

Generation of serine/threonine check points in HN(C)N spectra

DINESH KUMAR[#], JEETENDER CHUGH[#] and RAMAKRISHNA V HOSUR*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400 005
e-mail: hosur@tifr.res.in

MS received 22 January 2009; revised 26 March 2009

Abstract. We describe here a simple modification of the HN(C)N experiment for the generation of serine/threonine check points in the three-dimensional experiment. The various 'triplet of residue' specific peak patterns in the spectra are documented for ease of analysis and sequential backbone resonance assignment. The performance of this experiment, referred to as HN(C)N-ST, is demonstrated using two proteins, one properly folded and the other completely denatured. It is noteworthy that, even in the denatured protein, where spectral dispersions are rather poor, about 90% of the sequential connectivities through the chain could be established from this single experiment. This would have great implications for structural genomics efforts.

Keywords. NMR; ¹⁵N and ¹³C labelled protein; resonance assignment; band-selective pulse; check points; HN(C)N; HN(C)N-ST.

1. Introduction

New developments in multi-dimensional NMR experiments have made a significant contribution to achieve the structural genomics/proteomics aims (i.e. to generate accurate three-dimensional models for all folded, globular proteins and domains in the protein universe to understand the relationship between protein sequence, structure and function).^{1–3} For this purpose, rapid and unambiguous resonance assignment has gained utmost importance. While the automated analysis of different spectra offers one approach, the design of experimental techniques to facilitate assignments and that too with as few experiments as possible is another direction of development. In this context, the HNN experiment and its complimentary experiment, HN(C)N,^{4–7} have proved extremely valuable for a variety of protein systems.^{8–17} The recent developments of HNN tuning to generate alanine⁶ and serine/threonine specific peak patterns⁷ have enhanced the speed of assignment quite substantially. These developments involved a simple modification to the pulse sequence.

Continuing such efforts for rapid resonance assignments, we have implemented here the tuning ideas in the HN(C)N experiment to generate additional serine/threonine specific peak patterns in the spectra. This is referred to as HNCN-ST and extends the HN(C)N-A experiment previously described.⁶ This implementation becomes extremely valuable, when the polypeptide chain has far removed glycines and alanines along the sequence, which could pose some difficulties in the HN(C)N alone based application. We have experimentally tested this on a properly folded protein (M-crystallin),¹⁸ and a completely denatured protein (9.7 M urea-denatured GTPase effector domain of Dynamin).¹⁹ We have shown that even in the latter case where the spectral dispersions are very poor, as can be expected for denatured proteins, 90% of the sequential connectivities could be established from this HN(C)N-ST experiment alone.

2. Materials and methods

2.1 Protein expression and purification

The ¹⁵N-¹³C labelled GED protein was expressed and purified as described elsewhere.¹⁹ ¹⁵N-¹³C labelled apo M-crystallin¹⁸ was obtained as a kind gift from Ravi P Barnwal.

*For correspondence

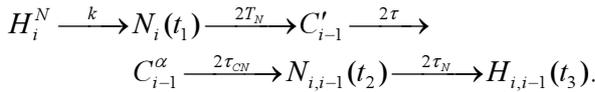
[#]Both the authors have contributed equally

bon channel was placed at 64 ppm (i.e. selectively inverting from 49 ppm to 79 ppm range of $C\alpha/C\beta$ resonances).

3. Results and discussion

3.1 Modification to the HN(C)N pulse sequence

The HN(C)N experiment (figure 1) employs the following magnetization transfer:²¹



The delays $2T_N$, $2\tau_C$, $2\tau_{CN}$ and $2\tau_N$ are the time periods during which magnetization transfer take place; $2\tau_{CN}$ should be chosen between 22 and 30 ms to get best transfers and obtain the characteristic peak patterns in the spectra. The choice would be a reasonable compromise between transfer efficiency and transverse relaxation losses. The pulse sequence in figure 1 differs slightly from the original sequence²¹ in that it uses WATERGATE²² sequence before detection for better water suppression and as a consequence the ‘sensitivity enhancement’ part has been dropped; frequency selection along the two indirect dimensions was done by States-TPPI method.²³ Further, the water flip back pulses have been used to provide sensitivity gain for labile NH groups avoiding saturation of the water magnetization.²⁴

The modification to the pulse sequence that we are proposing here, for the specific purpose described above, is indicated by the encircled pulse in figure 1. This 180° refocusing pulse during τ_{CN} constant time evolution (figure 1) is made band selective so that C^α resonances of all except glycines and C^β resonances of only serines and threonines are refocused as shown earlier.⁷ This means only serines and threonines go through all the evolution and transfer steps as in the original pulse sequence. The manifestation of this in the peak patterns for the various triplet sequences is described in the following sections.

3.2 Theory

In the HN(C)N experiments, in a chain of three residues, $i-1$ to i to $i+1$, the intensities of the diagonal (I_i^d) and the cross (I_{i-1}^c) peaks in the (F_2 , F_3) planes of the HN(C)N spectrum are given by:²⁵

$$I_i^d = A_1 A_2 A_4 A_6 K_1 K_2 \quad (1)$$

$$I_{i-1}^c = -A_1 A_3 A_5 A_6 K_1 K_3, \quad (2)$$

where I_i^d and I_{i-1}^c are the intensities of the diagonal and cross peaks, respectively,

$$A_1 = \sin^2(2\pi l_{i-1} \tau_C) \sin(2\pi m_i T_N) \quad (3)$$

$$A_2 = \cos(2\pi q_{i-1} \tau_{CN}) \cos(2\pi p_{i-1} \tau_{CN}) \quad (4)$$

$$A_3 = \sin(2\pi p_{i-1} \tau_{CN}) \sin(2\pi q_{i-1} \tau_{CN}) \quad (5)$$

$$A_4 = \cos(2\pi p_i \tau_N) \sin(2\pi q_{i-1} \tau_N) \quad (6)$$

$$A_5 = \sin(2\pi p_{i-1} \tau_N) \sin(2\pi q_{i-2} \tau_N) \quad (7)$$

$$A_6 = \cos(2\pi n_{i-1} \tau_{CN}) \quad (8)$$

$$K_1 = \exp(-2T_N R_2^{N_i} - 2\tau_C R_2^{C'_{i-1}} - 2\tau_{CN} R_2^{C^\alpha_{i-1}}) \quad (9)$$

$$K_2 = \exp(-2\tau_N R_2^{N_i}), \quad K_3 = \exp(-2\tau_N R_2^{N_{i-1}}), \quad (10)$$

where $2\tau_C$, $2\tau_{CN}$, $2T_N = A + B + C$ and $2\tau_N = D + E + F$ are the delays required for effecting the magnetization transfers as in figure 1. $R_2^{N_i}$, $R_2^{N_{i-1}}$, $R_2^{C'_{i-1}}$ and $R_2^{C^\alpha_{i-1}}$ are the various transverse relaxation rates corresponding to their respective superscripts, and

$$\begin{aligned} l_i &= {}^1J(C_i^\alpha - C'_i); \quad m_i = {}^1J(N_i - C'_{i-1}); \\ p_i &= {}^1J(C_i^\alpha - N_i); \quad q_i = {}^2J(C_i^\alpha - N_{i+1}); \\ n_i &= {}^1J(C_i^\alpha - C_i^\beta). \end{aligned} \quad (11)$$

In equation set 11, 1J and 2J are one-bond and two-bond coupling constants corresponding to the respective terms in brackets. The value of ${}^1J(N_i - C'_{i-1})$ is generally equal to ~ 15 Hz, thus the delay $2T_N$ should be ~ 33 ms for optimum transfer. However, the average value of one bond or two bond N- C^α coupling constant for the α and β types of structures are slightly different.²⁶ for α helices, ${}^1J_{C^\alpha-N}$ and ${}^2J_{C^\alpha-N}$ are in the ranges of 8–10 and 4–6 Hz, respectively, and for β structures these values are in the ranges of 10–13 and 6–9 Hz, respectively. Thus, based on the coupling constants, the delays $2\tau_{CN}$, and $2\tau_N$ become protein-dependent and thus have to be optimized. The optimum value of $2\tau_N$ for denatured state of a protein (~ 10 – 15 kDa size) generally comes out to be between 24–30 ms. Among the various evolutions and the transfers occurring through the pulse sequence, the evolution during the period $2\tau_{CN}$ is the most crucial from the sensitivity point of view, as, it is the period during which magnetization resides on the C^α carbon which has the

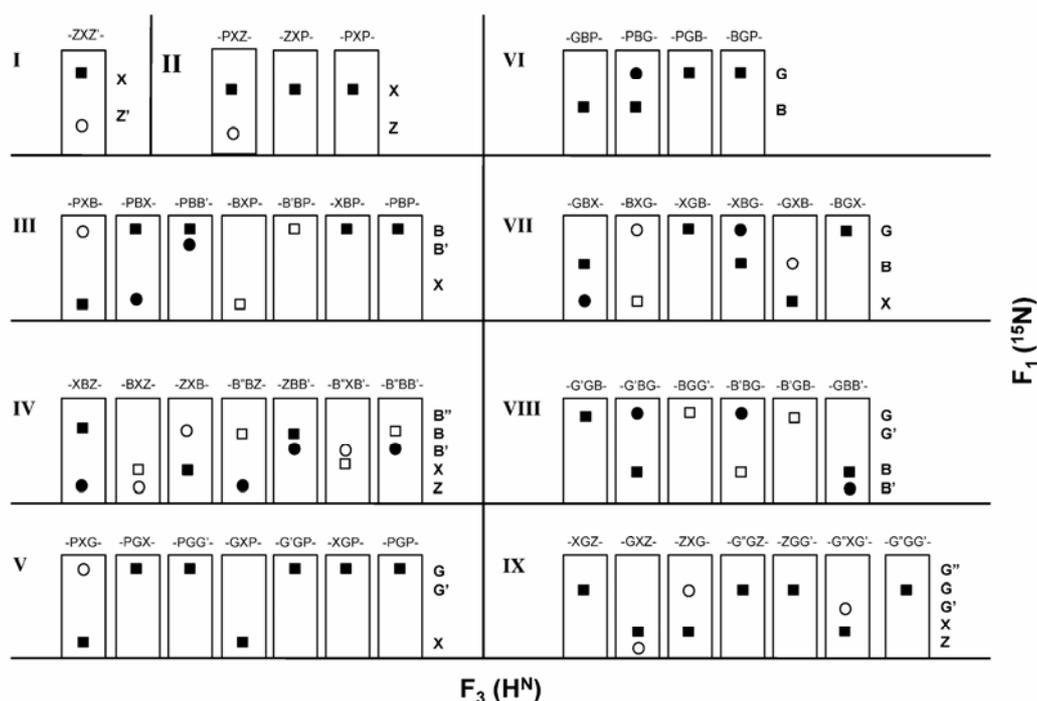


Figure 2. Schematic patterns in the $F_1(^{15}\text{N})$ – $F_3(\text{H}^{\text{N}})$ planes at the $F_2(^{15}\text{N})$ chemical shift of the central residues in the triplets mentioned on the top of each panel in HN(C)N–ST spectra for various special triplet sequences of categories (i)–(ix) (see text). X, Z, and Z' are any residue other than proline, glycine, serine/threonine. G, G' and G'' are glycines, P is proline, and B, B' and B'' stand for serines/threonines in the sequence. Diagonal peaks are shown by squares and sequential peaks by circles where black filled circles/squares are for positive peaks and open circles/squares are for negative peaks. In all the cases the peaks are aligned at the $F_3(\text{H}^{\text{N}})$ chemical shift of the central residue.

fastest transverse relaxation rate, and its minimization is crucial for highest sensitivity in the spectrum.

The C^α evolution during $2\tau_{\text{CN}}$ has another important consequence. A careful examination of the (equations 3–8) reveals that for the optimum choices of the transfer periods, the functions A1 to A5 are all positive while, A6 is negative. This means depending on the presence or absence of A6 (which originate from evolution under C^α – C^β coupling during $2\tau_{\text{CN}}$ period) the cross peaks can have negative or positive signs, respectively. This is precisely the basis of generation of different patterns of peaks due to glycines which do not have a C^β carbon.²¹ These generate the check points for sequential walks in the HN(C)N spectrum. It can be recognized easily that by tampering with the C^α evolution, especially with the C^α – C^β coupling evolution, during this period, it should be possible to manipulate the peak patterns for specific residues. In this context the 180° pulse (encircled in figure 1), which serves the purpose of (i) refocusing C^α chemical shifts and (ii) keeping ac-

tive evolutions under C^α – C^β coupling and C^α –N coupling, plays a crucial role. What this implies is that if a particular evolution under C^α – C^β coupling during $2\tau_{\text{CN}}$ period is prevented either by selective unlabelling of C^β carbon or by specific decoupling of C^β , then additional check points can be created in the 3D spectrum; the easy option here is to decouple the C^β and this is the basis of the present modification. These become particularly necessary if the protein sequence does not contain glycines for long stretches.

Among the various amino acid residues, the C^β chemical shifts of alanines and serines/threonines are very distinct. The alanine C^β 's resonate most up-field (18–22 ppm)^{7,27} while serine/threonine C^β 's resonate most down-field (60–79 ppm)^{7,27} as compared to all other residues (27–43 ppm).^{7,27} Next, the glycines not only lack β -carbons but also their C^{α} 's resonate most up-field (40–45 ppm).^{7,27} as compared to all other residues (50–70 ppm).^{7,27} Thus, on the basis of the ranges and distinctiveness of the C^α and

C^β chemical shifts of the different residue types, we consider two types of manipulations:

(i) The 180° pulse can be made band-selective not to invert the C^β of alanines. This will eliminate C^α - C^β coupling evolution for these residues and make them look like glycines in the 3D spectrum. This work has recently been described in detail by Chatterjee *et al*⁶ and the experiment has been named as HN(C)N-A.

(ii) The 180° band-selective pulse can be applied to selectively refocus all the non-glycine C^α carbons (note that any attempt to include C^α carbons of glycines results in inversion of several C^β carbons as well) in the range of 49–79 ppm,^{7,27} to avoid C^α - C^β coupling evolution for all the residues except for serines and threonines – for which the C^β (60–79 ppm)^{7,27} also gets refocused during the same time and C^α - C^β coupling evolves in a normal manner. This is the modification we are currently concerned with and we refer to this experiment as HN(C)N-ST.

The HN(C)N-ST modification has, however, another consequence. Since the 180° pulse excludes C^α of glycines as well, there will be no evolution under C^α -N coupling for glycines. As a result there will be no transfers of magnetization from the glycines to the neighbouring residues. Thus, for the stretch $-G_{i-1}X_i-$, the density operator for this stretch at point f in figure 1 comes out to be:

$$\sigma_f = 2H_{iz}N_{iy}A_{7B}A_{8B}A_{9B} \cos(\Omega_{N_i}t_2) \cos(\Omega_{N_i}t_1) \cos(2\tau_{CN}\Omega_{C_{i-1}^\alpha})K_1K_2, \quad (12)$$

where Ω_{N_i} and $\Omega_{C_{i-1}^\alpha}$ are the chemical shifts of N_i and C_{i-1}^α nuclei, respectively, and

$$\begin{aligned} A_{7B} &= \sin(2\pi l_{i-1}\tau_C) \sin(2\pi m_i T_N); \\ A_{8B} &= \cos(2\pi p_i \tau_N) \cos(2\pi q_{i-1} \tau_N); \\ A_{9B} &= \sin(2\pi J_{C^\alpha CO}(\tau_{CN} - \tau_C)). \end{aligned} \quad (13)$$

All other terms have their meaning as shown above. Note that there is no operator term in equation (13), describing transverse magnetization of N_{i-1} . Consequently, there will be no $(i-1)$ cross peak in the i th (F_2, F_3) plane (where $i-1$ is a glycine) and hence no $(i+1)$ cross peak in the i th (F_1, F_3) plane (where i is a glycine). The peak patterns in the different planes of the 3D spectrum for doublets of residues containing glycines will get modified compared to those in the normal HN(C)N spectrum. We have calculated all

these patterns by setting the appropriate coupling constants to zero in equations (3–8) when evolutions under them have to be ignored. The expected peak patterns in the (F_1, F_3) planes of the 3D spectrum are shown in figure 2. Note that the phases are adjusted here such that the patterns for the general doublets which do not involve a serine/threonine or a glycine are identical to those in the normal HN(C)N spectrum. This is necessary for convenience of assignment.

We have categorized different triplets of residues in nine following ways: (i) ZXZ'; (ii) PXZ, ZXP, PXP; (iii) PXB, PBX, PBB', BXP, B'BP, XBP, PBP; (iv) XBZ, BXZ, ZXB, B''BZ, ZBB', B''ZB', B''BB'; (v) PXG, PGX, PGG', GXP, G'GP, XGP, PGP; (vi) GBP, PBG, PGB, BGP; (vii) GBX, BXG, XGB, XBG, GXB, BGX; (viii) G'GB, G'BG, BGG', B'BG, B'GB, GBB'; and (ix) XGZ, GXZ, ZXG, G''GZ, ZGG', G''XG', G''GG', where X, Z and Z' are any residue other than proline, glycine, serine/threonine. G, G' and G'' are glycines, P is proline, and B, B' and B'' stand for serine/threonine in the sequence (figure 2). Category (i) is a general one not containing any glycine, proline, serine/threonine and has been included to be able to distinguish the special patterns from the general pattern, category (ii) has prolines but no glycines and serines/threonines; category (iii) has combinations of prolines and serines/threonines; category (iv) has serines/threonines but no glycines and prolines; category (v) has combinations of prolines and glycines; category (vi) has combinations of glycines, serines/threonines and prolines; category (vii) has combinations of glycines and serines/threonines but no prolines; category (viii) has combination of two glycines and one/serine/threonine and vice versa, and category (ix) has glycines and no prolines or serines/threonines. Diagonal peaks are shown by squares and sequential peaks by circles; filled circles/squares are for positive peaks and open circles/squares are for negative peaks. These patterns enable unambiguous discrimination between the triplets having glycines, triplets having serines/threonines and triplets having (X) as the central residue. However, careful examination reveals that for the triplets where glycine or proline is at the $i-1$ position the patterns will be similar and this can lead to some ambiguities (see for example, PXZ vs GXZ; PBX vs GBX; among the triplets PGX, PGG, GXP, GGP, PGP, GGZ, GGG and GGB etc.). In these situations the ambiguities can be readily resolved by comparing with the corresponding patterns in the HN(C)N spectrum.

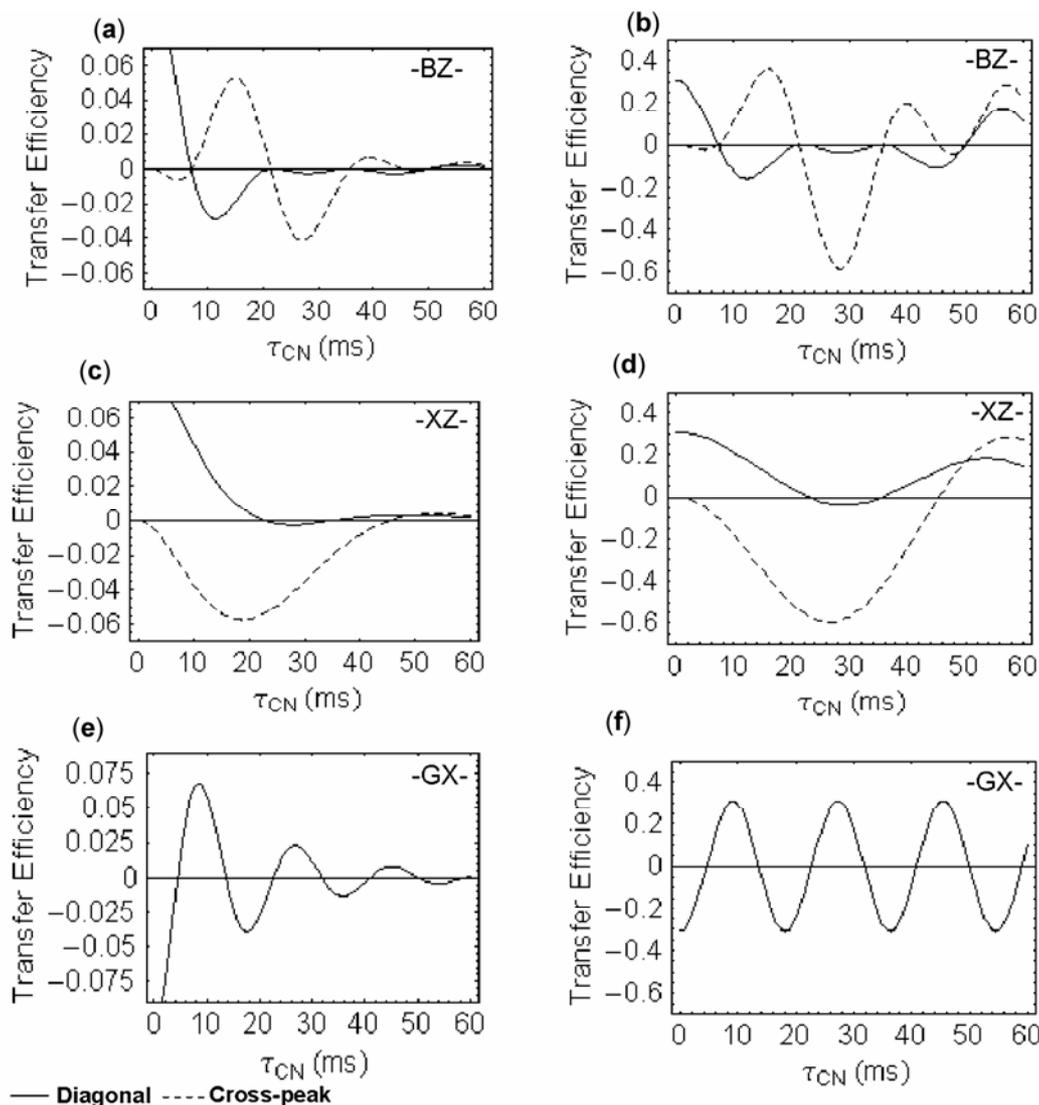


Figure 3. Plots of HN(C)N coherence transfer efficiencies. The transfer functions for the diagonal peak (J_i^d) (solid lines in the plot, eqn (1)) and the cross peak (J_{i-1}^d) (dashed lines in the plot, eqn (2)) are described in the text. (a), (c), and (e) are the transfer efficiencies calculated with relaxation terms included, while (b), (d) and (f) are the transfer efficiencies calculated without the relaxation terms. Here, (a) and (b) are for a stretch of amino acid –BZ– where B are either serine/threonine and Z is for any residue except proline. (c) and (d) are for –XZ– stretch where Z is for any residue except proline while X is for any residue other than serine, threonine, proline or glycine. (e) and (f) are for –GX– stretch where G is for glycine and Z is for any residue except proline. The transfer efficiencies are plotted as a function of τ_{CN} . The plots were calculated by using $J_{C^\alpha C^\beta}$, $J_{C^\alpha C^O}$, and J_{NCO} values of 35, 55 and 15 Hz respectively. The $^1J_{C^\alpha N}$, and $^2J_{C^\alpha N}$, values have been chosen to be 11 Hz and 7 Hz, respectively. The values of T_N , τ_N and τ_C used in the transfer functions for HN(C)N–ST were 15.0 ms, 15.0 ms and 4.5 ms, respectively. The values of transverse relaxation terms R_2^N , R_2^C and R_2^α used here are 30, 10 and 15 s^{-1} , respectively. Solid and dashed lines represent diagonal and sequential peaks, respectively.

For the HN(C)N experiment, an inversion band width of roughly 25 ppm (50–75 ppm) is required. Different kinds of band selective inversion pulses have been described in the literature.^{20,28} Among the

various possibilities, sech/tanh inversion pulses have the flattest inversion profiles over 90% of the band width.²⁹ However, these pulses are long ranging from 5 to 10 ms to meet the desired band-width. The

other desirable option is broadband inversion pulses (BIP).²⁸ These pulses are short, simplest to calculate, but their excitation profiles are not uniformly flat. However, as compared to BIP pulses, Gaussian Cascade pulses (Q3.1000)²⁰ were found to have better inversion profiles, and also the desired bandwidth (~30 ppm) was obtained from only a 650 μ s long Q3.1000 pulse (at 13.98 W). Therefore, the Gaussian Cascade pulses were used for the present experiment.

3.3 Transfer efficiencies and peak intensities

In the three-dimensional spectrum of HN(C)N–ST (also in HN(C)N) the peaks appear at the following coordinates:

$$F_1 = N_i; (F_3, F_2) = (H_i, N_i), (H_{i-1}, N_{i-1}),$$

$$F_2 = N_i; (F_3, F_1) = (H_i, N_i), (H_i, N_{i+1}).$$

Thus, in the (F_2, F_3) plane there occurs a diagonal peak ($F_1 = F_2 = N_i$) for each residue and one sequential peak to $(i - 1)$, while in the (F_1, F_3) plane there occurs a diagonal peak ($F_1 = F_2 = N_i$) and one sequential peak to $(i + 1)$. The transfer efficiencies which dictate the intensities of the diagonal and sequential peaks, can be calculated from the equations described in the previous section. Appropriate adjustments in the coefficients will have to be made for the special cases as has been mentioned for the calculation of the peak patterns. However, for this calculation, we may recall that magnetization originating on residue i results in only two peaks at the end of the pulse sequence, one to itself and the other to the $i - 1$ residue. Therefore, for every case, we need to calculate the intensities of one diagonal peak and one sequential peak only. Thus considering different doublets of residues, covering the general and all the special situations, we explicitly calculated the peak intensities as a function of τ_{CN} , the most crucial adjustable parameter and these are shown in figure 3. The calculations are shown with and without explicit inclusion of relaxation effects.

The HN(C)N–ST experiment has one particular advantage compared to normal HN(C)N, and that is in the sensitivities of the diagonal and the sequential peaks. This arises from the fact that the term Γ_6 in (7) will be unity for all non-serine/threonine residues. For the typical 12–14 ms choice of τ_{CN} τ_{CN} value and a C^α – C^β coupling constant of 35 Hz, this yields a roughly 15% gain in sensitivity.

3.4 Experimental demonstration

We have tested the HN(C)N–ST experiment with different proteins and we show here results on two proteins under different conditions: (i) M-crystallin protein (apo form; folded; 85 residues) and (ii) urea-denatured GED of dynamin (denatured protein; 138 residues); the assignments of which have already been published^{18,19} and thus they served as an excellent test cases. In either case almost all the expected correlations were observed in the HN(C)N–ST spectrum. Figures 4a and b show a comparison of the HN(C)N and HN(C)N–ST patterns respectively, for the stretch Ala3 to Tyr11 in the case of urea-denatured GED.¹⁹ The type of triplet stretch has been shown at the top of each panel. In the first strip (Ala3 diagonal), diagonal peak is positive in sign due to $i-1$ serine (Ser2) while peak corresponding to Ser4 residue is also of positive sign, this belongs to triplet –B''XB'– (panel IV of figure 2). In the last strip (Tyr11 diagonal), no sequential peak is there since the residue at position 12 is a proline (Pro12) which does not produce any peak, this belongs to triplet –ZXP– (panel II in figure 2). In the seventh strip of HN(C)N–ST spectrum (Gly9 diagonal), only negative diagonal peak is there while there is no sequential $i + 1$ cross-peak corresponding to Val10 (position highlighted by dashed circle); this belongs to triplet –XGZ– (panel IX in figure 2). However, the latter is present in normal HN(C)N spectrum as can be seen from figure 4b. Figure 4c depicts the different triplet patterns along the sequence.

To highlight the value of HN(C)N–ST in the sequential assignment of any protein, a comparison of the number of check points in HN(C)N, HN(C)N–A and HN(C)N–ST experiments for GED, has been depicted in figure 5a. Clearly, the larger number and good distribution of the checkpoints across the length of the chain provide great boost to the speed and correctness of the connectivities. Indeed, as can be seen in figure 5b, 90% of the sequential connectivities could be unambiguously established from this spectrum alone even for urea-denatured GED.

Figure 6 shows a sequential walk illustration through HN(C)N–ST in the case of the folded protein, M-crystallin; assignments used here have already been reported.¹⁸ It also demonstrates the significance of the distinctive patterns as check points during the sequential walks. The Gly62 plane (figure 6) shows only one peak as the residue 63 being a proline does not have amide proton and thus the peak correspond-

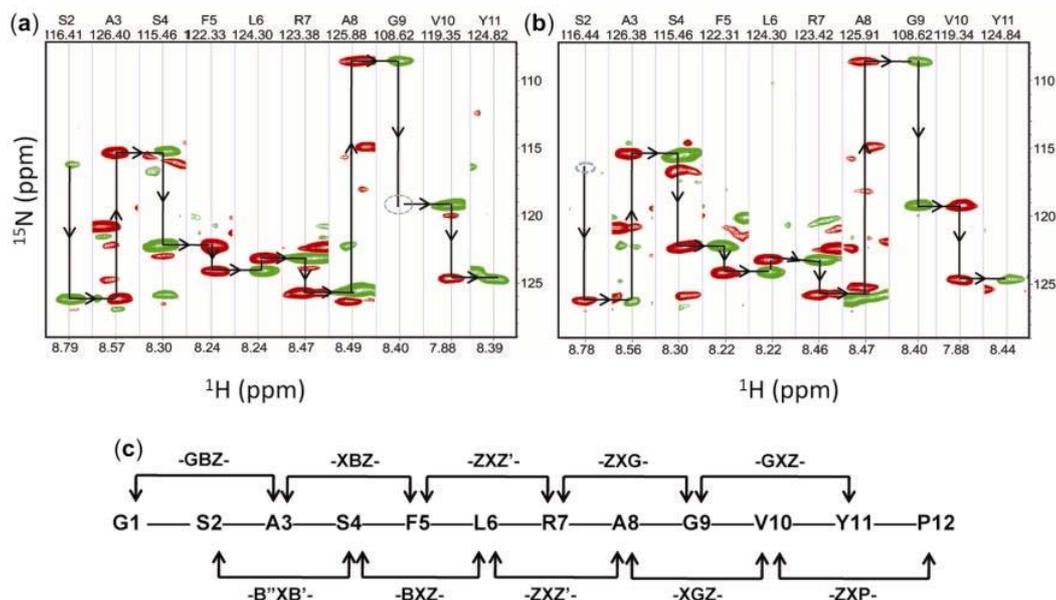


Figure 4. Comparison of $F_1(^{15}\text{N})-F_3(^1\text{H})$ strips of the (a) HN(C)N-ST and the (b) HN(C)N spectra for the residues Ala3 to Tyr11 in 9.7 M urea-denatured state GED protein. Each residue belongs to which strip in figure 4, has been clearly indicated in panel (c). The $F_2(^{15}\text{N})$ values, which help to identify diagonal peaks, are shown at the top for each strip. The green and red contours indicate positive and negative peaks, respectively. The residues corresponding to diagonal peaks in each strip have been shown at the top of each strip. The distinguishing behaviour for all the serine (Ser) and threonine (Thr) residues is clearly evident from the different patterns of positive and negative peaks depending on the position of serine or threonine. These residues also serve as starting points for the sequential walk, as if serine/threonine is in $i - 1$ position, both diagonal and sequential peaks for i th residue will come negative in HN(C)N-ST spectra. Position for Val10 cross-peak on the Gly9 plane is encircled by dashed line to show that the peak is absent because $^1J_{\text{NC}\alpha}$ coupling is not refocused (see text).



Figure 5. (a) The residues serving as start/check points (including prolines which has been underlined to emphasize their position as stop/check points) in different types of HN(C)N experiments have been marked in red along the amino acid sequence of GED of dynamin and their numbers have been shown on the right hand side of the sequence. (b) Summary of all the observed sequential connectivities along the sequence of the protein. The thick line indicates the presence of connectivity.

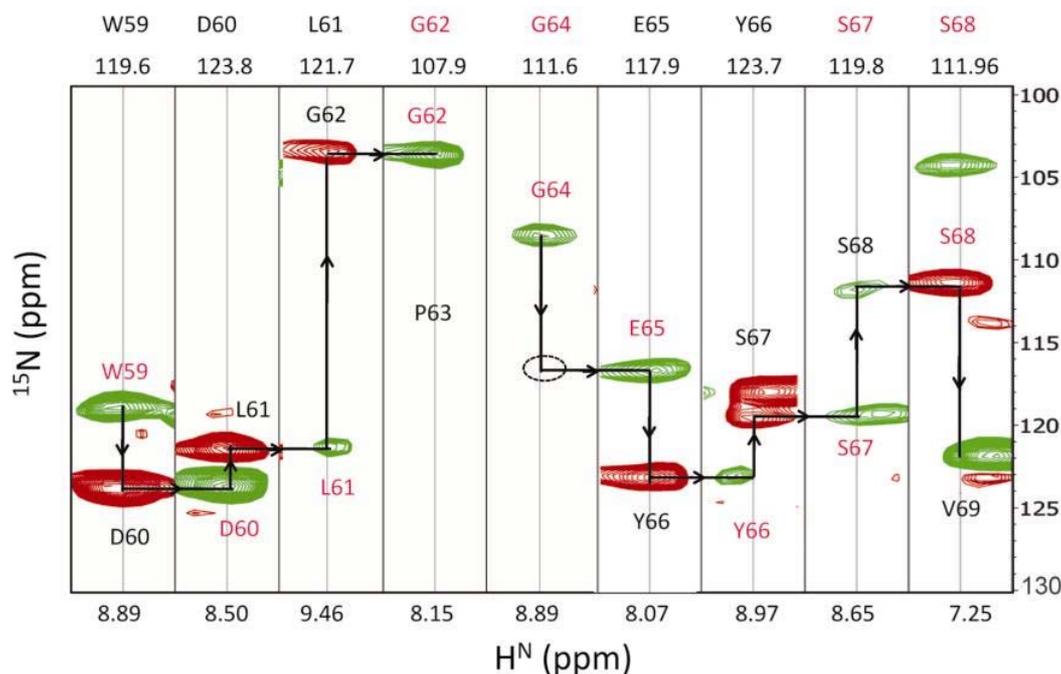


Figure 6. Illustrative of sequential walk through $F_1(^{15}\text{N})-F_3(\text{H}^{\text{N}})$ planes of HN(C)N-ST spectra for M-crystallin protein (apo form; folded) in 50 mM Tris buffer containing 150 mM KCl at pH 7.5. The $F_2(^{15}\text{N})$ values, which help to identify diagonal peaks, and the residues corresponding to diagonal peaks are shown at the top for each strip. The green and red contours indicate positive and negative peaks, respectively. The position for $i + 1$ peak (E65) in the G64 diagonal-plane is encircled to emphasize the absence of that peak due to selective refocusing of all non-glycine C^α carbons (see text). G62, G64 S67, and S68 labels at the top are in red to emphasize the distinctive features arising due to these residues (can be related to figure 2).

ing to $i + 1$ position will be absent (dashed circle), this belongs to triplet -XGP- depicted in panel V of figure 2. The E65 plane has typical doublet, but note here that the diagonal is positive; this plane belongs to triplet -GXZ- (panel IX of figure 2) which is equivalent to triplet -PXZ- (panel II of figure 2). This is due to decoupling of C^α -N couplings during the τ_{CN} period, hence no magnetization transfer occurs from X (Glu) to Gly. Thus, glycine will behave like proline if it is at first position of any triplet. Similarly, $i + 1$ residue will behave like proline if glycine is in the middle position of the triplet. The pattern for Tyr66 belongs to triplet -ZXB- (panel IV of figure 2) explicating that only first and middle residues of the triplet are involved in deciding the signs of diagonal and cross peaks in the (F_1 , F_3) planes; this is evident from the similar patterns for triplets: -ZXG- (panel IX in figure 2), -ZXP- (panel II in figure 2), and -ZXB- (panel IV in figure 2). Further, the pattern for Ser67 plane (triplets -ZBB'-, panel IV, figure 2) is as must be expected for the HN(C)N-ST.

4. Conclusions

In conclusion, we believe that the HN(C)N-ST experiment together with HN(C)N and HN(C)N-A will provide a substantial boost to the sequential assignment process in labelled proteins. It will be particularly of greater value if the protein sequence has none or few far apart glycines and alanines along the amino acid sequence of the polypeptide chain as has been seen in the case of GED.

Acknowledgements

We thank Government of India for providing financial support to the National Facility for High Field NMR at Tata Institute of Fundamental Research, India. We acknowledge Dr Shilpy Sharma for providing us [^{15}N , ^{13}C] labelled unfolded (9.7 M urea-denatured) GED sample and Atul K Srivastava for the useful discussions during the manuscript preparation.

References

1. Yee A, Gutmanas A and Arrowsmith C H 2006 *Curr. Opin. Struct. Biol.* **16** 611
2. Wuthrich K 1986 *NMR of proteins and nucleic acids* (New York: John Wiley & Sons)
3. Wagner G 1997 *Natl. Struct. Biol.* **4 Suppl.** pp 841–844
4. Bhavesh N S, Panchal S C and Hosur R V 2001 *Biochemistry* **40** 14727
5. Chatterjee A, Bhavesh N S, Panchal S C and Hosur R V 2002 *Biochem. Biophys. Res. Commun.* **293** 427
6. Chatterjee A, Kumar A and Hosur R V 2006 *J. Magn. Reson.* **181** 21
7. Chugh J, Kumar D and Hosur R V 2008 *J. Biomol. NMR* **40** 145
8. Bhavesh N S, Sinha R, Mohan P M and Hosur R V 2003 *J. Biol. Chem.* **278** 19980
9. Bhavesh N S, Juneja J, Udgaonkar J B and Hosur R V 2004 *Protein Sci.* **13** 3085
10. Chatterjee A, Mridula P, Mishra R K, Mittal R and Hosur R V 2005 *J. Biol. Chem.* **280** 11369
11. Chatterjee A, Krishna Mohan P M, Prabhu A, Ghosh-Roy A and Hosur R V 2007 *Biochimie* **89** 117
12. Chugh J, Chatterjee A, Kumar A, Mishra R K, Mittal R and Hosur R V 2006 *FEBS J.* **273** 388
13. Chugh J, Sharma S and Hosur R V 2007 *Biochemistry* **46** 11819
14. Chugh J, Sharma S, Kumar D and Hosur R V 2009 *Biomolecular NMR Assignment*
15. Kumar A, Srivastava S, Kumar M R, Mittal R and Hosur R V 2006 *J. Mol. Biol.* **361** 180
16. Kumar A, Srivastava S, Mishra R K, Mittal R and Hosur R V 2006 *Biophys. J.* **90** 2498
17. Kumar D, Kumar A, Misra J R, Chugh J, Sharma S and Hosur R V 2008 *1H, 15N, 13C resonance assignment of folded and 8 M urea denatured state of SUMO from Drosophila melanogaster* pp. 13–15
18. Barnwal R P, Jobby M K, Sharma Y and Chary K V 2006 *J. Biomol. NMR* **36 Suppl 1** 32
19. Chugh J, Sharma S, Kumar D and Hosur R V 2009 *Biomolecular NMR Assignment* **3** 13–16
20. Emsley L and Bodenhausen G 1990 *Gaussian pulse cascades: New analytical functions for rectangular selective inversion and in-phase excitation in NMR* pp. 469–476
21. Panchal S C, Bhavesh N S and Hosur R V 2001 *J. Biomol. NMR* **20** 135
22. Piotto M, Saudek V and Sklenar V 1992 *J. Biomol. NMR* **2** 661
23. Marion D, Ikura M, Tschudin R and Bax A 1989 *Application to the study of hydrogen exchange in proteins*, pp. 393–399
24. Grzesiek S and Bax A 1993 *J. Am. Chem. Soc.* **115** 12593
25. Bhavesh N S, Chatterjee A, Panchal S C and Hosur R V 2003 *Biochem. Biophys. Res. Commun.* **311** 678
26. Wirmer J and Schwalbe H 2002 *J. Biomol. NMR* **23** 47
27. BMRB DataBase 2009 http://www.bmrb.wisc.edu/ref_info/statful.htm#1
28. Smith M A, Shaka A J and Hu H 2001 *J. Magn. Reson.* **151** 269
29. Silver M S, Joseph R I and Hoult D I 1985 *Phys. Rev.* **A31** 2753
30. Shaka A J, Keeler J and Freeman R 1983 *Evaluation of a new broadband decoupling sequence: WALTZ-16* pp. 313–340
31. Shaka A J, Keeler J, Frankiel T and Freeman R 1983 *An improved sequence for broadband decoupling: WALTZ-16* pp. 335–338
32. Shaka, A J, Barker P B and Freeman R 1985 *Computer-optimized decoupling scheme for wideband applications and low-level operation* pp. 547–552