Journal of Biomolecular NMR, **15:** 173–176, 1999.
KLUWER/ESCOM

Assignment of cytosine N3 resonances in nucleic acids via intrabase three-bond coupling to amino protons

Simon Rüdisser, Jeffrey G. Pelton* & Ignacio Tinoco Jr.
Department of Chemistry, University of California and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720-1460, U.S.A.

Received 6 July 1999; Accepted 11 August 1999

Key words: cytosine, group I intron, ^{1}H-^{15}N HSQC, ^{15}N NMR, RNA

Abstract

Coherences were observed between ^{15}N3 of cytosine and its *trans* amino proton (H42) using a modified gradient-based heteronuclear single quantum coherence (HSQC) pulse sequence optimized for three-bond proton-nitrogen couplings. The method is demonstrated with a 22-nucleotide RNA fragment of the P5abc region of a group I intron uniformly labeled with ^{15}N. Use of intraresidue ^{15}N3-amino proton couplings to assign cytosine ^{15}N3 signals complements the recently proposed J_{NN} HNN COSY [Dingley, A.J. and Grzesiek, S. (1998) *J. Am. Chem. Soc.*, **120**, 8293–8297] method of identifying hydrogen-bonded base pairs in RNA.

Abbreviations: HSQC, heteronuclear single quantum coherence; INEPT, insensitive nuclei enhanced by polarization transfer.

Recently, direct observation of J_{NN} couplings through hydrogen bonds has been reported for RNA (Dingley and Grzesiek, 1998) and DNA (Pervushin et al., 1998). Correlation of hydrogen-bonded nitrogens gives important constraints for structure calculations, but poses the problem of assigning the nitrogens involved in hydrogen bond formation. For G·C base pairs (Figure 1A) the guanine imino nitrogen can readily be assigned from one-bond ^{1}H-^{15}N correlation experiments once the imino proton has been assigned. However, to our knowledge, a method for assigning the cytosine ^{15}N3 resonances in nucleic acids has not been reported. Here we present a new and easy method for the unambiguous assignment of cytosine ^{15}N3 resonances by correlating them to intrabase non-hydrogen-bonded (*trans*) amino protons via a three-bond $^{3}J_{NH}$ coupling. The correlations are observed for a 22-nucleotide RNA molecule using a modified ^{1}H-^{15}N HSQC pulse sequence (Mori et al., 1995). Other nitrogens have been correlated with non-exchangeable proton resonances in RNA (Sklenár et al., 1994).

*To whom correspondence should be addressed. E-mail: jgpelton@lbl.gov

Assignment of the cytosine ^{15}N3 signals relies on the previous assignment of the cytosine amino proton resonances. These signals can usually be assigned via NOEs to H5 protons or to the imino proton of the hydrogen-bonded guanosine (Mueller et al., 1995; Varani et al., 1996). Alternatively, the amino protons can be assigned from triple-resonance HNCCCH (Simorre et al., 1995) or HCCNH-TOCSY (Fiala et al., 1996; Sklenár et al., 1996) experiments which correlate the amino H4 and base H6 resonances. The relative merits of these sequences are discussed in the review by Wijmenga and Van Buuren (1998).

All experiments were done on a Bruker AMX 600 or DRX 500 spectrometer at 30 °C. The ^{15}N-labeled 22-nucleotide RNA (Figure 1B) was enzymatically synthesized from a synthetic DNA template using T7 RNA polymerase (Wyatt, 1991) with ^{15}N-labeled NTPs. The RNA concentration was 2.2 mM in 10 mM sodium phosphate (pH 5.6), 200 mM NaCl, 0.1 mM EDTA in 90% H_2O/10% D_2O. NMR data were processed with Felix 95.0 (MSI Inc.). Proton chemical shifts were indirectly referenced to DSS using the water resonance (4.72 ppm). Nitrogen chem-

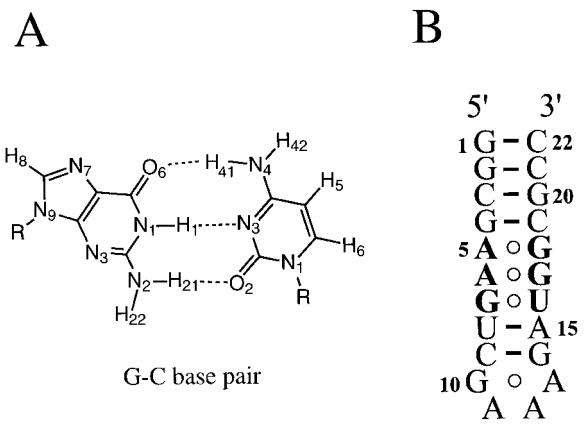

Figure 1. (A) A G·C base pair with the atom numbering as recommended by Markley et al. (1998). Open circles indicate base-pair mismatches. (B) The 22-nucleotide RNA (P5ga) which is part of the P5abc region of the group I intron from *Tetrahymena thermophila.* Almost complete ^{1}H and ^{15}N assignments have been obtained (unpublished results).

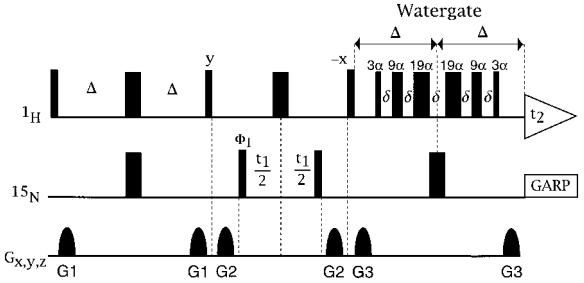

Figure 2. Fast HSQC (FHSQC) pulse sequence with a composite WATERGATE 3-9-19 pulse train (Mori et al., 1995). Unless indicated, all pulses are applied along the x axis. ϕ_1 and the receiver phase are cycled (x, −x). Narrow and wide pulses denote 90° and 180° pulses, respectively. Gradients G_1(x,y,z), G_2(x,y,z) and G_3(x,y,z) were set to 5, 3.6 and 7.1 G/cm, respectively. The WATERGATE delay δ was set to 312.4 μs (500 MHz), which corresponds to an excitation maximum at 8.00 ppm ($6\alpha = 180°$). To obtain long-range ^{1}H-^{15}N correlations, the delay in which ^{1}H magnetization becomes antiphase with respect to nitrogen (2Δ) was set to 42 ms. Quadrature detection in t_1 is achieved by States-TPPI modulation of the first nitrogen 90° pulse (ϕ_1) and the receiver (Marion et al., 1989).

ical shifts were indirectly referenced to ammonia (Wishart et al., 1995).

The 22-nucleotide RNA molecule (P5ga) used in our studies (Figure 1B) is part of the P5abc region of the group I intron from *Tetrahymena thermophila.* It contains two G·A base mismatches that bind Mg^{2+} when the sequence is part of the larger P4-P6 domain in a crystal (Cate et al., 1996). Our NMR structure shows that these G·A mismatches form imino hydrogen-bonded base pairs (unpublished data). In imino G·A mismatches the guanine H1 is hydrogen-bonded to adenine N1 and one adenine amino proton is hydrogen-bonded to guanine O6 (Burkard et al., 1999). This hydrogen bonding pattern is different from that observed in the crystal structure of the P4-P6 domain (Cate et al., 1996).

The fast HSQC (FHSQC) as reported by Mori et al. (1995) is useful for the study of exchangeable protons and scalar-coupled nitrogens in aqueous solutions with high sensitivity. An important advantage of the FHSQC experiment over other gradient-based HSQC methods is that water magnetization is returned to the +z axis at the start of acquisition without the need for water-selective flip-back pulses (Grzesiek and Bax, 1993). Remaining water magnetization is suppressed by a 3-9-19 WATERGATE (Sklenár et al., 1993) sequence during the ^{15}N-^{1}H refocusing delay. Detailed descriptions of the HSQC (Bax et al., 1990) and FHSQC (Mori et al., 1995) experiments have been reported. The pulse sequence used for observation of H4/^{15}N3 correlations is shown in Figure 2. The sequence is the same as for the FHSQC, except that the delay during which protons become antiphase with respect to the ^{15}N3 nitrogens (2Δ; INEPT delay; Figure 2) was modified. For our 22-nucleotide RNA fragment, maximum signal due to the long-range proton-nitrogen couplings was achieved with the INEPT delay set to 42 ms. Correlations between H4 protons and their directly attached nitrogens (^{15}N4) were also observed; however, these signals were strongly attenuated due to refocusing of the single-bond proton-nitrogen coupling ($^{1}J_{NH} = 91$ Hz (Roberts et al., 1981)) and due to proton-proton coupling between H41 and H42. For larger RNA molecules, maximum signal may be achieved with a shorter INEPT delay, due to increased relaxation during this period. Alternatively, the two-bond ^{15}N4-^{15}N3 coupling ($^{2}J_{NN} = 5.8$ Hz (Rüterjans et al., 1978)) might be used to correlate ^{15}N4 and ^{15}N3 signals (Wijmenga and Van Buuren et al., 1998; Majumdar et al., 1999). The long-range FHSQC experiment is expected to be somewhat less sensitive than the long-range HSQC (Sklenár et al., 1994), which was used to obtain nitrogen to non-exchangeable proton correlations in RNA, because of the relatively large amino proton linewidths.

Figures 3B and 3D show portions of a long-range FHSQC spectrum of the 22-nucleotide RNA acquired in 90% H_2O/10% D_2O at 30 °C. Cross peaks between cytosine ^{15}N4 nitrogens and their directly attached protons are shown in Figure 3B. In G·C base pairs (Figure 1A) one cytosine amino proton, H41, is hydrogen-bonded to guanine O6 and typi-

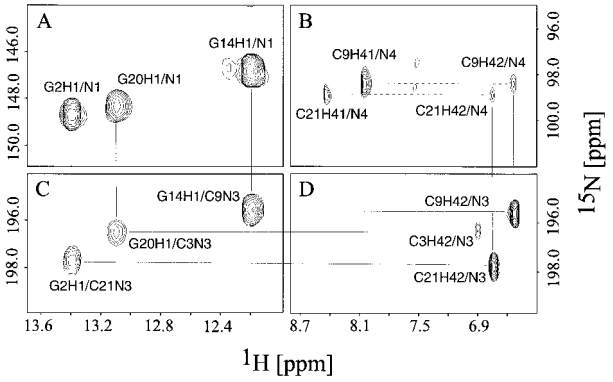

Figure 3. Expansion of a J_{NN} HNN COSY spectrum of P5ga. Magnetization is transferred from the guanine imino proton to the directly attached nitrogen ^{15}N1 (A) and from ^{15}N1 to the hydrogen-bonded cytosine ^{15}N3 (C). (B and D) Expansion of a long-range FHSQC spectrum of P5ga. Cross peaks between cytosine ^{15}N4 and the hydrogen-bonded H41 and non-hydrogen-bonded H42 protons are shown (B). Correlations between cytosine ^{15}N3 at about 197 ppm and H42 due to the three-bond coupling between the *trans* amino proton and ^{15}N3 (D). This allows assignment of the ^{15}N3 resonance involved in hydrogen bonding in G·C base pairs, provided the amino protons have been assigned. No cross peaks between ^{15}N3 and the *cis* H41 proton were observed. The J_{NN} HNN COSY (Dingley and Grzesiek, 1998) pulse sequence was slightly modified by replacing the original WATERGATE pulse train with a 1-1 jump and return water suppression sequence. The proton sweep width was set to 12195.1 Hz (600 MHz) and the nitrogen sweep width was set to 7905.1 Hz. The proton carrier was centered at H_2O and the nitrogen carrier was centered at 185 ppm. The J_{NN} HNN COSY experiment was recorded with a 4.5 ms INEPT transfer time and 30 ms J_{NN} magnetization transfer time. The excitation maximum was set to 12.7 ppm. A total of 208 scans were signal averaged for each of 97 complex t_1 (^{15}N) points. A total of 512 complex points were acquired in the t_2(^{1}H) dimension. A long-range FHSQC (B and D) was acquired at 500 MHz with a proton sweep width of 10000.0 Hz and a nitrogen sweep width of 5473.4 Hz. The ^{1}H carrier was set to H_2O and the ^{15}N carrier was set to 146 ppm. The ^{15}N rf field strength was 6.8 kHz. Nitrogen decoupling was achieved by GARP (Shaka et al., 1985) modulation of a 3.8 kHz rf field. A total of 48 scans were signal averaged for each of 112 complex t_1 (^{15}N) points. A total of 512 complex points were acquired in the t_2 (^{1}H) dimension. The INEPT delay (2Δ) was set to 42 ms.

cally resonates at lower field than the non-hydrogen-bonded amino proton, H42. Hydrogen-bonded and non-hydrogen-bonded protons are in slow exchange due to slow kinetics of rotation about the carbon-nitrogen bond and therefore two resonances, one for each type of proton, can usually be observed (McConnel and Seawell, 1973; McConnell, 1984). The guanine and adenine amino protons in nucleic acids are usually highly exchange-broadened by internal rotation about the carbon-nitrogen bond and often show only a single resonance (McConnell, 1984; Mueller et al., 1995).

The non-hydrogen-bonded cytosine amino protons show cross peaks to nitrogens resonating at about 197 ppm as shown in Figure 3D. From chemical shift considerations these nitrogens can only be ^{15}N3 resonances of the corresponding cytosine. Adenine ^{15}N1 nitrogens resonate near 224 ppm and guanine ^{15}N1 nitrogens resonate near 137 ppm (Markowski et al., 1977). We see cytosine amino H42 to ^{15}N3 correlations, but not cytosine amino H41 to ^{15}N3 cross peaks (Figure 3D). According to the Karplus equation, the three-bond coupling from ^{15}N3 to the *trans* H42 proton should be larger than the coupling from ^{15}N3 to the *cis* H41 proton (Roberts et al., 1981). This was observed previously in 5-azacytidine. The $^{3}J_{N3H42}$ coupling constant was measured to be 5 Hz, whereas $^{3}J_{N3H41}$ was too small to be observed (Roberts et al., 1981). As expected, we therefore do not observe cross peaks between H41 and ^{15}N3 in Figure 3D.

To test whether the correlations in Figure 3D are actually due to exchangeable amino protons we recorded an FHSQC spectrum in 100% D_2O and could not observe the correlations shown in Figure 3D. In addition we acquired a long-range FHSQC spectrum of the 22-nucleotide RNA in 90% H_2O/10% D_2O without decoupling in t_2 and observed a splitting due to $^{1}J_{NH}$ of the H42-^{15}N3 cross peaks in the proton dimension. This splitting corresponds to coupling of H42 with its directly attached nitrogen (^{15}N4) and further confirms our assignments. We also observed H8-^{15}N7/^{15}N9 and H1′-^{15}N9 correlations for purines, H6-^{15}N1 and H1′-^{15}N1 correlations for pyrimidines, and H2-^{15}N1/^{15}N3 correlations for adenine in the long-range FHSQC. These types of cross peaks have been observed previously in long-range HSQC spectra (Sklenár et al., 1994).

We were prompted to look for methods of assigning ^{15}N3 cytosine signals in order to confirm base pairing patterns suggested by guanine ^{15}N1 to cyto-

sine ^{15}N3 correlations observed in J_{NN} HNN COSY experiments. The J_{NN} HNN COSY experiment (Dingley and Grzesiek, 1998) correlates the guanine or adenine ^{15}N1 nitrogens with its hydrogen-bonded cytosine or uracil ^{15}N3 via a J_{N1N3} scalar coupling. Intra-residue GH1/^{15}N1 signals and the corresponding inter-residue GH1/C^{15}N3 obtained from the J_{NN} HNN COSY experiment applied to P5ga are shown in Figures 3A and C. Cytosine amino H4/^{15}N4 and H4/^{15}N3 correlations obtained from the long-range FHSQC experiment are shown in Figures 3B and D. As can be seen in Figures 3C and D, the ^{15}N3 resonance provides a linkage between the guanine imino (J_{NN} HNN COSY) and cytosine amino (long-range FHSQC) ^{1}H-^{15}N signals. Together, these data provide for the unambiguous assignment of hydrogen-bonded G·C base pairs, whenever both the imino and amino resonances are assigned.

With the present method we were able to assign ^{15}N3 signals for three of the five Watson–Crick G·C base pairs (C3, C9, C21) for the 22-nucleotide RNA (Figure 1B). The closing G·C base pair (G1·C22) and the G·C base pair next to the G·A mismatches (G4·C19) are in intermediate exchange, i.e. conformational dynamics and/or exchange with the solvent which leads to broad imino and amino proton resonances. Corresponding peaks can therefore not be observed either in the long-range FHSQC experiment or in the J_{NN} HNN COSY experiment. The C3H42 to C3^{15}N3 cross peak (Figure 3D) is also weak, presumably due to dynamics.

The long-range FHSQC method for the assignment of cytosine ^{15}N3 resonances should also be applicable to uniformly ^{15}N-labeled DNA molecules, and should provide information that can be used in conjunction with the DNA $^{h}J_{NN}$ HNN TROSY experiment (Pervushin et al., 1998) for the identification of hydrogen-bonded base pairs.

Acknowledgements

We thank Ms. Barbara Dengler for managing the laboratory, and Mr. David Koh for synthesizing DNA templates. This research was supported in part by National Institute of Health Grant GM 10840, by the Department of Energy Grant DE-FG03-86ER60406, and through instrumentation grants from the Department of Energy (DE-FG05-86ER75281) and from the National Science Foundation (DMB 86-09305). S.R. is grateful for a postdoctoral fellowship from the 'Fonds zur Förderung der wissenschaftlichen Forschung' of Austria, project Nr. J1499-PHY.

References

Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn. Reson.*, **86**, 304–318.

Burkard, M.E., Turner, D.H. and Tinoco, I. Jr. (1999) In *The RNA World* (Eds, Gesteland, R.F., Cech, T.R. and Atkins, J.F.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 677.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) *Science*, **273**, 1678–1684.

Dingley, A.J. and Grzesiek, S. (1998) *J. Am. Chem. Soc.*, **120**, 8293–8297.

Fiala, R., Jiang, F. and Patel, D.J. (1996) *J. Am. Chem. Soc.*, **118**, 689–690.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Majumdar, A., Kettani, A. and Skripin, E. (1999) *J. Biomol. NMR*, **14**, 67–70.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *J. Mol. Biol.*, **280**, 933–952.

Markowski, V., Sullivan, G.R. and Roberts, J.D. (1977) *J. Am. Chem. Soc.*, **99**, 714–718.

McConnell, B. (1984) *J. Biomol. Struct. Dyn.*, **1**, 1407–1421.

McConnell, B. and Seawell, P.C. (1973) *Biochemistry*, **12**, 4426–4434.

Mori, S., Abeygunawardana, C., O'Neil Johnson, M. and van Zijl, P.C.M. (1995) *J. Magn. Reson.*, **B108**, 94–98.

Mueller, L., Legault, P. and Pardi, A. (1995) *J. Am. Chem. Soc.*, **117**, 11043–11048.

Pervushin, K., Ono, A., Fernández, C., Szyperski, T., Kainosho, M. and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 14147–14151.

Roberts, J.D., Sullivan, G.R., Pang, P.P. and Leonard, N.J. (1981) *J. Org. Chem.*, **46**, 1014–1016.

Rüterjans, H., Blomberg, F. and Büchner, P. (1978) In *Proc. Eur. Conf. NMR Macromol.* (Ed., Filippo, C.), Lerici, Rome, Italy, pp. 319–335.

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.

Simorre, J.-P., Zimmermann, G.R., Pardi, A., Farmer II, B.T. and Mueller, L. (1995) *J. Biomol. NMR*, **6**, 427–432.

Sklenár, V., Dieckmann, T., Butcher, S.E. and Feigon, J. (1996) *J. Biomol. NMR*, **7**, 83–87.

Sklenár, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1994) *J. Biomol. NMR*, **4**, 117–122.

Sklenár, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson.*, **A102**, 241–245.

Varani, G., Aboul-ela, F. and Allain, F.H.-T. (1996) *Prog. NMR Spectrosc.*, **29**, 51–127.

Wijmenga, S.S. and Van Buuren, B.N.M. (1998) *Prog. NMR Spectrosc.*, **32**, 237–387.

Wishart, D.S., Bigam, C.Y., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.

Wyatt, J.R., Chastain, M. and Puglisi, J.D. (1991) *BioTechniques*, **11**, 764–769.