Journal of Biomolecular NMR, **15:** 207–211, 1999.
KLUWER/ESCOM
© 1999 *Kluwer Academic Publishers. Printed in the Netherlands.*

Observation of internucleotide NH···N hydrogen bonds in the absence of directly detectable protons

Ananya Majumdar*, Abdelali Kettani, Eugene Skripkin & Dinshaw J. Patel
Cellular Biochemistry and Biophysics Program, Box 557, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, U.S.A.

Received 10 August 1999; Accepted 13 September 1999

Key words: A · (G · G · G · G) · A hexad, G · C · G · C tetrad, G · G · G · G tetrad, HNN-COSY, $^{2h}J_{NN}$

Abstract

Several structural motifs found in nucleic acids involve N-H · · · N hydrogen bonds in which the donor hydrogens are broadened to extinction due to chemical or conformational exchange. In such situations, it is impossible to use the well-established HNN-COSY or soft HNN-COSY experiments, which report the presence of the hydrogen bond directly on the donor proton(s). We present a pulse sequence, H(CN)N(H), for alleviating this problem in hydrogen bonds of the type $N_dH \cdots N_a$-CH, in which the donor N_d nitrogen is correlated with the corresponding *non-exchangeable* C-H proton associated with the acceptor N_a nitrogen. In this way, missing $N_dH \cdots N_a$ correlations in an HNN-COSY spectrum may be recovered from CH-N_d correlations in the H(CN)N(H) spectrum. By correlating a different set of nuclei relative to the HNN-COSY class of experiments, the H(CN)N(H) experiment also serves to remove ambiguities associated with degeneracies in HNN-COSY spectra. The technique is demonstrated on $d(GGAGGAG)_4$,a quadruplex containing a novel A · (G · G · G · G) · A hexad and on $d(GGGCAGGT)_4$, containing a G · C · G · C tetrad, in which missing $NH_2 \cdots N7$ correlations are retrieved via H8-(N2,N6) correlations in the H(CN)N(H) spectrum.

Introduction

The recent discovery of internucleotide $^{2h}J_{NN}$ in nucleic acids (Dingley and Grzesiek, 1998; Pervushin et al., 1998) and inter-residue $^{3h}J_{NC}$ in proteins (Cordier and Grzesiek, 1999; Cornilescu et al., 1999a), yielding direct evidence for hydrogen bonds, has added a new dimension to biomolecular structure determination by NMR (Cornilescu et al., 1999b; Dingley et al., 1999; Kettani et al., 1999a, b; Wang et al., 1999). In nucleic acids, and their complexes with peptides and proteins, a variety of structural motifs have been observed, consisting of extensively hydrogen bonded networks. These include tetrads, triads (Patel et al., 1999) and arginine fork alignments (Patel, 1999). New motifs, characteristic of DNA and RNA polymorphism, are increasingly being discovered. Elucidating the hydrogen bonding pattern in these molecules enables their structures to be determined far more accurately than previously used indirect methods such as NOEs, chemical shifts and exchange rates of hydrogen bonded protons.

Despite the tremendous success of established techniques such as HNN-COSY (Dingley and Grzesiek, 1998) and soft HNN-COSY (Majumdar et al., 1999) to obtain correlations between $^{2h}J_{NN}$ coupled ^{15}N nuclei, serious problems are posed in cases where the donor protons are engaged in intermediate chemical or conformational exchange. Most severely affected are the amino (NH_2) protons, which undergo intermediate rotational exchange about the exocyclic C-N bond. Even at temperatures as low as 0 °C, these protons are often broadened to extinction, thus precluding the observation of the appropriate $^{2h}J_{NN}$ correlation directly on the amino protons. This phenomenon has been commonly observed in a number of nucleic acid structures, especially with hydrogen bonds in-

*To whom correspondence should be addressed. E-mail: majumdar@sbnmr1.ski.mskcc.org

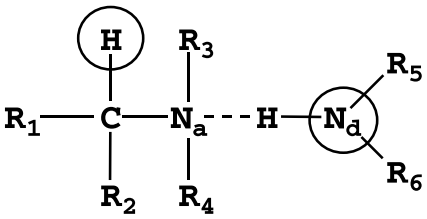

Figure 1. Schematic of the hydrogen bonded network to which the H(CN)N(H) pulse sequence of Figure 2 may be applied. Nuclei that are correlated by the sequence are enclosed in circles.

volving guanine amino protons. Inability to identify these hydrogen bonds invariably introduces ambiguity in structure determination. Fortunately, this limitation may be overcome fairly easily in the case of hydrogen bonds of the type $N_dH\cdots N_a$-CH, as shown in Figure 1. We present a pulse sequence called H(CN)N(H), which bypasses the exchangeable N_dH proton(s), and reports the $^{2h}J_{NN}$ coupling on the *non-labile* CH proton associated with the acceptor N_a nitrogen. The H(CN)N(H) experiment has the following benefits: (a) it allows observation of hydrogen bonded nuclei in the absence of observable hydrogen bonded protons; (b) it may be performed at higher temperatures with concomitant gains in sensitivity due to more favorable relaxation effects, and (c) since different pairs of nuclei are correlated relative to the HNN-COSY or soft HNN-COSY experiment, it may be used for resolving ambiguities due to degeneracies in HNN-COSY type of spectra. Specifically, in this work, the H(CN)N(H) sequence is used to correlate the H8 proton (ω_2) to the nitrogen (N2,N6) amino donor (ω_1) of guanine and adenine residues, in order to retrieve missing NH_2-N7 correlations in corresponding soft HNN-COSY spectra. The technique is demonstrated on $d(GGAGGAG)_4$, a quadruplex containing a novel $A\cdot(G\cdot G\cdot G\cdot G)\cdot A$ hexad (Kettani et al., 1999b) and on $d(GGGCAGGT)_4$, containing a $G\cdot C\cdot G\cdot C$ tetrad.

Methods

Figure 2 shows the pulse sequence used in the H(CN)N(H) experiment. Magnetization transfer follows an out-and-back pathway as outlined below:

$$-H_y \xrightarrow{J_{NH}} H_xN_{az} \xrightarrow{90^H_y,90^{Na}_x} H_zN_{ay} \xrightarrow{J_{NN}} -H_zN_{ax}N_{dz}$$

$$\xrightarrow{90^{Na}_y,90^{Nd}_x} -H_zN_{az}N_{dy}(t_1) \xrightarrow{90^{Na}_y,90^{Nd}_x} -H_zN_{ax}N_{dz} \xrightarrow{J_{NN}}$$

$$-H_zN_{ay} \xrightarrow{90^H_{-x},90^{Na}_x} -H_yN_{az} \xrightarrow{J_{NH}} H_x(t_2)$$

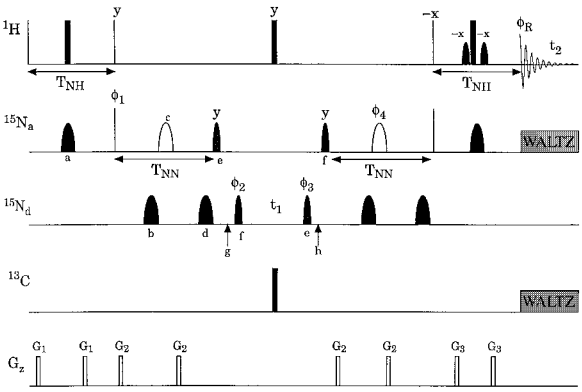

Figure 2. Pulse sequence for the H(CN)N(H) experiment. Narrow and wide pulses (both square and shaped) represent 90° and 180° pulses, respectively, with phase *x*, unless indicated otherwise. The ^{1}H carrier was placed at 4.8 ppm (water). The ^{13}C carrier was initially placed at 155 ppm and shifted to 140 ppm prior to detection. The ^{15}N carrier was placed in the center of the N7 region (235 ppm) prior to the ^{15}N evolution period until point g (indicated by the first arrow), when it was switched to the center of the N2,N6 region (80 ppm). The carrier was switched back again to 235 ppm at point h (indicated by the second arrow), after the ^{15}N evolution period. Rf field strengths employed were: ^{1}H: hard pulses: 40 kHz, qsneeze pulses for WATERGATE (Piotto et al., 1992): 1.1 kHz (4.3 ms); ^{13}C: 18 kHz, WALTZ16 (Shaka et al., 1983) decoupling during acquisition (t_2): 3.8 kHz; ^{15}N: hard pulses: 6.3 kHz, WALTZ16 decoupling during acquisition (t_2): 1.2 kHz. ^{15}N selective pulses, as indicated by the letters (a): 0.9 ms (3.7 kHz peak rf) iburp1 (Geen and Freeman, 1991), (b): 0.9 ms iburp1, phase-modulated at −160 ppm (Boyd and Scoffe, 1989; Patt, 1992), (c): 0.62 ms (3.8 kHz peak rf) rsnob (Kupce et al., 1995), (d): same as (b), for Bloch–Siegert shift compensation (Vuister and Bax, 1992), (e): 1.2 ms (4.0 kHz peak rf) time-reversed qsneeze pulses (Kupce and Freeman, 1995) for $M_{xy}\rightarrow M_z$ rotations, (f): 1.2 ms qsneeze pulses for $M_z\rightarrow M_{xy}$ rotations. All shaped pulse durations were optimized for 600 MHz (^{1}H) frequency. Delays used were: T_{NH} = 28.0 ms, T_{NN} = 45.0 ms. Phase cycles: $\phi_1 = x, -x$, $\phi_2 = 2(x), 2(-x)$, $\phi_3 = 4(x), 4(-x)$, $\phi_4 = 8(x), 8(y)$, ϕ_R = ABBA, where A = x, −x, −x, x, B = −x, x, x, −x. Quadrature detection in the ω_1 dimension was achieved by States-TPPI phase cycling of ϕ_2 (Marion et al., 1989). Rectangular z-gradients G_1-G_3 (20 G/cm) were applied for 1.0 (G_1), 1.0 (G_2) and 0.4 (G_3) ms.

For the specific case of H = H8, N_a = N7 and N_d = (N2, N6), transfer from H8 to N7 via the long-range $^2J_{N7H8}$ coupling (∼10 Hz) turned out to be more efficient than the alternative H8-C8-N7 pathway. Since N7 resonances (∼240 ppm) appear far downfield from the N2 (G) or N6 (A) amino nitrogen (∼80–100 ppm), evolution under the $^{2h}J_{NN}$ coupling and subsequent N7-N2 coherence transfer was achieved in a pseudo-heteronuclear manner as described previously (Majumdar et al., 1999). Although non-exchangeable protons were being detected, we found water suppression to be most effective via a

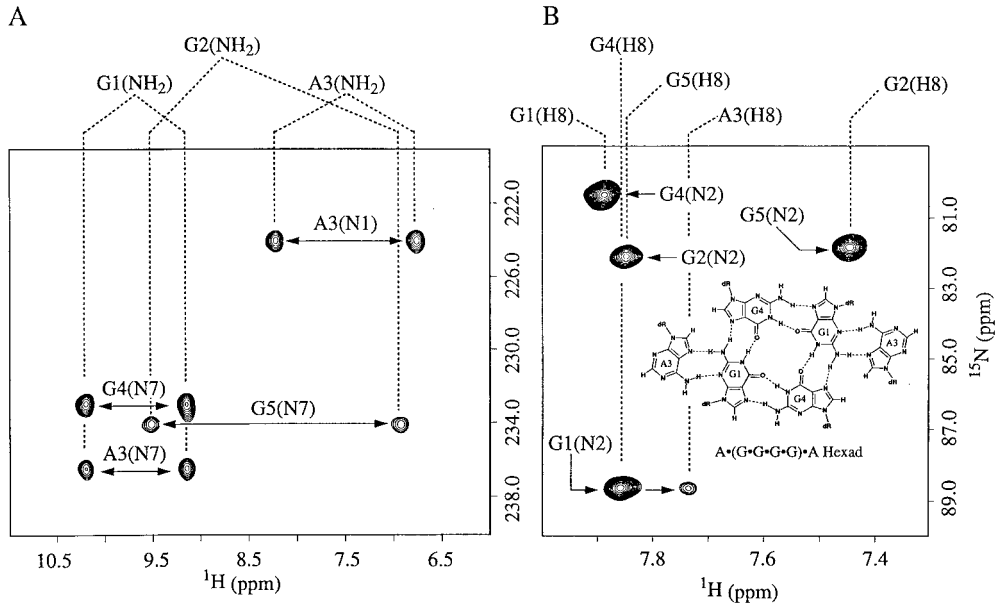

Figure 3. (A) Spectrum showing inter- and intranucleotide NH_2 (ω_2)-(N1,N7) (ω_1) cross peaks using the soft HNN-COSY experiment, recorded on a 1.5 mM sample of uniformly (^{15}N, ^{13}C) labeled $d(GGAGGAG)_4$ at 0 °C on a Varian Inova spectrometer operating at 600 MHz (^1H frequency). The spectrum consisted of 864 (t_2) and 40 (t_1) complex data points with 480 transients/FID. Spectral widths of 2000 Hz (t_1 max: 20.0 ms) and 12 000 Hz ($t_2$max: 72.0 ms) were employed. The final 2D matrix consisted of 1024 (ω_2) $*$ 128 (ω_1) points. (B) H(CN)N(H) spectrum recorded on the same sample at 20 °C showing the internucleotide H8 (ω_2)-N2(ω_1) cross peaks. The spectrum consisted of 736 (t_2) $*$ 40 (t_1) complex points, with 144 transients/FID. Spectral widths of 1000 Hz (t_1 max: 40 ms) and 8000 Hz ($t_2$max: 92 ms) were used, with a relaxation delay of 2.4 s, resulting in a total acquisition time of 8.5 h. The final 2D matrix consisted of 1024 (ω_2)$*$128 (ω_1) points. All spectra in Figures 3 and 4 were processed using Felix 97.0 (Biosym).

WATERGATE sequence (Piotto et al., 1992) following the final hard 90° proton pulse, which returns the bulk of the water magnetization to the +z axis. Despite the long duration of the pulse sequence (∼ 150 ms, minimum), the ability to perform the experiment at higher temperatures (20 °C in this case) resulted in adequate sensitivity on molecules of M_r 9–10 kDa at concentrations of 1.5 mM.

Results

Figures 3A and 3B show comparisons of soft HNN-COSY spectra correlating NH_2 protons (ω_2) with the N7 (ω_1) nitrogen (Figure 3A) and the H(CN)N(H) experiment correlating the H8 protons (ω_2) of adenine and guanine residues with the N2,N6 (ω_1) nitrogens, in the $d(GGAGGAG)_4$ quadruplex. This molecule has been demonstrated to adopt a unique structural motif consisting of a central A · (G · G · G · G) · A hexad involving G1, G4 and A3 residues (inset in Figure 3B). Hexad formation involves recognition of the exposed minor groove of opposing guanines within the G1 · G4 · G1 · G4 tetrad through sheared G1 · A3 formation. Symmetry related A · (G · G · G · G) · A hexads are stacked on G2 · G5 · G2 · G5 tetrads (Kettani et al., 1999b). The A3(NH_2)-G1(N3) correlation was obtained using the HNN-COSY experiment. The soft HNN-COSY spectrum (Figure 3A) recorded at 0 °C shows some of the cross peaks which identify the NH_2-N7 hydrogen bonds that are consistent with the model: G1(NH_2)-G4(N7) and G1(NH_2)-A3(N7) from the hexad and G2(NH_2)-G5(N7) from the G-tetrad. The G2(NH_2)-G5(N7) correlation was weak due to intermediate rotational broadening of the G2 amino protons and required CPMG-based modifications (Mueller et al., 1995) of the experiment in order to be observable. However, the amino protons of G4 and G5 were broadened beyond limits of detection, resulting in the G4(NH_2)-G1(N7) and G5(NH_2)-G2(N7) cross peaks being totally absent from the soft HNN-COSY spectrum. As shown in Figure 3B, these cross peaks may be completely 'recovered' in the H(CN)N(H) spectrum, recorded at 20 °C. For example, the G1(NH_2)-G4(N7) correlation in the soft-HNN COSY spectrum appears as a G4(H8)-G1(N2) correlation in the H(CN)N(H) spectrum. Similarly, the miss-

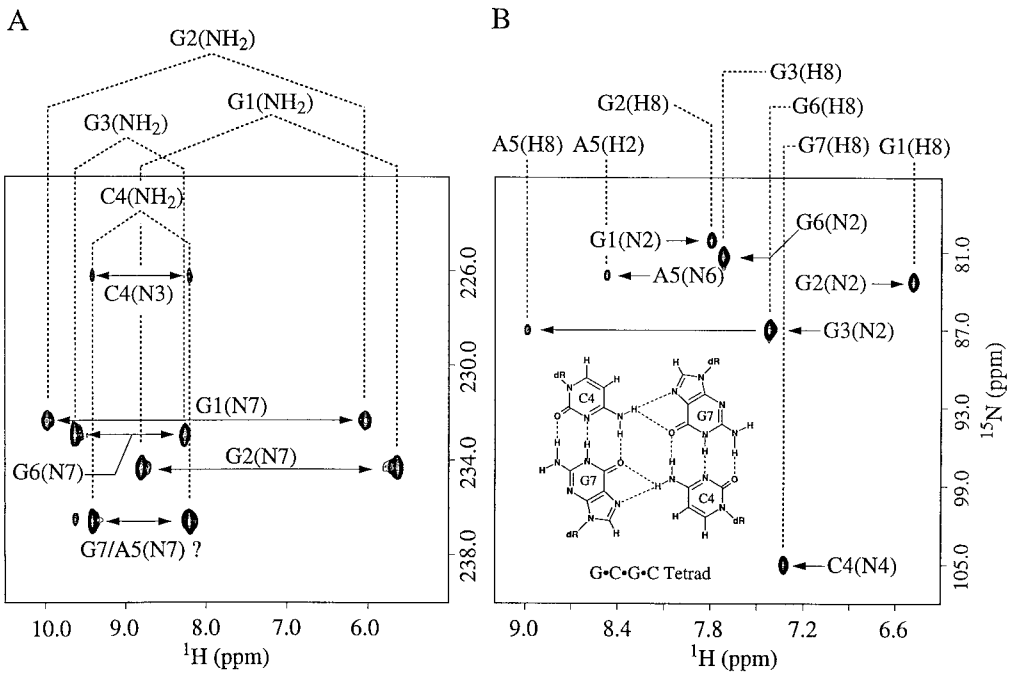

Figure 4. (A) Spectrum showing inter- and intranucleotide NH_2 (ω_2)-(N1,N7) (ω_1) cross peaks using the soft HNN-COSY experiment, recorded on a 1.5 mM sample of uniformly (^{15}N, ^{13}C) labeled d(GGGCAGGT)$_4$ at 0 °C. The spectrum consisted of 896 (t_2) and 64 (t_1) complex data points with 144 transients/FID. Spectral widths of 1824 Hz ($t_1$max: 34.5 ms) and 12500 Hz ($t_2$max: 72 ms), and a relaxation delay of 1.5 s were employed. The final 2D matrix consisted of 1024 (ω_2)$*$ 128(ω_1) points. (B) H(CN)N(H) spectrum recorded on the same sample at 20 °C showing internucleotide H8 (ω_2)-N2 (ω_1) cross peaks and the intranucleotide H2 (ω_2)-N6 (ω_1) correlation for A5. The spectrum consisted of 608 (t_2)$*$ 60 (t_1) complex points, with 160 transients/FID. Spectral widths of 2000 Hz ($t_1$max: 30 ms) and 8000 Hz ($t_2$max: 76 ms) were used, with a relaxation delay of 2 s, resulting in a total acquisition time of 8 h. The final 2D matrix consisted of 1024 (ω_2) $*$ 128 (ω_1) points.

ing G4(NH_2)-G1(N7) and G5(NH_2)-G2(N7) cross peaks appear as G1(H8)-G4(N2) and G2(H8)-G5(N2) correlations, respectively. The N2 assignments were obtained using an H(N1)(C2)N2 correlation spectrum (available as Supplementary material). The G2 (NH_2)-G5(N7) cross peak, which is weak in the soft HNN-COSY spectrum due to the broad G2 (NH_2) proton, appears as a strong peak in the H(CN)N(H) spectrum. Interestingly, the G1(NH_2)-A3(N7) cross peak in the soft HNN-COSY spectrum is stronger than its A3(H8)-G1(N2) counterpart in the H(CN)N(H) spectrum, probably due to a smaller H8-N7 coupling in the adenine residue. The H(CN)N(H) and HNN-COSY experiments therefore complement each other effectively. In conjunction, the two techniques yielded all the N $\cdots$ N hydrogen bonds consistent with the NMR structure of this molecule.

As mentioned above, another aspect of the complementary nature of the two experiments is that they correlate two different pairs of nuclei. This feature allows the H(CN)N(H) experiment to resolve assignment-related ambiguities in the HNN-COSY experiment. Figures 4A and 4B compare the soft HNN-COSY (Figure 4A) and H(CN)N(H) spectra (Figure 4B) recorded on the G-quadruplex d(GGGCAGGT)$_4$ for which the exact fold is not known at present, but preliminary data establishes G $\cdot$ C $\cdot$ G $\cdot$ C tetrad formation involving dimerization through major groove edges of individual G7 $\cdot$ C4 Watson–Crick base pairs. The G7(N1)-C4(N3) pairing was obtained from HNN-COSY data. However, the C4(NH_2)-G7(N7) correlation from the soft HNN-COSY spectrum was ambiguous (Figure 4A), due to degeneracy of the N7 nitrogens of G7 and A5. This degeneracy is easily lifted in the H(CN)N(H) spectrum (Figure 4B) where the distinct chemical shift of C4(N6) clearly establishes the G7(H8)-C4(N6) correlation. In addition, the missing G6(NH_2)-G3(N7) correlation in the soft HNN-COSY spectrum is recovered via the G3(H8)-G6(N2) cross peak.

Discussion

We have presented a technique which effectively complements the soft HNN-COSY experiment by obtaining hydrogen bond information via non-exchangeable protons. The DNA systems on which the technique has been demonstrated are relatively large (9–10 kDa), by nucleic acid standards. Direct detection of hydrogen bonded protons in the HNN-COSY class of experiments requires low temperatures, usually around 0 °C. With increasing molecular weight, the increased line widths at such low temperatures will eventually become the limiting factor. The H(CN)N(H) experiment, which may be carried out at higher temperatures, is likely to serve as a more sensitive alternative. Although the focus in this work has been on NH_2-N7 hydrogen bonds in G · G, G · A and G · C base pairs, the technique is applicable to any $N_dH \cdots N_a$-CH type of linkage (Figure 1). For example, N(6)H2-N7 hydrogen bonds in A · A mismatches, which were previously studied using the soft HNN-COSY sequence (Majumdar et al., 1999), may also be observed using the H(CN)N(H) experiment. Similarly, NH-N1 linkages in A · U or A · T base pairs, or NH_2-N1 linkages in G · A mismatches (Jiang et al., 1996) may be detected using the non-labile H2 proton of adenine. The technique also provides a method for probing hydrogen bonds between basic side-chains and intracyclic nitrogens of nucleotides in nucleic acid-(peptide/protein) complexes. In addition to providing structural information, this technique is likely to be of significance in monitoring hydrogen bonds during folding/unfolding studies of nucleic acids and their complexes.

Acknowledgements

This research was supported by NIH grant no. GM-34504 to D.J.P.

References

Boyd, J. and Scoffe, N. (1989) *J. Magn. Reson.*, **85**, 406–413.
Cordier, F. and Grzesiek, S. (1999) *J. Am. Chem. Soc.*, **121**, 1601–1602.
Cornilescu, G., Hu, J.S. and Bax, A. (1999a) *J. Am. Chem. Soc.*, **121**, 2949–2950.
Cornilescu, G., Ramirez, B.E., Frank, M.K., Clore, G.M., Gronenborn, A.M. and Bax, A. (1999b) *J. Am. Chem. Soc.*, **121**, 6275–6279.
Dingley, A.J. and Grzesiek, S. (1998) *J. Am. Chem. Soc.*, **120**, 8293–8297.
Dingley, A.J., Masse, J.E., Peterson, R.D., Barfield, M., Feigon, J. and Grzesiek, S. (1999) *J. Am. Chem. Soc.*, **121**, 6019–6027.
Geen, H. and Freeman, R. (1991) *J. Magn. Reson.*, **93**, 93–141.
Jiang, F., Kumar, R.A., Jones, R.A. and Patel, D.J. (1996) *Nature*, **382**, 183–186.
Kettani, A., Bouaziz, S., Skripkin, E., Majumdar, A., Wang, W., Jones, R. and Patel, D.J. (1999a) *Structure*, **7**, 803–815.
Kettani, A., Gorin, A., Majumdar, A., Skripkin, E., Zhao, H., Jones, R. and Patel, D.J. (1999b) *Structure*, submitted.
Kupce, E., Boyd, J. and Campbell, I.D. (1995) *J. Magn. Reson.*, **B106**, 300–303.
Kupce, E. and Freeman, R. (1995) *J. Magn. Reson.*, **A112**, 134–137.
Majumdar, A., Kettani, A. and Skripkin, E. (1999) *J. Biomol. NMR*, **14**, 67–70.
Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.
Mueller, L., Legault, P. and Pardi, A. (1995) *J. Am. Chem. Soc.*, **117**, 11043–11048.
Patel, D.J., Bouaziz, S., Kettani, A. and Wang, Y. (1999) In Oxford Handbook of Nucleic Acid Structures (Ed., Neidle, S.), Oxford University Press, Oxford, pp. 389–453.
Patel, D.J. (1999) *Curr. Opin. Struct. Biol.*, **9**, 74–87.
Patt, S.L. (1992) *J. Magn. Reson.*, **96**, 94–102.
Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M. and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 14147–14151.
Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
Shaka, A.J., Keeler, J., Frenkiel, T. and Freeman, R. (1983) *J. Magn. Reson.*, **52**, 335–338.
Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.
Wang, Y.X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S.J., Huang, S.L., Torchia, D., Grzesiek, S. and Bax, A. (1999) *J. Biomol. NMR*, **14**, 181–184.