

Addressing the overlap problem in the quantitative analysis of two dimensional NMR spectra: Application to ¹⁵N relaxation measurements

Vitali Tugarinov^a, Wing-Yiu Choy^a, Eriks Kupče^b & Lewis E. Kay^a

^aProtein Engineering Network Centres of Excellence and the Departments of Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, Ontario, Canada M5S lA8 ^bVarian Ltd., Eynsham, Oxford, U.K.

Received 30 April 2004; Accepted 9 July 2004

Key words: HNCO, ¹⁵N relaxation, plane projections, resonance overlap

Abstract

A quantitative analysis of 2D ${}^{1}\text{H}{}^{15}\text{N}$ spectra is often complicated by resonance overlap. Here a simple method is presented for resolving overlapped correlations by recording 2D projection planes from HNCO data sets. Applications are presented involving the measurement of ${}^{15}\text{N}$ T_{1 ρ} relaxation rates in a high molecular weight protein, malate synthase G, and in a system that exchanges between folded and unfolded states, the drkN SH3 domain. By supplementing relaxation data recorded in the conventional way as a series of 2D ${}^{1}\text{H}{}^{-15}\text{N}$ data sets with a series of a pair of projection planes the number of dynamics probes is increased significantly for both systems studied.

Over the past decade a large number of 2D heteronuclear correlation experiments have been developed for quantifying a wide range of molecular properties. Such 2D spectra have the advantage in that they can be recorded rapidly and in that they are intrinsically more sensitive than their higher dimensionality counterparts, but a major limitation with quantification results from resonance overlap. Possible solutions involve the use of new processing schemes which can improve resolution (Gutmanas et al., 2004) or extending the 'classical' 2D experiments to 3D, using sensitive HNCO data sets, for example (Caffrey et al., 1998; Tugarinov and Kay, 2003). Here we describe an alternative strategy that is based very closely on the reduced-dimensionality methodology of Szyperski and coworkers (Kim and Szyperski, 2003, 2004) and on the projection-reconstruction approach of Kupce and Freeman (2003a-c, 2004a,b).

To illustrate the utility of the approach, we focus on the measurement of protein backbone $^{15}N T_{1\rho}$ relaxation rates. Typically, such rates are measured from a series of 2D $^{1}H^{-15}N$ correlation maps and resolution is critical in applications involving high molecular weight monomeric proteins or protein systems with substantial populations of disordered or partially unfolded states. It would be advantageous, therefore, to introduce a third chemical shift without increasing the dimensionality of the experiments, by simultaneous incrementation of a pair of indirect evolution delays, t_1 and t_2 , in a 3D-type experiment. Here we increment $t_1({}^{13}C')$ and $t_2({}^{15}N)$ in HNCO spectra that have been adapted for ${}^{15}N$ relaxation measurements. The utility of the methodology is demonstrated with applications to malate synthase G (Howard et al., 2000; Tugarinov et al., 2002), a 723-residue enzyme (MSG, 82 kDa) and to the N-terminal SH3 domain from drk (drkN SH3) that exists in equilibrium between folded and unfolded states under physiological buffer conditions (Zhang and Forman-Kay, 1995; Zhang et al., 1997).

Figure 1 illustrates the TROSY-HNCO based pulse scheme that has been developed to measure ¹⁵N T_{1ρ} rates in high molecular weight proteins. The design of the pulse scheme follows very closely from 2D ¹⁵N T_{1ρ} and 3D HNCO sequences that exist in the literature (Peng and Wagner, 1992; Yang and Kay, 1999) and only the salient features will be described here. Briefly, at point *a* in the scheme transverse ¹⁵N magnetization from spin *j* is locked along an effective field that makes an angle $\theta = \tan^{-1} (v_1/\Delta v_j)$ with respect to the *z*-axis, where v₁ is the strength of the spin-lock field



Figure 1. TROSY-based pulse scheme for the measurement of $^{15}NT_{10}$ relaxation rates from HNCO-plane projections. All narrow (wide) rectangular pulses are applied with flip angles of 90° (180°) along the x-axis unless indicated otherwise. The ¹H, ¹⁵N and ¹³C' carriers are positioned at 4.7, 119 and 176 ppm, respectively. Water magnetization is preserved and placed along the +z-axis at the end of the sequence; the pulse labeled 'w' prior to G3 is water selective (1.5 ms). In order to suppress cross-correlation/cross-relaxation effects a small number (\leq 3) of ¹H 180° pulses are applied during the spin lock interval T (Kay et al., 1992b; Palmer et al., 1992). Denoting the maximum value of T as T_{max} , m=1, 2 or 3 ¹H pulses are applied if $0 \le T < T_{\text{max}}/3$, $T_{\text{max}}/3 \le T < 2T_{\text{max}}/3$ or $2T_{\text{max}}/3 \le T \le T_{\text{max}}$, respectively. The total number of ¹H pulses is kept constant at 3 by applying (3-m) pulses prior to G5. 90° ¹³C' pulses are applied with a field of $\Delta/\sqrt{15}$, while 180° ¹³C' and ¹³C^{α} pulses employ a field of $\Delta/\sqrt{3}$, where Δ is the difference (Hz) between the centers of the ¹³C^{α} (58 ppm) and ¹³CO (176 ppm) chemical shift regions (Kay et al., 1990). ¹³C^{α} pulses are applied off-resonance using phase modulation of the carrier (Boyd and Soffe, 1989; Patt, 1992). The vertical arrow denotes the position of the Bloch-Siegert shift compensation pulse (Vuister & Bax, 1992). Nitrogen pulses are applied with a field strength of 7 kHz, while an ¹⁵N spin-lock (phase x) between 1.8–2 kHz was employed. The phase cycling used is: $\phi 1 = x, -x; \phi 2 = x;$ $\phi 3 = 2(x), 2(y), 2(-x), 2(-y); \\ \phi 4 = x, -x; \\ \phi 5 = 4(y), 4(-y); \\ \phi 6 = 8(x), 8(-x); \\ \phi 7 = x; \\ \operatorname{rec} = x, -x, -x, x, -x, x, -x. \\ \operatorname{The delays used are:}$ $\tau_a = 2.3 \text{ ms}, \varepsilon = 2.68 \text{ ms}, T_N = 13 \text{ ms}, \beta = 250 \text{ } \mu\text{s}, \sigma = 2^* \text{pwn} - \text{pwca180}, \text{ where 'pwn' ('pwca180') is the length of the } ^{15} \text{N } 90^\circ (^{13} \text{C}^{\alpha})^{13} \text{ } \text{C}^{\alpha}$ 180°) pulse (¹⁵N and ¹³C^{α} pulses are applied simultaneously in the center of t_1 , and t_1 , t_2 are incremented by $sw^{-1}sin(\delta)$, $sw^{-1}cos(\delta)$ respectively, as described in the text. Delays ζ and ξ are set to $1/(2\pi^*\nu_1)$ and $1/(2\pi^*\nu_1) - 4^*$ pwn/ π , respectively, with ν_1 the spin-lock field strength (Griesinger and Ernst, 1987; Yamazaki et al., 1994). The durations and strengths of the pulsed field z-gradients are: $G_1 = (0.5 \text{ ms}, 1000 \text{ ms})$ 8 G/cm, G2 = (0.5 ms, 5 G/cm), G3 = (1 ms, 15 G/cm), G4 = (0.6 ms, -10 G/cm), G5 = (0.8 ms, -7.5 G/cm), G6 = (0.75 ms, 10 G/cm), G6 = $G7 = (1 \text{ ms}, 10 \text{ G/cm}), G8 = (1.25 \text{ ms}, 30 \text{ G/cm}), G9 = (0.4 \text{ ms}, 2.9 \text{ G/cm}), G10 = (0.4 \text{ ms}, 5.35 \text{ G/cm}), G11 = (62.5 \mu \text{s}, 28.7 \text{ G/cm}).$ Shared constant-time evolution is implemented in t_2 (Grzesiek and Bax, 1993; Logan et al., 1993); for evolution times $t_2 > 2^*T_N$ the scheme is modified as indicated. (It is worth noting that the use of linear prediction can distort peak intensities in the case where data is acquired in the shared constant-time mode). Four data sets are recorded for each simultaneous increment of (t_1, t_2) : (a) $\phi 2 = x$, $\phi 7 = x$; (b) $\phi 2 = y$, $\phi 7 = x$; (c) $\phi 2 = x$, $\phi 7 = -x$ with the sign of gradient G8 inverted; and (d) $\phi 2 = y$, $\phi 7 = -x$ with the sign of G8 inverted. Pairs a and c, b and d are processed using the enhanced sensitivity pulsed field gradient method (Kay et al., 1992a; Schleucher et al., 1993) and subsequently combined as described in the text. A non-TROSY version of this scheme has also been developed and is available from the authors upon request.

applied for time T and Δv_i is the offset of the spin (Griesinger and Ernst, 1987; Yamazaki et al., 1994). Subsequently, the flow of magnetization follows that which is typical for an HNCO. The values of t_1 and t_2 are incremented by $sw^{-1} * \sin(\delta)$ and $sw^{-1} * \cos(\delta)$ respectively, where sw is the spectral width in the indirectly detected dimension and δ is a 'projection angle' (see below). Note that the value of t_2 can exceed $2T_N$, Figure 1, so that resolution is not limited by the requirement to refocus ¹⁵N magnetization that is coupled to ${}^{13}C'$, which occurs during the interval where ¹⁵N chemical shift is recorded. In this regard it is also possible to use the first 2T_N period to record ¹⁵N chemical shift as well (McCoy, 1998), although this has not been implemented here. For each simultaneous increment of (t_1, t_2) a set of four spectra are obtained in exactly the same manner as for the regular enhanced sensitivity 3D-HNCO experiment (Yang and Kay, 1999) and these spectra are manipulated to

produce four data sets modulated in the indirect dimension by

$$a = \cos(\Omega_{\rm N} t_2) \cos(\Omega_{\rm C'} t_{-1} t_1),$$

$$b = \cos(\Omega_{\rm N} t_2) \sin(\Omega_{\rm C'} t_{-1} t_1),$$

 $c = \sin(\Omega_{N,i}t_2) \cos(\Omega_{C',i-1}t_1)$, and

 $d = \sin(\Omega_{N,i}t_2)\sin(\Omega_{C',i-1}t_1)$, (software available upon request). Subsequent addition and subtraction of the data sets as described in the caption to Figure 1 and in detail by Kupce and Freeman (2004b) gives two hypercomplex sets of interferograms that are modulated in t by $\cos[(\Omega_{N,i}\cos\delta\pm\Omega_{C',i-1}\sin\delta)t]$ $(a\mp d)$, $\sin[(\Omega_{N,i}\cos\delta\pm\Omega_{C',i-1}\sin\delta)t]$ $(c\pm b)$, where $\Omega_{N,i}$ and $\Omega_{C',i-1}$ are frequency offsets of residues i and i-1 from the ¹⁵N and ¹³C' carriers, respectively. Finally, Fourier transformation leads to spectra with correlations at $(\Omega_{+i}, \Omega_{HN})$ and $(\Omega_{-i}, \Omega_{HN})$, where

$$\Omega_{\pm,i} = \Omega_{\mathrm{N},i} \cos(\delta) \pm \Omega_{\mathrm{C}',i-1} \sin(\delta) \tag{1}$$

In this manner a pair of projection planes is obtained where, for each Ω_{HN} , correlations that would normally



Figure 2. Computed number of resolved cross-peaks expected in projection data sets of MSG and of the drkN SH3 domain that are *unique* from those resolved in ¹H-¹⁵N correlation maps, $\delta = 0$, as a function of δ . A total of 690 and 114 backbone amide probes are available in MSG and in the drkN SH3 domain, respectively. $\delta = 0^{\circ}$ and $\delta = 90^{\circ}$ correspond to ¹H-¹⁵N and ¹H-¹³C' plane projections, respectively. The computation is based on Equation 1 and the known ¹H,¹⁵N and ¹³C' chemical shifts of each protein (i.e., assignment tables) and assumes (for MSG) that pairs of peaks are overlapped if their centers are separated by less than 30 and 50 Hz in the indirect and direct (¹H) dimensions, respectively. In the case of the drkN SH3 domain peaks separated by less than 12 and 20 Hz in F₁,F₂ are considered overlapped. Thus, for a given value of δ a 'computed' peak-table is generated for the projection data set. For each correlation not resolved in regular ¹H-¹⁵N correlation maps ($\delta = 0$) a search through the 'computed table' establishes whether it is resolved in the projection (based on the criteria listed above).

appear in a $(\Omega_{C',i-1}, \Omega_{N,i})$ plane of the 3D HNCO are projected along a line that makes an angle of $\pm \delta$ with respect to the ¹⁵N axis. (Note that in previous publications Kupce and Freeman (2003b,c) define the projection angle α with respect to the ¹³C' axis, *i.e.*, $\alpha = 90^{\circ}$ - δ). The intensity of each of the correlations is attenuated by the factor $exp(-T/T_{10})$, corresponding to the decay of magnetization during T, from which the value of $T_{1\rho}$ is extracted on a per-residue basis, as in conventional ¹H-¹⁵N correlation experiments. Note that if a set of projections were acquired with different values of δ for each value of T, a set of 3D HNCO relaxation data sets could be reconstructed, following the procedures of Kupce and Freeman (2003a,b) and Coggins et al. (2004). We prefer not to do this here since the process of reconstruction may distort crosspeak intensities, depending on the algorithm used, and because for large proteins the sensitivity of a given plane projection decreases significantly with δ (see below).

Figure 2 plots the number of unique cross-peaks that would be expected to be observed in projection data sets of MSG and of the drkN SH3 domain as a function of δ (i.e., peaks that are overlapped in ¹H-¹⁵N correlation maps, $\delta = 0$, but are resolved in either the $+\delta$ or $-\delta$ projection). The number of unique correlations increases rapidly with δ for values less than approximately 10° and then levels off. However, the effective decay rate in the indirect dimension, t, increases with δ as $R_{2,C'} \sin \delta$ (in the ¹⁵N dimension constant-time acquisition is employed for $t_2 < 2T_N$), and in applications involving large proteins at high magnetic fields $R_{2,C'}$ can be large $(R_{2,C'} \sim 100 \text{ s}^{-1})$, while by comparison $R_{2,N}(\text{TROSY}) \sim 15 \text{ s}^{-1}$ for MSG at 37 °C, 800 MHz (Tugarinov et al., 2002)). Average decreases in signal-to-noise by factors of 1.4 and 1.7 were obtained for spectra of MSG recorded at 37 °C, 800 MHz with δ = 10° and 20°, respectively, relative to $\delta = 0^{\circ}$. We have, therefore, chosen to record a pair of data sets with $\delta = 10^{\circ}$ (i.e., $\pm 10^{\circ}$ projections) as a compromise between the need to resolve

domain (Zhang et al., 1997) was used, 5° C, 500 MHz. Acquisition times of (N-1)sins/sw and (N-1)coss/sw were employed in t_1 , t_2 respectively, where N is the number of complex (t_1 , t_2) points and sw is the spectral width in the indirect dimension of the projection data set; ($t_1 \max, t_2 \max,$ Figure 3. Comparison of selected regions from spectra corresponding to the first relaxation delay (T = 2 ms) of conventional 2D ¹H-¹⁵N T₁₀ data sets (a,f,k) and 2D HNCO plane projections, $+10^{\circ}$ (b,g.)) and -10° (c,h.m), illustrating how overlap in ¹H-¹⁵N data sets can be resolved in the projections. (Note that correlations in the indirect dimension of each projection are at positions given by Equation 1; indirect detection axes are labeled as for the conventional data set). Relaxation decay curves for residues that become resolved in the projections are indicated for G183 (d,e) and A321 (i,j) of MSG and for N29, E40 of the drKN SH3 domain (n,o). The derived relaxation times are indicated at the bottom of each plot. A 0.9 mM sample of U-1²H, ¹⁵N, ¹³C] MSG was employed, 37 °C, 800 MHz, prepared as described previously (Tugarinov et al., 2002), while a 0.6 mM (0.3 mM in folded and unfolded protein) U-¹⁵N, ¹³C sample of the drkN SH3 relaxation delay of 2 s was used along with N=80(MSG) and 72(drkN SH3) for net acquisition times of 2.8 (MSG) and 1.3 hours/spectrum.





Figure 4. ¹⁵N $T_{1\rho}$ relaxation times (ms) in MSG vs. residue number. Residues whose correlations are resolved in conventional 2D ¹H-¹⁵N data sets and for which accurate $T_{1\rho}$ values can be extracted are shown in black. Additional probes available from at least one of the $\pm 10^{\circ}$ projections are shown in red. Errors in $T_{1\rho}$ were obtained from 120 Monte-Carlo simulations (Kamith and Shriver, 1989) using the noise-floor as an estimate of errors in peak intensities.

as many additional peaks as possible and sensitivity concerns. In this regard the use of cryo- or cold-probes should prove particularly advantageous.

Figure 3A illustrates the utility of the projection approach for G183 (a-e) and A321 (f-j) of MSG. The correlation derived from G183 is partially overlapped with cross-peaks from G602 and G218 in the standard 2D ${}^{1}\text{H}$ - $\overline{{}^{15}}\text{N}$ correlation map (a), but becomes well resolved in both $\pm 10^{\circ}$ projections (b,c). Similarly the partial overlap involving A321 and V389 in the ¹H-¹⁵N plane (f) is resolved in both projection-spectra (g,h), allowing the quantification of T_{10} values for these residues. Residues N29 and E40 from the unfolded state of the drkN SH3 domain are completely overlapped in the 2D ¹H-¹⁵N map, yet are resolved in the $+10^{\circ}$ projection so that ¹⁵N T₁ values can be obtained, Figure 3B. It is noteworthy that all spectra recorded on MSG were of the TROSY variety (Pervushin et al., 1997) (Figure 3A), while non-TROSY schemes were used in the case of the SH3 domain (Figure 3B).

Correlations in crowded regions of spectra of MSG and of the drkN SH3 domain were quantified only in cases where the separation between nearest neighbors in *both* direct and indirect dimensions is more than 0.75 of the sum of their linewidths. In this manner an *additional* 114 peaks could be resolved for MSG from the $\pm \delta$ projections (relative to conventional ¹H-¹⁵N maps) and 45 of these correlations had sufficient signal-to-noise for accurate quantification of T₁_ρ values. By means of comparison, using the same criteria for resonance overlap in 2D ¹H-¹⁵N T₁_ρ spectra, 197 cross-peaks were resolved and 119 of these could be fit reliably. The smaller fraction of peaks that could be well fit in the case of the projection data sets reflects the fact that these experiments are on average 3-4 fold less sensitive ($\delta = 10^{\circ}$) than the corresponding ¹H-¹⁵N maps. In this regard it should be noted that the sensitivity/unit measuring time is no worse in the projection 2D data sets than for corresponding 3D HNCO-based spectra (Caffrey et al., 1998; Tugarinov and Kay, 2003). In cases where sensitivity is not limiting, such as in many applications involving (partially) unfolded proteins, or involving proteins of intermediate size, the projection methodology offers distinct advantages over 3D approaches since relaxation data can be obtained in shorter measurement periods, with projection angles chosen suitably to remove overlap.

Figure 4 plots ¹⁵N $T_{1\rho}$ values *vs.* residue number in MSG, with black (red) circles delineating those residues for which relaxation times were obtained from standard ¹H-¹⁵N (projection) spectra. The projection data sets were also particularly useful in the case of the drkN SH3 domain where 18 additional peaks belonging to the unfolded state could be quantified, increasing by a factor of 1.8 the number of probes that are available relative to ¹H-¹⁵N experiments. In addition, a further 10 peaks from the folded state (+24%) were resolved in at least one of the projections.

In summary, a simple approach for increasing the number of residues available for quantification in 2D correlation maps is presented and illustrated in the context of ^{15}N spin relaxation measurements. It is

clear that this methodology can be applied to a large number of different experiments and that for cases involving large proteins or unfolded protein states the benefits will be significant.

Acknowledgements

This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) to L.E.K. V.T. acknowledges the support of the Human Frontiers Science Program. W.-Y.C. is a recipient of a CIHR Senior Research Fellowship. The authors thank Irina Bezsonova for the gift of a drKN SH3 sample. Useful discussions with Professor Ray Freeman are acknowledged. L.E.K. holds a Canada Research Chair in Biochemistry.

References

- Boyd, J. and Soffe, N. (1989) J. Magn. Reson., 85, 406-413.
- Caffrey, M., Kaufman, J., Stahl, S.J., Wingfield, P.T., Gronenborn, A.M. and Clore, G.M. (1998) J. Magn. Reson., 135, 368–372.
- Coggins, B.E., Venters, R.A. and Zhou, P. (2004) J. Am. Chem. Soc., 126, 1000–1001.
- Griesinger, C. and Ernst, R.R. (1987) J. Magn. Reson., 75, 261-271.
- Grzesiek, S. and Bax, A. (1993) J. Biomol. NMR, 3, 185–204.
- Gutmanas, A., Tu, L., Yu, V., Orekhov, V.Y. and Billeter, M. (2004) J. Magn. Reson. 167, 107–113.
- Howard, B.R., Endrizzi, J.A. and Remington, S.J. (2000) *Biochemistry*, **39**, 3156–3168.
- Kamith, U. and Shriver, J.W. (1989) J. Biol. Chem., 264, 5586– 5592.

- Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) J. Magn. Reson., 89, 496–514.
- Kay, L.E., Keifer, P. and Saarinen, T. (1992a) J. Am. Chem. Soc., 114, 10663–10665.
- Kay, L.E., Nicholson, L.K., Delaglio, F., Bax, A. and Torchia, D.A. (1992b) J. Magn. Reson., 97, 359–375.
- Kim, S. and Szyperski, T. (2003) J. Am. Chem. Soc., 125, 1385– 1393.
- Kim, S. and Szyperski, T. (2004) J. Biomol. NMR., 28, 117-130.
- Kupče, E. and Freeman, R. (2003a) *J. Biomol. NMR*, **27**, 101–113.
- Kupče, E. and Freeman, R. (2003b) *J. Am. Chem. Soc.*, **125**, 13958– 13959.
- Kupče, E. and Freeman, R. (2003c) J. Biomol. NMR, 27, 383-387.
- Kupče, E. and Freeman, R. (2004a) J. Am. Chem. Soc., 126, 6429– 6440.
- Kupče, E. and Freeman, R. (2004b) J. Biomol. NMR, 28, 391–395.
- Logan, T.M., Olejniczak, E.T., Xu, R.X. and Fesik, S.W. (1993) J. Biomol. NMR, 3, 225–231.
- McCoy, M.A. (1998) J. Magn. Reson., 130, 341-345.
- Palmer, A.G., Skelton, N.J., Chazin, W.J., Wright, P.E. and Rance, M. (1992) *Mol. Phys.*, **75**, 699–711.
- Patt, S.L. (1992) J. Magn. Reson., 96, 94-102.
- Peng, J.W. and Wagner, G. (1992) J. Magn. Reson., 98, 308-332.
- Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) Proc. Natl. Acad. Sci. USA, 94, 12366–12371.
- Schleucher, J., Sattler, M. and Griesinger, C. (1993) Angew. Chem. Int. Ed. Engl., 32, 1489–1491.
- Tugarinov, V. and Kay, L.E. (2003) J. Mol. Biol., 327, 1121-1133.
- Tugarinov, V., Muhandiram, R., Ayed, A. and Kay, L.E. (2002) J. Am. Chem. Soc., 124, 10025–10035.
- Vuister, G.W. and Bax, A. (1992) J Magn. Reson., 98, 428-435.
- Yamazaki, T., Muhandiram, R. and Kay, L.E. (1994) J. Am. Chem. Soc., 116, 8266–8278.
- Yang, D. and Kay, L.E. (1999) J. Biomol. NMR, 13, 3-10.
- Zhang, O. and Forman-Kay, J.D. (1995) *Biochemistry*, **34**, 6784–6794
- Zhang, O., Forman-Kay, J.D., Shortle, D. and Kay, L.E. (1997) J. Biomol. NMR, 9, 181–200.