complimented by different versions of 3D/4D HNCOCANH experiments which correlate $^{15}N_i/^{1}H^{N}_i$ shifts with those of $^{15}N_{i-1}/^{1}H^{N}_{i-1}$ (Bracken et al. 1997; Fiorito et al. 2006; Hiller et al. 2007; Panchal et al. 2001; Shirakawa et al. 1995). Further, a 3D HNCACB/3D CBCA (CO)NH is often used along with these experiments for amino acid type identification based on $^{13}C^{\beta}$ chemical shifts.

In the case of large and/or unstructured proteins, it is desirable to correlate $^{15}N_i/^{1}H^{N}_i$ resonances simultaneously with ^{15}N and $^{1}H^{N}$ of residues $i - 1$ and $i + 1$ in a single experiment. While the 4D experiment mentioned above (Ikegami et al. 1997) affords this correlation it needs a long *minimal* measurement time to obtain sufficient resolution. Further, a sequentially linked tri-peptide segment often cannot be mapped onto the primary sequence uniquely and longer stretches are thus required. We propose here a reduced dimensionality (RD) 3D HN(CA)NH experiment which achieves the dual objective of: (1) rapid and high resolution data acquisition and (2) the ability to sequentially connect longer segments of amino acid residues starting from a given $^{15}N_i/^{1}H^{N}_i$ (the resonances of $^{15}N_{i-2}/^{1}H^{N}_{i-2}$ and $^{15}N_{i+2}/^{1}H^{N}_{i+2}$ in addition to $^{15}N_{i+1/i-1}/^{1}H^{N}_{i+1/i-1}$) providing thus a stretch of penta-peptide which can be uniquely mapped onto the sequence. Reduced dimensionality NMR involves joint sampling of two or more chemical shifts in a single indirect dimension thereby reducing the measurement time by an order of magnitude while retaining the high dimensional shift correlations (Atreya and Szyperski 2005; Szyperski et al. 2002). In the approach presented here, a 3D HN(CA)NH (providing 4D spectral information) combined with a 3D CBCA(CO)NH experiment (for amino acid type identification) is sufficient and provides a robust and efficient approach for sequential assignments. We demonstrate the method on three different systems representing, respectively, large, unstructured and well folded proteins: per-deuterated maltose-binding protein (MBP; 42 kDa), the linker (intrinsically disordered) domain of Insulin-like growth factor binding protein-2 (L-IGFBP2; 12 kDa) and ubiquitin (8.6 kDa). In the case of ubiquitin, we show that the RD experiment can be acquired rapidly using the method of longitudinal ^{1}H relaxation optimization (L-optmization) (Atreya and Szyperski 2004; Pervushin et al. 2002) with a measurement time of 2.5 h which is useful for high throughput resonance assignments. Further, due to the high precision of shifts obtained, the spectral analysis can be easily automated. We describe below the methodology used and the results obtained on the different systems.

Methods and materials

NMR Sample preparation and data collection/analysis

Ubiquitin containing plasmid (PGLUB) was transformed into *E. coli* BL21 cells. Cells were grown at 37°C in M9 minimal medium consisting of 1 g/L of $^{15}NH_4Cl$ and 4 g/L of ^{13}C-Glucose. Protein expression was induced at midlog phase ($O.D_{600}$ ~0.6) by addition of 1.0 mM isopropyl β-D-thiogalactoside (IPTG). Cells were harvested by centrifugation and solubilized in acetate buffer (5 mM EDTA, 50 mM Na acetate, pH 5). Following sonication, the supernatant containing the protein was loaded on to a pre-equilibrated ion exchange column (SP Sepharose fast flow from GE) and the protein eluted with a salt gradient of 0–0.6 M NaCl. For NMR studies, a sample containing ~2.0 mM of protein in 50 mM Phosphate buffer (95% H_2O/5% $^{2}H_2O$; pH 6.0) was prepared. A 0.5 mM sample of per-deuterated triply labeled ($^{13}C/^{15}H/^{2}H$) MBP was purchased from PROTERA (http://www.proterasrl.com).

The gene corresponding to the central flexible (linker) domain of IGFBP-2 [residues 97–190 from the full length protein (Swain et al. 2010)] was PCR-amplified from the plasmid corresponding to the full length protein and cloned into a pET3a vector with a GST-fusion tag at the N-terminus of the protein resulting in 105 residues (~12 kDa). The constructs were transformed into BL21(DE3) cells. The cells were grown at 37°C in M9 minimal medium consisting of 1 g/L of $^{15}NH_4Cl$ and 4 g/L of ^{13}C-Glucose and induced with 0.5 mM IPTG. The protein was purified by binding the cell lysate to GST-beads followed by cleavage and separation of the GST-bound protein by HRV-3C protease using a protocol described earlier (Swain et al. 2010). The final NMR sample contained ~1.0 mM protein in 50 mM Na-phosphate buffer (pH 6.0) and 100 mM NaCl in 95% H_2O/5% $^{2}H_2O$.

All NMR experiments were performed at 25°C on Bruker 700 and 800 MHz spectrometers equipped with a

cryogenic probe (acquisition parameters and measurement times for the different experiments are provided in Table S1 of the Supporting Information). Data were processed with NMRPipe (Delaglio et al. 1995) and analyzed using XEASY (Bartels et al. 1995). Since the spectrum is acquired in a RD manner (Szyperski et al. 2002) the raw-data does not require any pre-processing and can be processed using methods used for conventional 3D FT-NMR spectra.

Results and discussion

Implementation of the experiments

Figure 1 shows schematically the magnetization transfer pathway implemented in 3D HN(CA)NH and Fig. 2 shows the radio-frequency (r.f.) pulse sequences for the non L- and L-optimized versions, respectively. Briefly, starting with the transfer of polarization from $^{1}H_i^N$ to $^{15}N_i$, the magnetization is transferred to $^{13}C_i^{\alpha}$ (using $^{1}J_{NC}^{\alpha}$) and $^{13}C_{i-1}^{\alpha}$ (using $^{2}J_{NC}^{\alpha}$), respectively. The magnetization is subsequently relayed to $^{15}N_{i+1}$ (from $^{13}C_i^{\alpha}$) and $^{15}N_{i-1}$ (from $^{13}C_{i-1}^{\alpha}$) before detecting on $^{1}H_{i+1}^N$ and $^{1}H_{i-1}^N$, respectively. As depicted in Fig. 1b, the magnetization originating on a given $^{15}N_i$ to $^{1}H_i^N$ is transferred to $^{15}N_{i+1}/^{1}H_{i+1}^N$ and $^{15}N_{i-1}/^{1}H_{i-1}^N$ and vice versa.

The chemical shift evolution periods of $^{1}H^N$ (semi-constant time) and ^{15}N (constant time) are co-incremented in the t_1 (ω_1) dimension with the ^{1}H shifts scaled by a factor 'κ' relative to ^{15}N which results in cross peaks with the following linear combinations: ω_1: $\Omega(^{15}N_{i-1}) \pm \kappa^*\Omega(^{1}H_{i-1}^N)$ and ω_1: $\Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{1}H_{i+1}^N)$ for a given ω_2: $^{15}N_i$ and ω_3: $^{1}H_i^N$. Owing to tenfold higher gyromagnetic ratio of ^{1}H relative to ^{15}N, the chemical shifts of ^{1}H are scaled 10 times over ^{15}N. Scaling can be tuned further using the factor 'κ' (Kim and Szyperski 2003; Szyperski and Atreya 2006), which allows one to increase the dispersion of peaks or to restrict the chemical shift evolution of ^{1}H to avoid loss in sensitivity due to transverse relaxation of ^{1}H. In the present study, $\kappa = 1$ was chosen.

One of the key factors which govern the intensity of the cross peaks is the total delay period used for transfer of magnetization from C^{α} to N (denoted as τ_4 in Fig. 2). Taking relaxation of C^{α} into account, the effective transfer function for the cross peak (the product operator terms at different points in the sequence have been described in detail in Weisemann et al. 1993 and Panchal et al. 2001) is given as:

$$\Gamma_{\text{cross-peaks}} = \sin(\pi *^{1} J_{C\alpha N} * \tau_4)\sin(\pi *^{2} J_{C\alpha N} * \tau_4) \cos(\pi *^{1} J_{C\alpha C\beta} * \tau_4)\exp(-\tau_4/T_2(C^{\alpha})) \quad (1)$$

A diagonal peak at ω_1: $\Omega(^{15}N_i) \pm \kappa^*\Omega(^{1}H_i^N)$ is also observed due to incomplete transfer of magnetization from $^{13}C^{\alpha}$ to ^{15}N and its intensity is governed by the following transfer function:

$$\Gamma_{\text{diagonal}} = -2 * \cos(\pi *^{1} J_{C\alpha N} * \tau_4)\cos(\pi *^{2} J_{C\alpha N} * \tau_4) \cos(\pi *^{1} J_{C\alpha C\beta} * \tau_4)\exp(-\tau_4/T_2(C^{\alpha})) \quad (2)$$

Figure 3 shows a plot of the transfer functions as a function of τ_4 and the $T_2(C^{\alpha})$. As evident from these plots, the intensity of cross peaks reaches a maximum (*in magnitude*) at two values of τ_4 (denoted as 1st and 2nd Maximum in Fig. 3a). The intensity of the cross peaks at τ_4 ~55 ms is half that value at τ_4 ~30 ms ignoring relaxation. However, this changes if losses due to relaxation are included (Fig. 3b). Tuning τ_4 to ~55 ms (2nd maximum) results in maximal cross peak and minimal diagonal peak intensity. On the other hand, a value of τ_4 ~30 ms (1st maximum) leads to 'diagonal peaks' at linear combinations: $\Omega(^{15}N_i) \pm \Omega(^{1}H_i^N)$. These peaks have a phase opposite to that of the cross peaks (Figure S3). This has two advantages: (1) it helps in identifying the diagonal peaks and (2) helps in linking the sequential residues ($i + 1$, $i - 1$) which give cross peaks at the position of diagonal peak in the RD dimension (Figure S3). However, a potential drawback is the overlap of the diagonal peak with cross peaks in case of degeneracy in ^{15}N and ^{1}H chemical shifts of $i + 1/i - 1$ with those of i. This can lead to spectral crowding and/or mutual cancellation of the peaks. Notably, the optimal value of τ_4

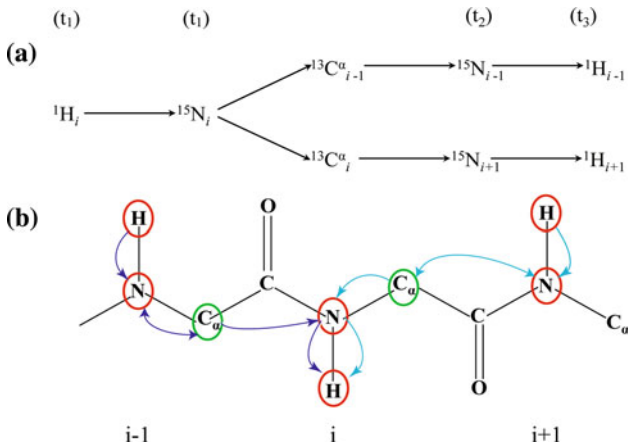

Fig. 1 **a** Schematic representation of the selected coherence transfer pathways employed in RD-3D HN(CA)NH. The magnetization flow from $^{1}H^N(i)$ is shown and the frequency labeling of the appropriate nuclei are indicated. **b** Schematic illustration of the magnetization transfer pathway implemented in 3D HN(CA)NH experiments. In **b** only pathways originating from $^{1}H(i-1)^N$ to $^{1}H(i+1)^N$ and detected on $^{1}H^N(i)$ are indicated. The *red circles* indicate the nuclei which are frequency labeled in the experiment. *Blue* and *violet arrows* indicate magnetization transfers and chemical shift evolution periods that relate residues $i + 1$ and $i - 1$, respectively

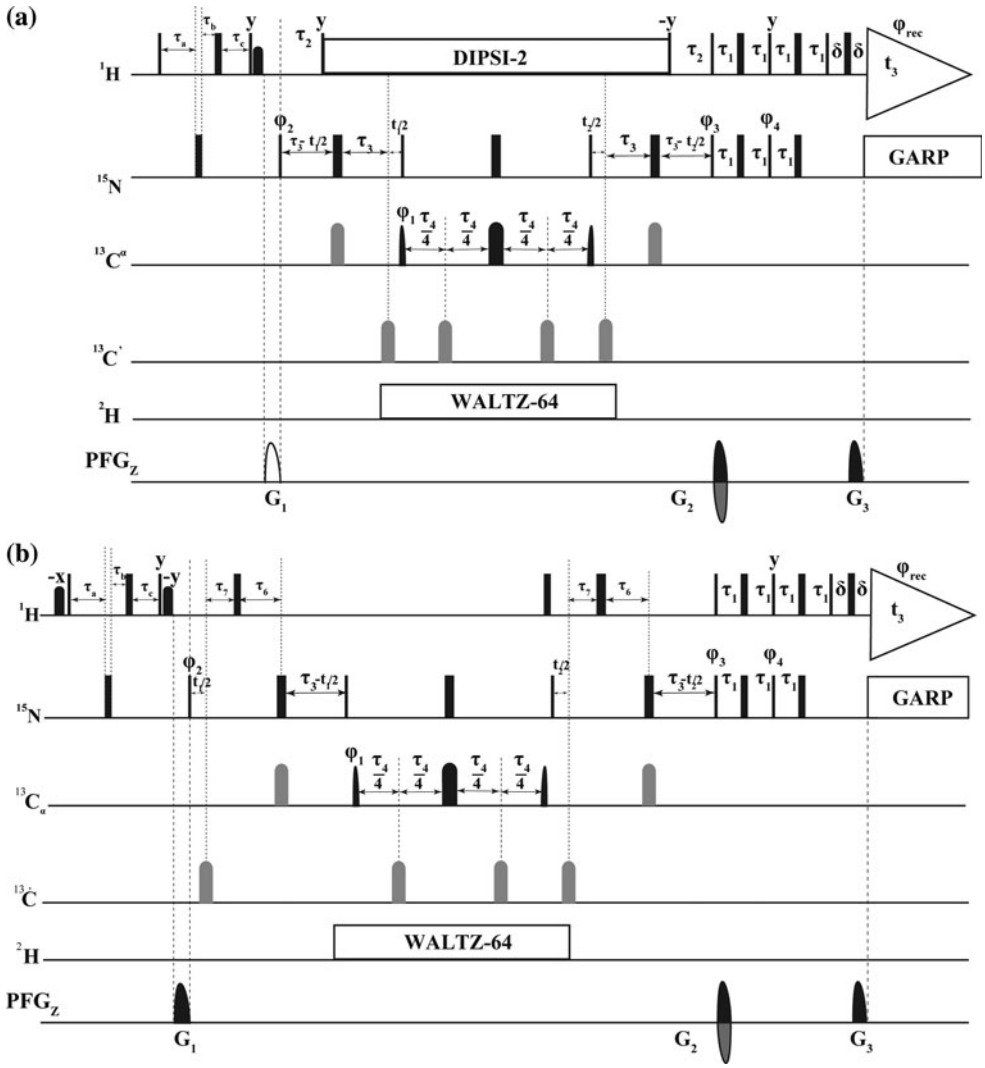

Fig. 2 R.f. pulse sequences of: **a** non-L 3D HN(CA)NH and **b** L-optimized 3D HN(CA)NH. Rectangular 90° and 180° hard pulses on ^{1}H and ^{15}N channel are indicated by *thin* and *thick vertical bars*, respectively. The representative pulse widths are 12.5, 35 and 10.5 μs for ^{1}H, ^{15}N and ^{13}C, respectively. The corresponding phases of the applied pulses are indicated *above* and in places where no r.f. phase is marked, the pulse is applied along *x*. The ^{1}H offset is placed at the position of the solvent line at 4.7 ppm. The ^{15}N carrier was adjusted according to the spectral width observed in ^{15}N dimension. The offset was set to 119, 120 and 117.5 ppm for MBP, L-IGFBP2 and Ubiquitin, respectively. The $^{13}C^{\alpha}$ r.f. carrier was placed at 54 ppm throughout the sequence and $^{13}C^{\alpha}$ carrier frequency was set at 176.0 ppm. The shaped pulse on $^{13}C^{\alpha}$ are of Gaussian cascade type (Cavanagh et al. 2007) with a pulse width of 274 and 219 μs, respectively, for 90° and 180° on resonance. The 180° off resonance pulse (Gaussian cascade) on $^{13}C^{\alpha}$ was applied for duration of 219 μs. In **a** DIPSI-2 (Cavanagh et al. 2007) is employed for decoupling ^{1}H during ^{15}N and ^{13}C shift evolution periods (r.f. strength = 6.25 kHz) and ^{2}H decoupling is achieved using WALTZ-64 (r.f. strength = 1.5 kHz) (Cavanagh et al. 2007). GARP was employed to decouple ^{15}N during acquisition (r.f. strength = 3 kHz). All pulsed z-field gradients (PFGs) are sinc shaped with gradient recovery delay of 200 μs. The duration of gradient was 1.0 ms each and the strengths of the PFGs were G1: 16 G/cm, G2: 43 G/cm, G3: 4.3 G/cm. The delays employed were: $\tau_a = 2.3$ ms, $\tau_b = 3$ μs, $\tau_c = 2.3$ ms, $\tau_1 = 2.3$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 12$ ms, $\tau_4 = 50$ ms and $\delta = 1.2$ ms. Phase cycling: $\varphi_1 = x, -x$; $\varphi_2 = 8(x), 8(-x)$; $\varphi_3 = 2(x), 2(-x)$; $\varphi_4 = 2(-y), 2(y)$ and $\varphi_{rec} = 2(x, -x, -x, x), 2(-x, x, x, -x)$. Quadrature detection in $t_2(^{15}N)$ is accomplished using the sensitivity enhanced scheme (Cavanagh et al. 2007), that is, by inverting the sign of gradient G_2 in concert with phases φ_4. Chemical shift evolution in ^{1}H channel (t_1) is achieved in a semi-constant manner by varying τ_a, τ_b and τ_c as: $\Delta\tau_a = \kappa^*\Delta t_1/2$; $\Delta\tau_b = \kappa^*\Delta t_1/2 - 2.3$ ms/(no. of complex points in t_1); $\Delta\tau_c = -2.3$ ms/(no. of complex points in t_1) where 'κ' is the scaling factor. At the same time ^{15}N chemical shift evolution period is co-incremented in a constant time manner leading to the linear combination: $\Omega(^{15}N) \pm \kappa^*\Omega(^{1}H^{N})$. The tunable scaling factor 'κ' was set to 1.0 in the present study. In **b** band-selective EBURP-2 and a time-reversed band-selective EBURP-2 pulse (Cavanagh et al. 2007), each of 1.0 ms duration centered at 2.2 ppm (covering a range of −1 to 5.5 ppm) is applied before first 90° pulse and immediately after the second 90° pulse on ^{1}H, respectively. The delays τ_6 and τ_7 were set to 9.2 and 2.6 ms, respectively, taking into account the duration of 219 μs for the 180° shape pulse on $^{13}C'$. For L-optimized sequence DIPSI-2 decoupling in ^{1}H channel was replaced by three 180° hard pulses

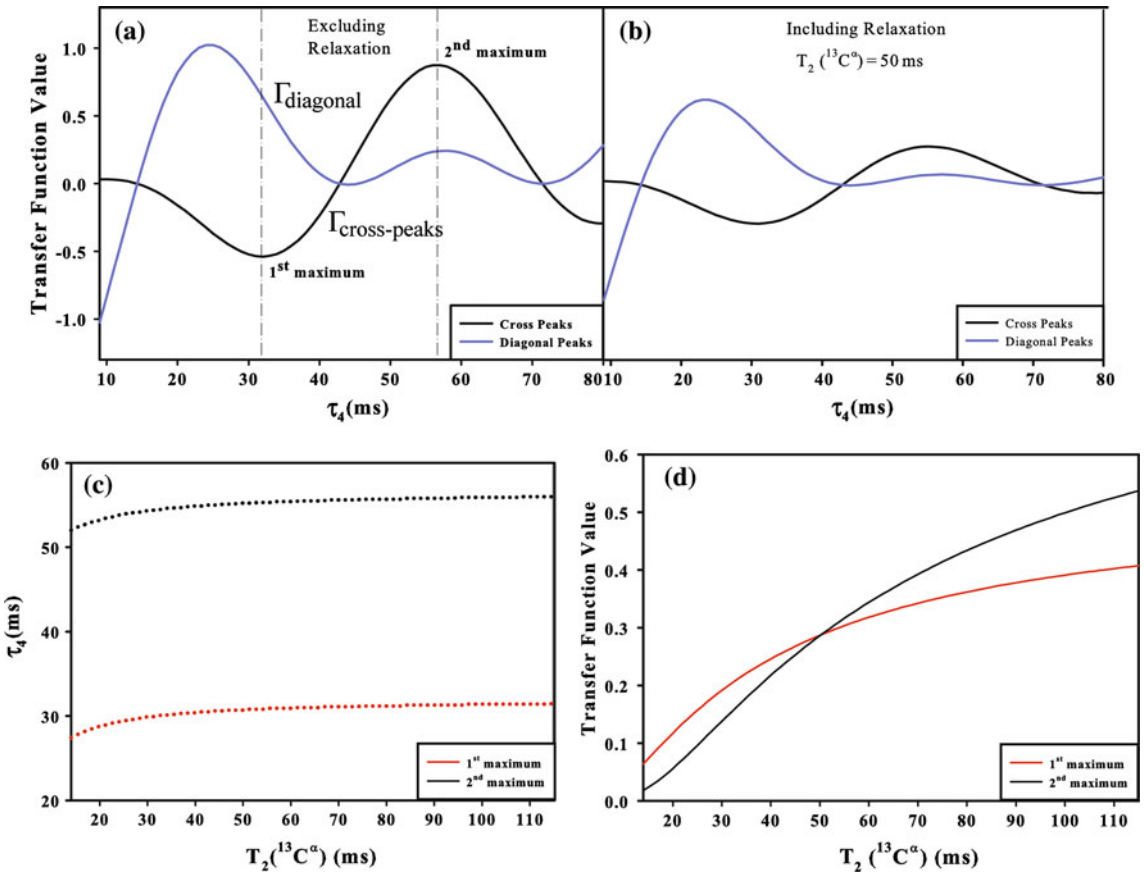

Fig. 3 Plots of the transfer functions (see Equations 1 and 2 in main text). **a** The plot of $\Gamma_{\text{cross-peaks}}$ (*black*) and $\Gamma_{\text{diagonal-peaks}}$ (*blue*) excluding the relaxation term, $\exp(-\tau_4/T_2(C^\alpha))$, as a function of the delay period τ_4 used for CA-N magnetization transfer (Fig. 2). The two maxima (*in magnitude*) of $\Gamma_{\text{cross-peaks}}$ are indicated as 1st and 2nd maximum, respectively, and shown by *dotted lines*. **b** The plot of $\Gamma_{\text{cross-peaks}}$ (black) and $\Gamma_{\text{diagonal-peaks}}$ (*blue*) including the relaxation term: $\exp(-\tau_4/T_2(C^\alpha))$ for $T_2(C^\alpha) = 50$ ms in the transfer functions. **c** Plot of τ_4 at the 1st and 2nd maximum of the transfer function shown in **a** as a function of $T_2(C^\alpha)$. **d** The value of $\Gamma_{\text{cross-peaks}}$ at τ_4 corresponding to the 1st and 2nd maximum of the transfer function shown in **a**. The values of coupling constant used are: $^1J_{C\alpha N} = 11$ Hz, $^2J_{C\alpha N} = 7$ Hz and $^1J_{C\alpha C\beta} = 35$ Hz

is nearly constant for a wide range of $T_2(C^\alpha)$ (Fig. 3c). Hence in all our data acquisition we have chosen a delay of 50 ms. For $T_2(C^\alpha)$ of <40 ms, it is advantageous to use the shorter value for the delay τ_4 where the relative sensitivity of the experiment is higher (Fig. 3d). The sensitivity of the experiment can also be increased for large systems using TROSY, which can be incorporated into the sequence similar to the way done for triple resonance experiments in general (Cavanagh et al. 2007).

A 3D (H)N(CA)NH providing ω_1: $\Omega(^{15}N_{i-1})$ and ω_1: $\Omega(^{15}N_{i+1})$ for a given ω_2: $^{15}N_i$ and ω_3: $^1H_i^N$ chemical shifts serves as a central peak spectrum (Atreya and Szyperski 2005; Kim and Szyperski 2003). Central peaks encode the chemical shift of the centre of the two linear combinations observed in the RD experiment (Szyperski et al. 1995, 1996). This serves to: (1) identify peak pairs if chemical shifts in all but the RD dimension are degenerate and (2) to recover sensitivity by symmetrization about the position of the central peaks (Szyperski et al. 1995). For the RD experiment proposed in the present study the central peak spectrum is not required as discussed below (see "Strategy for sequential assignments" below). If desired, the central peak spectrum can also be obtained for 'free' from the ^{15}N –steady state magnetization as shown previously for RD NMR experiments (Szyperski et al. 1996, 2002). Further, in the current implementation, the $^1H^N$ shift evolution during t_1 is not detected in quadrature like in the G-matrix Fourier transform (GFT) approach (Atreya and Szyperski 2005; Kim and Szyperski 2003). This is owing to the fact that the combination of shifts: $\Omega(^{15}N) \pm \Omega(^1H^N)$ in the RD dimension (ω_1) lie in distinct enhanced spectral regions as a result of chemical shifts scaling of ^1H (10 times) over ^{15}N and hence do not require editing (discussed below).

In the case of L-optimization (Fig. 2b), the r.f. pulse sequence is implemented in a way so as to preserve the magnetization of aliphatic and water ^1H spins along + z

with minimal perturbation (Atreya and Szyperski 2004; Pervushin et al. 2002). This is accomplished by: (1) removal of ^{1}H decoupling and (2) incorporating band selective ^{1}H r.f. pulses with excitation in the region of −1 to 5.5 ppm placed at the start of the pulse sequence and immediately after the second 90° pulse on ^{1}H (Fig. 2b). This is similar to that implemented by Diercks et al. (2005) where it was shown to maximally retain the magnetization of ^{1}H (aliphatic/water) along z-axis.

3D HN(CA)NH spectrum is acquired typically with a t_{max} of 10 ms resulting in line-widths of <90 Hz in the RD (ω_1) dimension. Higher resolution can be achieved by utilizing the complete constant time evolution period for ^{15}N. However, this results in longer acquisition times for $^{1}H^{N}$ which limits the sensitivity due to its T_2 relaxation. The chemical shift evolution of $^{1}H^{N}$ can be tuned using the scaling factor, κ (Fig. 2). The acquisition times in the RD dimension can thus be adjusted depending on the protein under investigation. The $^{1}H^{N}$ chemical shift is extracted from $\Omega(^{15}N) \pm \Omega(^{1}H^{N})$ by subtracting the two linear combinations followed by scaling down by a factor of 10 for $\kappa = 1$ (due to ten-fold higher gyromagnetic ratio of ^{1}H relative to ^{15}N). This scales down the error in $\Omega(^{1}H^{N})$ associated with incorrect picking of peak positions. Moreover, the precision of $\Omega(^{15}N)$ and $\Omega(^{1}H^{N})$ shifts is enhanced by the fact that chemical shifts are multiply encoded [i.e., in the two linear combinations: $\Omega(^{15}N) \pm \Omega(^{1}H^{N})$]. This corresponds to performing statistically independent multiple measurements resulting in an increase in precision by a factor $\sqrt{2}$ for ^{15}N and $^{1}H^{N}$ (Atreya and Szyperski 2005; Kim and Szyperski. 2003). Overall (^{15}N, $^{1}H^{N}$) shifts from the RD dimension are obtained with a tolerance of less than ± 0.5 ppm for ^{15}N and ± 0.05 ppm for $^{1}H^{N}$. This resolution is sufficient to map the (^{15}N, $^{1}H^{N}$) shifts obtained for residues $i - 1$ and $i + 1$ on to 2D [^{15}N, ^{1}H] HSQC and helps in ruling out incorrect pairing of $\Omega(^{15}N) \pm \Omega(^{1}H^{N})$ as discussed below (see "Strategy for sequential assignments"). Such precision of shifts also affords automated assignment of this spectrum. Notably, a 4D spectrum acquired using the conventional sampling scheme would require an order of magnitude more of minimal measurement time to achieve an equivalent resolution in the indirect ^{15}N and $^{1}H^{N}$ dimensions.

Strategy for sequential assignments

Backbone sequential walk using 3D HN(CA)NH is accomplished in concert with 3D CBCA(CO)NH (for spin system identification). This is depicted schematically in Fig. 4. We first discuss below a case where the residue, 'i': (a) does not contain a Pro residue as $i - 1$ or $i + 1$ and (b) its ($^{15}N_i$, $^{1}H_i^{N}$) shifts do not overlap with those of another residue in the 2D [^{15}N, ^{1}H] HSQC spectrum. These cases are discussed later as special cases.

1. Starting first from a given ($^{15}N_i$, $^{1}H_i^{N}$) peak in the 2D [^{15}N, ^{1}H] HSQC spectrum, four peaks corresponding to $\Omega(^{15}N_{i-1}) \pm \Omega(^{1}H_{i-1}^{N})$ and $\Omega(^{15}N_{i+1}) \pm \Omega(^{1}H_{i+1}^{N})$ are expected to be observed in the RD (ω_1) dimension (provided $i - 1$ or $i + 1$ is not a proline). If the central peak spectrum is available, the four individual chemical shifts corresponding to $^{15}N_{i-1}$, $^{15}N_{i+1}$, $^{1}H_{i-1}^{N}$ and $^{1}H_{i+1}^{N}$ can be obtained in a straightforward manner from sums and differences of the linear combinations. If the central peak spectrum is not available, all (four) possible combinations of the peaks observed in the RD dimension (ω_1) are taken and four pairs of (^{15}N, $^{1}H^{N}$) shifts representing $i - 1$ and $i + 1$ are collected.
2. Only two of the four pairs of (^{15}N, $^{1}H^{N}$) shifts are correct combinations. In most cases only the two correct combinations will correspond to a peak in the 2D HSQC spectrum. The correct pairs are verified further by referring to the respective 2D (ω_1, ω_3) strips of each of the pair of (^{15}N, $^{1}H^{N}$) shifts calculated. In the 2D strips corresponding to the correct combination of ($^{15}N_{i+1}$, $^{1}H_{i+1}^{N}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}^{N}$), a correlation to ($^{15}N_i$, $^{1}H_i^{N}$) (indicated as red in Fig. 4a) will be observed. Only the correct combination of ($^{15}N_{i+1}$, $^{1}H_{i+1}^{N}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}^{N}$) will yield such correlations.
3. Once the ($^{15}N_i$, $^{1}H_i^{N}$) correlation is identified in the 2D (ω_1, ω_3) strips of ($^{15}N_{i+1}$, $^{1}H_{i+1}^{N}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}^{N}$), the other correlation observed in these strips will automatically correspond to ($^{15}N_{i+2}$, $^{1}H_{i+2}^{N}$) and ($^{15}N_{i-2}$, $^{1}H_{i-2}^{N}$), respectively. Thus, a penta-peptide stretch comprising ($^{15}N_k$, $^{1}H_k^{N}$) ($k = i - 2, \ldots, i + 2$) is identified. The procedure can be repeated for longer stretches of assignments. However, as discussed below based on statistical analysis, a penta-peptide can be uniquely mapped onto the primary sequence. Note that the central peak information is only required for the starting ($^{15}N_i$, $^{1}H_i^{N}$) strip and not required thereafter.
4. The next step is to map the penta-peptide stretch on to the primary sequence for sequence specific resonance assignments. This requires the information of the amino acid type corresponding to each residue in the stretch. A 3D CBCA(CO)NH spectrum is used for this purpose, which helps in identifying the amino acid type for residue $k - 1$ ($k = i - 2, \ldots, i + 2$) corresponding to each ($^{15}N_k$, $^{1}H_k^{N}$) of the penta-peptide. In our approach, based on the $^{13}C^{\beta}$ chemical shifts, the 20 amino acids are divided into 7 different categories (Atreya et al. 2000; Atreya and Chary. 2002). Amino acids in each category are assigned a single digit code (0–6): 0-Proline; 1-Alanine; 2-Glycine; 3-Serine; 4-Threonine;

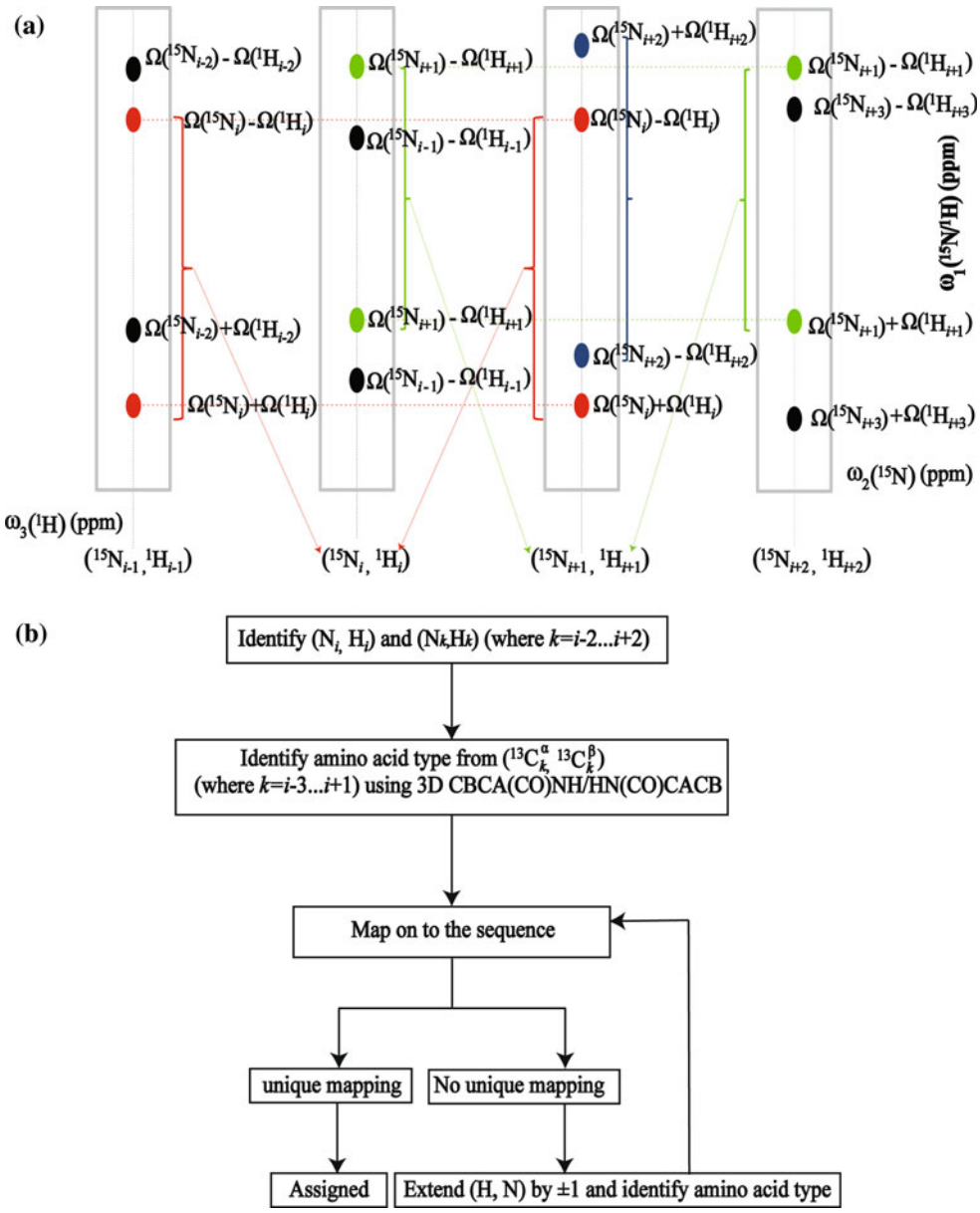

Fig. 4 **a** A schematic representation of the peak pattern in 3D HN(CA)NH. Each 2D (ω_1, ω_3) strip shown corresponds to one residue (say 'i') taken at the respective $\omega_2(^{15}N_i)$ and $\omega_3(^{1}H_i^N)$ chemical shifts and consists of four cross-peaks along the RD (ω_1-) dimension at $\Omega(^{15}N_{i+1}) \pm \kappa^*\Omega(^{1}H_{i+1}^N)$ and $\Omega(^{15}N_{i-1}) \pm \kappa^*\Omega(^{1}H_{i-1}^N)$. From the different linear combinations, chemical shifts of ($^{15}N_{i+1}$, $^{1}H_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}$) are calculated. Verification of ($^{15}N_{i+1}$, $^{1}H_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}$) is carried out by referring to 2D (ω_1, ω_3) strips corresponding to ($^{15}N_{i+1}$, $^{1}H_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}$) both of which should contain a cross-peak at $\Omega(^{15}N_i) \pm \kappa^*\Omega(^{1}H_i^N)$. The remaining cross peaks in 2D (ω_1, ω_3) strips of ($^{15}N_{i+1}$,$^{1}H_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}$) then correspond to $\Omega(^{15}N_{i+2}) \pm \kappa^*\Omega(^{1}H_{i+2}^N)$ and $\Omega(^{15}N_{i-2}) \pm \kappa^*\Omega(^{1}H_{i-2}^N)$, respectively. Thus a sequentially connected segment of five residues: ($^{15}N_k$,$^{1}H_k$) ($k = i - 2, \ldots, i + 2$) is obtained, which is assigned using the strategy depicted in **b**. Note that the verification of ($^{15}N_{i+1}$, $^{1}H_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H_{i-1}$) chemical shifts is not necessary if a central peak spectrum is acquired (see text). **b** Illustrates the strategy for sequential assignment using a combination of 3D HN(CA)NH and 3D CBCA(CO)NH/HN(CO)CACB

5-Arginine, Cysteine, Glutamine, Glutamic acid, Histidine, Lysine, Methonine, Tryptophan and Valine; 6-Aspartic Acid, Asparagine, Isoleucine, Leucine, Phenylalanine and Tyrosine. These classifications have been described in detail earlier (Atreya et al. 2000; Atreya and Chary 2002). Each residue in the penta-peptide is assigned the above codes based on the observed $^{13}C^{\beta}$ chemical shifts. Simultaneously, the primary sequence is converted to a sequence of one-digit codes. The penta-peptide is then matched on to the primary sequence. A statistical analysis (discussed below) reveals that such penta-peptide segments almost always match uniquely on the sequence despite the fact that the 20 amino acids are reduced to seven categories.

Note that the procedure of identifying ($^{15}N_k$, $^{1}H_k^N$) ($k = i - 2, \ldots, i + 2$) (step 3) is non-directional. That is, the correlations assigned to ($^{15}N_{i+1}$, $^{1}H_{i+1}^N$), ($^{15}N_{i+2}$, $^{1}H_{i+2}^N$) can be assigned to ($^{15}N_{i-1}$, $^{1}H_{i-1}^N$), ($^{15}N_{i-2}$, $^{1}H_{i-2}^N$) and vice versa. Thus, the stretch of five sequentially connected residues can be mapped on the primary sequence in the *reverse* direction. The chances of the penta-peptide mapping successfully on the sequence in both the forward and reverse direction are rare as revealed by the statistical analysis. The overall strategy is summarized in Fig. 4b.

In the case of the residue $i - 1$ or $i + 1$ being a proline, the corresponding (^{15}N, $^{1}H^N$) correlations are not observed on ($^{15}N_i$, $^{1}H_i^N$). The sequential walk is then carried out in one direction and a tetra-peptide stretch of residues is

assigned and mapped on the sequence. In the event of overlap of two (^{15}N, ^{1}H^{N}) peaks (residues *i* and *j*) in the 2D [^{15}N, ^{1}H] HSQC spectrum, a maximum of 16 possible pairs of (^{15}N, ^{1}H^{N}) can be formed from the sums and differences of the different linear combination. The four correct combinations corresponding to (^{15}N$_{i-1}$, ^{1}H$^{N}_{i-1}$), (^{15}N$_{i+1}$, ^{1}H$^{N}_{i+1}$) and (^{15}N$_{j-1}$, ^{1}H$^{N}_{j-1}$), (^{15}N$_{j+1}$, ^{1}H$^{N}_{j+1}$) are identified by referring to the respective 2D (ω_1, ω_3) strips of each of the pair of (^{15}N, ^{1}H^{N}) shifts calculated. In the 2D strips corresponding to the correct combination of (^{15}N$_{i+1/j+1}$, ^{1}H$^{N}_{i+1/j+1}$) and (^{15}N$_{i-1/j-1}$, ^{1}H$^{N}_{i-1/j-1}$), a correlation to (^{15}N$_{i/j}$, ^{1}H$^{N}_{i/j}$) should be observed. Only the correct combination of (^{15}N, ^{1}H^{N}) will yield such correlations. Note that these ambiguities are resolved in a straightforward manner if the central peak spectrum (providing ^{15}N chemical shifts only in the ω_1 dimension) is available.

Another possible scenario is that the chemical shifts of (^{15}N$_{i+1}$, ^{1}H$^{N}_{i+1}$) and (^{15}N$_{i-1}$, ^{1}H$^{N}_{i-1}$) are similar such that their respective linear combinations of $\Omega(^{15}\mathrm{N}) \pm \Omega(^{1}\mathrm{H}^{N})$ overlap. This yields a peak pattern similar to that observed when the residue $i-1$ or $i+1$ is a proline. However, a statistical analysis of resonance assignments of ~200 proteins from the BMRB data base reveals that this happens rarely (<0.1%) in any given protein.

Application of 3D HN(CA)NH

Figures 5 and Figure S1 (Supporting Information) show representative 2D (ω_1, ω_3) strip plots of spectra acquired for MBP and L-IGFBP2 in ~50 and ~15 h, respectively (Table S1). A yield of 90 and 100% were observed, respectively, for the two proteins. The distribution of

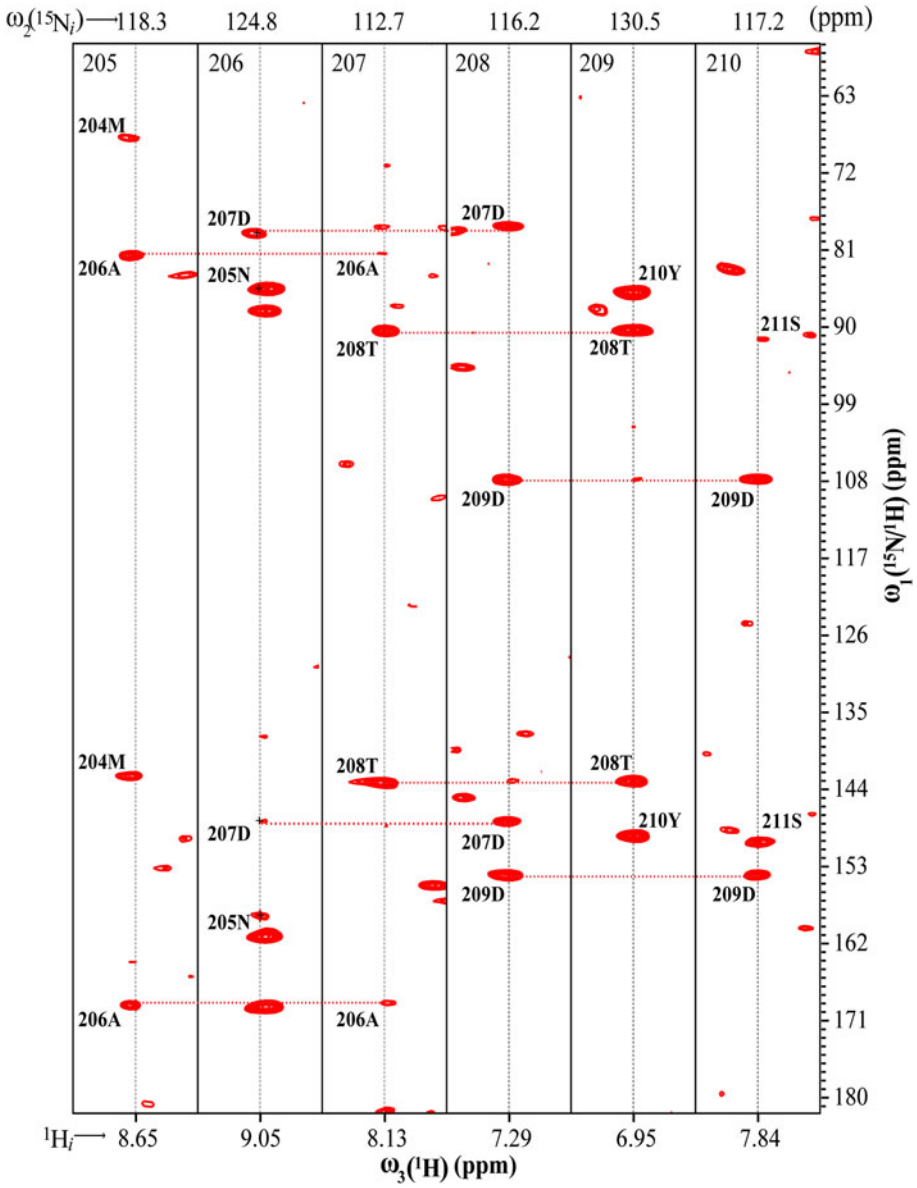

Fig. 5 2D (ω_1, ω_3) strip plots of 3D HN(CA)NH spectra acquired for MBP showing the residue stretch N205-Y210. The corresponding ^{1}H$_i$ and ^{15}N$_i$ chemical shifts for each strip are indicated below and above of the corresponding strip, respectively. In the 2D strips corresponding to (^{15}N$_{i+1}$, ^{1}H$^{N}_{i+1}$) and (^{15}N$_{i-1}$, ^{1}H$^{N}_{i-1}$), correlations to (^{15}N$_i$, ^{1}H$^{N}_i$) are indicated in *red dotted lines*

signal-to-noise (S/N) for the different cross peaks in spectra acquired for the three proteins are shown in Figure S2 of Supporting Information. The experiment is particularly useful for assignment of unfolded or intrinsically unstructured polypeptides which lack good dispersion in $^{1}H^{N}$, but retain good resolution in ^{15}N shifts (Dyson and Wright 2005). This experiment was used for the assignment of intrinsically disordered L-IGFBP2 (105 residues; 12 kDa) which contains stretches of sequences with low complexity. A noteworthy example is that of a stretch of four consecutive Gly: G50-G53 with overlapping (^{15}N, $^{1}H^{N}$) chemical shifts. In the 3D HNCACB/CBCA(CO)NH spectra, which yields di-peptide segments the pairs: G50-G51, G51-G52 and G52-G53 could not be resolved. Using 3D HN(CA)NH and 3D CB(CA)CONH the tri-peptide stretch L49-G50-G51 and G52-G53-S54 could be assigned unambiguously from the correlations to M-48 and S-54 observed on L49 and G53, respectively. This also gave the (^{15}N, $^{1}H^{N}$) assignments of G51 and G52 which were otherwise difficult to obtain. This is illustrated in Fig. 6. The chemical shift assignments of L-IGFBP2 were obtained using these set of experiments.

Figure 7 shows representative 2D (ω_1, ω_3) strip plots of spectra of L-optimized 3D HN(CA)NH acquired for

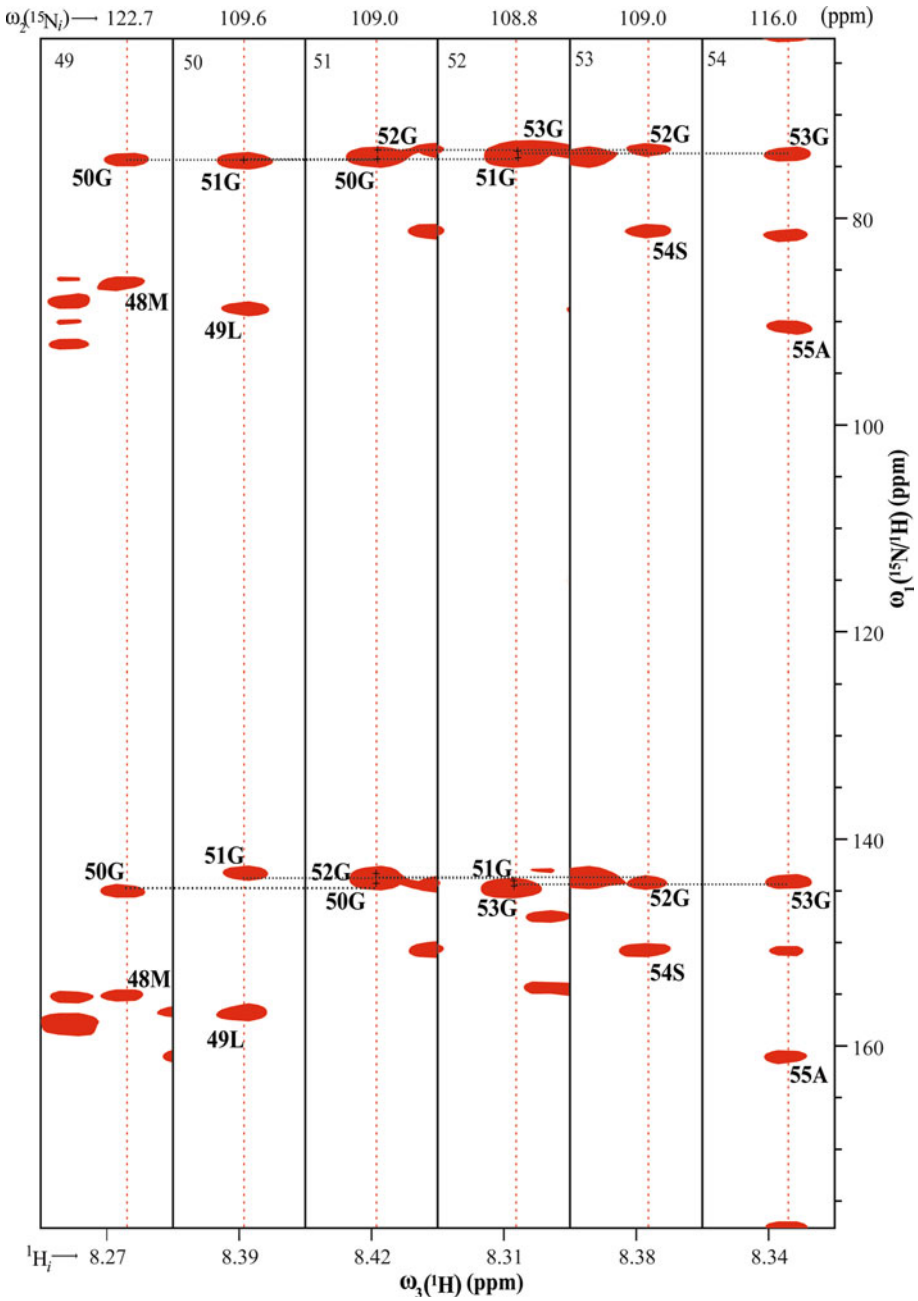

Fig. 6 2D (ω_1, ω_3) strip plots of 3D HN(CA)NH spectra acquired for L-IGFBP-2 illustrating the use of 3D HN(CA)NH in resolving ambiguities in a stretch of residues involving amino acid repeats. L-IGFBP2 contains a segment ^{49}L-G-G-G-G-S^{54}. They are unambiguously assigned using 3D HN(CA)NH which shows correlations to both ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$). The $^{1}H_i$ and $^{15}N_i$ chemical shifts for each strip are indicated below and above of the corresponding strip, respectively. In the 2D strips of ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$), correlations to ($^{15}N_i$, $^{1}H^{N}_{i}$) are indicated in *black dotted lines*

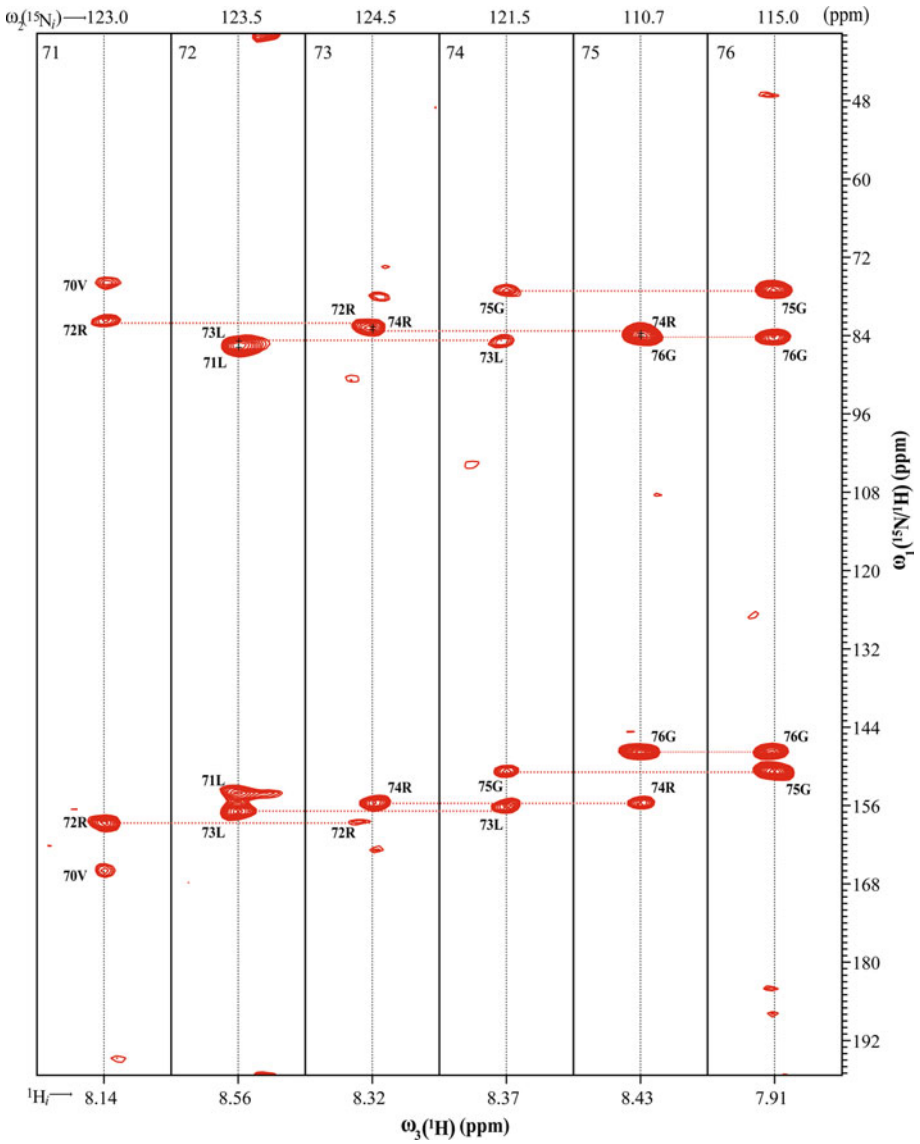

Fig. 7 2D (ω_1, ω_3) strip plots of 3D HN(CA)NH spectra acquired for Ubiquitin using L-optimzation (Fig. 2b) showing the residue stretch R72-G76. The corresponding $^{1}H_i$ and $^{15}N_i$ chemical shifts for each strip are indicated below and above of the corresponding strip, respectively. In the 2D strips of ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$), correlations to ($^{15}N_i$, $^{1}H^{N}_i$) are indicated in red dotted lines. Note that in the 2D (ω_1, ω_3) strip of G76, which is the final (C-terminal) residue of the protein, linear combinations corresponding to ($^{15}N_i$, $^{1}H^{N}_i$) are also observed. This happens because the residue at the C-terminus end of a protein does not contain $^{15}N_{i+1}$ and the magnetization is retained on $^{13}C_i$ (see Fig. 1)

ubiquitin in ~2.5 h (Table S1) (with a relaxation delay of 200 ms between scans). Such rapid data acquisition is useful for high-throughput resonance assignments and structure determination (Liu et al. 2005). Note that measurement time can be reduced further by reducing the spectral widths in the RD dimension (ω_1) so that the complete set of peaks comprising one of the linear combinations $\Omega(^{15}N_{i+1}) \pm \Omega(^{1}H^{N}_{i+1})$ get aliased. However, in such cases the two linear combinations have to be edited into different spectra using the GFT mode of data acquisition (Kim and Szyperski. 2003). This entails phase sensitive sampling of ^{1}H chemical shifts resulting in slight increase in measurement time due to an additional 2-step phase cycle required for quadrature detection of ^{1}H. Note that in the 2D (ω_1, ω_3) strip of G76, which is the final (C-terminal) residue of the protein, linear combinations corresponding to ($^{15}N_i$, $^{1}H^{N}_i$) are also observed. This happens owing to the fact that the residue at the C-terminus end of the protein does not contain $^{15}N_{i+1}$. Hence, the magnetization from $^{13}C^{\alpha}_i$ is retained on $^{15}N_i$ (Fig. 1). This results in the observed ($^{15}N_i$, $^{1}H^{N}_i$) correlations.

We performed a statistical analysis to estimate the extent to which penta-peptide segments identified from 3D HN(CA)NH for a given residue ($^{15}N_i$, $^{1}H^{N}_i$) match uniquely onto the primary sequence. A penta-peptide is chosen owing to the fact that verification of the correctness (in absence of a central peak spectrum) of ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$) shifts calculated from the linear combinations observed for a given ($^{15}N_i$, $^{1}H^{N}_i$) requires a referral to the 2D (ω_1, ω_3) strips at the corresponding ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$) shifts. This immediately yields the correlations: ($^{15}N_{i+2}$, $^{1}H^{N}_{i+2}$) and ($^{15}N_{i-2}$, $^{1}H^{N}_{i-2}$) in the respective strips. If a penta-peptide assignment is not feasible (arising from breaks due to proline/missing peaks

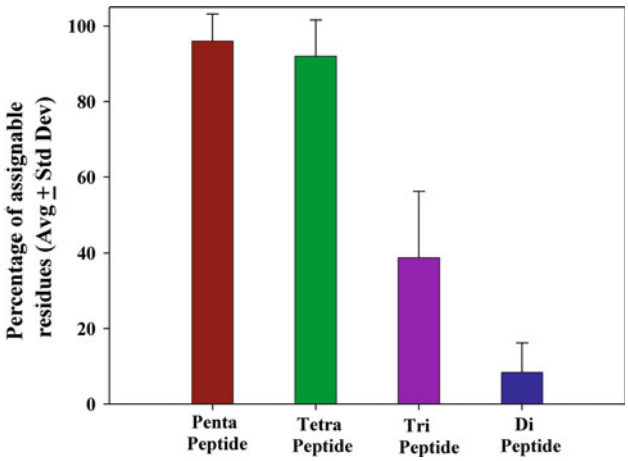

Fig. 8 Statistical analysis of protein sequences indicating the uniqueness of penta-peptide segments in proteins compared to smaller stretches. For the statistical analysis, ~200,000 non-redundant sequences from the RefSeq data base (http://www.ncbi.nlm.nih.gov/RefSeq/) were taken varying from 50 to 500 amino acids in length. Each polypeptide was converted to a sequence of one-digit codes based on the 7 amino acid categories described in text. Four different calculations involving mapping of: (1) penta-peptide, (2) tetra-peptide, (3) tri-peptide and (4) di-peptide segments on the primary sequence were carried out as described in the text

etc.), shorter stretches of residues can be mapped onto the sequence. For the statistical analysis, ~200,000 non-redundant sequences from the RefSeq data base (http://www.ncbi.nlm.nih.gov/RefSeq/) were taken varying from 50 to 500 amino acids in length. Each polypeptide was converted to a sequence of one-digit codes based on the seven amino acid categories described above (Atreya et al. 2000; Atreya and Chary. 2002). Four different calculations involving mapping of: (1) penta-peptide, (2) tetra-peptide, (3) tri-peptide and (4) di-peptide segments on the primary sequence were carried out as follows. Using a moving window of peptide of length X (where $X = 1.5$), all unique X-residue segments in the primary sequence were first obtained. Following this, subsequent rounds of search for uniqueness involving shorter peptides of length X–K ($2 \leq K < X$) was carried out for residues which were not found to be unique in the previous round. As noted above, 3D HN(CA)NH does not contain directional information. That is, the correlations assigned to ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$), ($^{15}N_{i+2}$, $^{1}H^{N}_{i+2}$) can be assigned to ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$), ($^{15}N_{i-2}$, $^{1}H^{N}_{i-2}$) and vice versa. Thus, the stretch of X sequentially connected residues can be mapped on the primary sequence in the *reverse* direction. This was taken into account in the above calculation for estimating the uniqueness. Figure 8 depicts the results of the analysis described above. In the case of di-peptide segments, only a handful of residues can be assigned uniquely. The uniqueness in mapping the residues is increased upon increase in the peptide length. If penta- and tetra-peptide stretches are taken into account, almost the complete polypeptide chain can be assigned unambiguously by categorizing amino acids into the 7-types based on their $^{13}C^{\beta}$ chemical shifts.

The experiment 3D HN(CA)NH is thus unique in terms of: (1) the correlations obtained with high resolution and (2) the straightforward manner of analysis. Alternative experiments proposed in recent years, which have similar magnetization transfer pathway include the 3D TROSY-hNcaNH (Frueh et al. 2006) and TROSY-hNCAnH (Frueh et al. 2009). The 3D HNN experiment (Panchal et al. 2001) provides similar correlations as in 3D HN(CA)NH. However, the analysis in 3D HNN is limited by the absence of direct ($^{15}N_{i+1}$, $^{1}H^{N}_{i+1}$) and ($^{15}N_{i-1}$, $^{1}H^{N}_{i-1}$) correlations which is overcome in 3D HN(CA)NH. A new method proposed recently by Takeuchi et al. (Takeuchi et al. 2011) uses a set of CACA-TOCSY based experiments which provide similar correlation as in 3D HN(CA)NH. However, these experiments require a sample with alternate ^{13}C–^{12}C labeling together with perdeuteration and limited in sensitivity due to long CA–CA isotropic mixing time.

Conclusions

In summary, we have developed an experiment for efficient and rapid sequential assignment in proteins. The experiment, 3D HN(CA)NH provides for a given residue 'i', both ^{15}N and $^{1}H^{N}$ chemical shifts of residues $i - 1$ and $i + 1$. The triplet of sequentially connected residues can be extended to longer stretches, which can then be mapped uniquely onto the primary sequence based on their amino acid types identified from $^{13}C^{\beta}$ chemical shifts. The experiment provides ^{15}N and $^{1}H^{N}$ shifts with high precision, which makes it easy to automate the assignment procedure. The efficacy of the experiment can be increased by combining this approach with other methods such as amino acid selective labeling or unlabeling (Krishnarjuna et al. 2011; Ohki and Kainosho 2008). Specific residues identified selectively in this manner can serve as starting points for the assignments. Rapid collection of the data is facilitated by L-optimization, which can be accelerated further using non-linear sampling schemes (Atreya and Szyperski 2005). Taken together, the experiment will provide an important tool for structure determination of proteins by NMR.

Acknowledgments The facilities provided by NMR Research Centre at IISc and National facility for high-field NMR at Tata Institute of Fundamental Research (TIFR; Mumbai) supported by Department of Science and Technology (DST), India is gratefully acknowledged. HSA acknowledges support from DST-SERC and Department of Biotechnology (DBT) research awards. GJ acknowledges fellowship from Council of Scientific and Industrial Research

(CSIR), India. We thank Dr. John Cort, Pacific Northwest National Laboratory, for providing the Ubiquitin plasmid and Prof. R. V. Hosur, TIFR for providing the MBP sample.

References

Atreya HS, Chary KVR (2002) Automated NMR assignments of proteins for high throughput structure determination: TATAPRO II. Curr Sci 83:1372–1376

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

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

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

Baran MC, Huang YJ, Moseley HNB, Montelione GT (2004) Automated analysis of protein NMR assignments and structures. Chem Rev 104:3541–3555

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

Bracken C, Palmer AG, Cavanagh J (1997) (H)N(COCA)NH and HN(COCA)NH experiments for ^{1}H–^{15}N backbone assignments in C-13/N-15-labeled proteins. J Biomol NMR 9:94–100

Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton NJ (2007) Protein NMR spectroscopy. Academic Press, San Diego

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

Diercks T, Daniels M, Kaptein R (2005) Extended flip-back schemes for sensitivity enhancement in multidimensional HSQC-type out-and-back experiments. J Biomol NMR 33:243–259

Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6:197–208

Fiorito F, Hiller S, Wider G, Wüthrich K (2006) Automated resonance assignment of proteins: 6D APSY-NMR. J Biomol NMR 35:27–37

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

Frueh DP, Arthanari H, Koglin A, Walsh CT, Wagner G (2009) A double TROSY hNCAnH experiment for efficient assignment of large and challenging proteins. J Am Chem Soc 131:12880–12881

Hiller S, Wasmer C, Wider G, Wüthrich K (2007) Sequence-specific resonance assignment of soluble nonglobular proteins by 7D APSY-NMR spectroscopy. J Am Chem Soc 129:10823–10828

Ikegami T, Sato S, Walchli M, Kyogoku Y, Shirakawa M (1997) An efficient HN(CA)NH pulse scheme for triple-resonance 4D correlation of sequential amide protons and nitrogens-15 in deuterated proteins. J Magn Reson 124:214–217

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

Krishnarjuna B, Jaipuria G, Thakur A, D'Silva P, Atreya HS (2011) Amino acid selective unlabeling for sequence specific resonance assignments in proteins. J Biomol NMR 49:39–51

Liu GH, Shen Y, Atreya HS, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TA, Arrowsmith CH, Montelione GT, Szyperski T (2005) NMR data collection and analysis protocol for high-throughput protein structure determination. Proc Natl Acad Sci 102:10487–10492

Ohki SY, Kainosho M (2008) Stable isotope labeling methods for protein NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 53:208–226

Panchal SC, Bhavesh NS, Hosur RV (2001) Improved 3D triple resonance experiments, HNN and HN(C)N, for H–N and N-15 sequential correlations in (C-13, N-15) labeled proteins: Application to unfolded proteins. J Biomol NMR 20:135–147

Pervushin K, Vogeli B, Eletsky A (2002) Longitudinal ^{1}H relaxation optimization in TROSY NMR spectroscopy. J Am Chem Soc 124:12898–12902

Shirakawa M, Walchli M, Shimizu M, Kyogoku Y (1995) The use of heteronuclear cross-polarisation for backbone assignment of H-2-labeled, N-15-labeled and C-13-labeled proteins—a pulse scheme for triple-resonance 4D correlation of sequential amide protons and N-15. J Biomol NMR 5:323–326

Swain M, Slomiany MG, Rosenzweig SA, Atreya HS (2010) High-yield bacterial expression and structural characterization of recombinant human insulin-like growth factor binding protein-2. Arch Biochem Biophys 501:195–200

Szyperski T, Atreya HS (2006) Principles and applications of GFT projection NMR spectroscopy. Magn Reson Chem 44:S51–S60

Szyperski T, Braun D, Fernandez C, Bartels C, Wüthrich K (1995) A novel reduced-dimensionality triple resonance experiment for efficient polypeptide backbone assignment, 3D COHNNCA. J Magn Reson B108:197–203

Szyperski T, Braun D, Banecki B, Wüthrich K (1996) Useful information from axial peak magnetization in projected NMR experiments. J Am Chem Soc 118:8146–8147

Szyperski T, Yeh DC, Sukumaran DK, Moseley HNB, Montelione GT (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. Proc Natl Acad Sci 99:8009–8014

Takeuchi K, Gal M, Takahashi H, Shimada I, Wagner G (2011) 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. J Biomol NMR 49:17–26

Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. Ann Rev Biochem 73:107–146

Weisemann R, Rüterjans H, Bermel W (1993) 3D triple-resonance NMR techniques for the sequential assignment of NH and N-15 resonances in N-15-labeled and C-13-labeled proteins. J Biomol NMR 3:113–120