Journal of Biomolecular NMR, **16:** 175–178, 2000.
KLUWER/ESCOM

Long-range imino proton-^{13}C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA

Anh Tuân Phan
Groupe de Biophysique de l'Ecole Polytechnique et de l'UMR 7643 du CNRS, F-91128 Palaiseau, France. E-mail: pat@pmc.polytechnique.fr

Received 27 September 1999; Accepted 6 December 1999

Key words: heteronuclear NMR, imino proton assignment, long-range J-coupling, through-bond correlation, unlabeled DNA, water suppression

Abstract

Thanks to rather large (5–9 Hz) long-range imino proton-^{13}C J-couplings, heteronuclear correlation experiments in H_2O provide unambiguous assignment of imino protons by intranucleotide through-bond connectivities to guanosine H8 and thymidine CH_3 protons, or sequence-specific assignment of non-exchangeable protons when the imino protons are identified independently. This method is demonstrated in the Dickerson dodecamer $[d(CGCGAATTCGCG)]_2$ and in a human telomeric fragment of 22 nucleotides.

Abbreviations: HMBC, heteronuclear multiple-bond correlation; HSQC, heteronuclear single quantum coherence; INEPT, insensitive nuclei enhanced by polarization transfer; JR, jump and return; JRHMBC, HMBC with JR solvent suppression; JRHSQC, HSQC with JR solvent suppression.

Imino protons are precious for NMR structural studies of nucleic acids. In the case of duplex structures, they can be assigned by the classical sequential NOE walk (Wüthrich, 1986; Wijmenga et al., 1993) between imino protons whose spectrum is often relatively well resolved. The correlation between imino and non-exchangeable protons then provides an assignment method for the latter, whose spectrum is usually less well resolved. This approach may be inadequate for non-canonical structures. In such cases, one would first assign the non-exchangeable protons by model-independent strategies based on through-bond correlation instead of NOEs. Next, the imino protons would be assigned by correlation with the non-exchangeable ones.

Methods for such correlations have been developed for ^{13}C-/^{15}N-labeled RNA involving up to four J-couplings between heteronuclei (Simorre et al., 1995,1996; Fiala et al., 1996; Sklenár et al., 1996; Wöhnert et al., 1999). However, isotopic labeling is not always possible or desirable, and for DNA it is still expensive. Fortunately, it is sometimes unnecessary if one can find a coherence transfer pathway involving only one heteronucleus. Thus, correlation of an imino proton to its geminal ^{15}N without isotopic enrichment can be used to distinguish imino protons of guanosine and uridine (Szewczak et al., 1993). Also, long-range correlation of adenosine H2 and H8 protons has been achieved with good sensitivity by natural abundance HMBC in D_2O solvent (van Dongen et al., 1996; Phan et al., 2000). A correlation of imino and non-exchangeable protons at natural abundance requires a large J-coupling of the imino proton to a distant heteronucleus. In this paper, we demonstrate such J-couplings to ^{13}C5 and present an application of long-range heteronuclear correlation experiments in H_2O to the correlation of imino and non-exchangeable base protons in unlabeled DNA. The measurement requires solvent suppression methods which avoid saturation of the water magnetization and which operate rapidly, for sensitive detection of quickly exchanging protons.

The sequence of Figure 1A, which will be designated JRHMBC, combines the refocused HMBC sequence (Bax and Summers, 1986; Zhu and Bax,

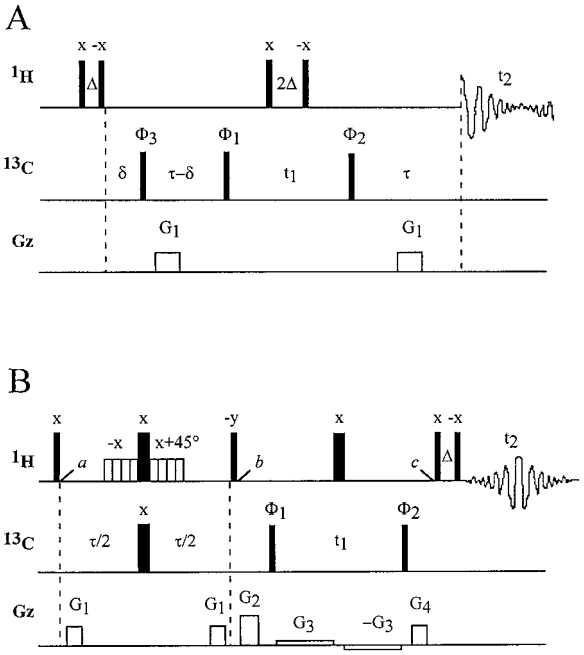

Figure 1. Pulse sequences for long-range heteronuclear correlation in H_2O. First and second lines: RF events; third line: z-axis pulsed-field gradient events. RF filled rectangles represent 180° (wide) and 90° (narrow) hard pulses; sets of open rectangles symbolize a 180° water-selective DANTE sequence (Morris and Freeman, 1978). The pulse phases are indicated. The phase cycle is: Φ_1 = x,−x; Φ_2 = 2(x),2(−x); Φ_3 = 4(x),4(−x); receiver = x,−x,−x,x. In addition, Φ_1 is phase-cycled in hypercomplex mode (States et al., 1982). (A) JRHMBC: δ = 2.5 ms; G_1 = (1 ms, 9 G/cm). (B) JRHSQC: G_1 = (0.5 ms, 9 G/cm); G_2 = (0.5 ms, 20 G/cm); G_3 = (t_1/2, 0.02 G/cm); G_4 = (0.5 ms, 9 G/cm).

1993) with JR water signal suppression (Plateau and Guéron, 1982) into a JR-echo method (Roy et al., 1984) supplemented by the use of gradients (Sklenár and Bax, 1987; Szewczak et al., 1993). The principal advantages of JR water suppression are simplicity (two hard pulses), the absence of phase-shift (and therefore baseline roll), and the short duration of the sequence, which makes it resistant to radiation damping, and good for detection of the fast-exchanging imino protons. The frequency of JR maximum sensitivity is set for optimal detection of imino protons. The H8 and CH_3 signals are not limiting, despite the reduction in signal by a factor of three to seven due to the $\sin^3$ profile of the JR-echo sequence (Figures 2, 3). One can also consider other water suppression procedures, using for instance water selective pulses which take place during the transfer delays (Gruschus and Ferretti, 1998).

The JRHMBC sequence may be compared to the sequence of Szewczak et al. (1993), from which it differs by the addition of the first ^{13}C pulse. This pulse, applied at time $\delta \sim 1/(2J_{one-bond})$, suppresses one-bond correlations in the 2D spectrum, with only slight perturbation of the long-range (small $J_{long-range}$) connectivities. This has proved useful in cases where one-bond correlation cross peaks overlapped with the cross peaks of interest due to spectral folding.

For long-range transfer, the standard delay $1/(2J_{long-range})$ is quite long and would lead to signal loss by transverse relaxation and by exchange with water in the case of exchangeable protons. The optimum delay τ is therefore shorter. The in-phase signal in JRHMBC (the anti-phase signal is not given for simplicity) is proportional to a factor given by:

$$A = \sin^3(\omega\Delta).\cos(\pi J_{long-range}\delta) . \sin(\pi J_{long-range}(\tau-\delta)).\sin(\pi J_{long-range}\tau) . \exp(-2\tau/T_2^{\#}) \quad (1)$$

where ω is the difference between the frequency of the observed protons and the carrier frequency which is that of the water protons, Δ is the JR delay. $T_2^{\#}$ is given by:

$$(T_2^{\#})^{-1} = (T_2)^{-1} + (\tau_{ex})^{-1} \quad (2)$$

where T_2 is the transverse relaxation time and τ_{ex} is the proton exchange time with water. The optimum delay τ for the in-phase signal is the value which maximizes the factor A in Equation 1 (van Dongen et al., 1996). Since it is shorter than $1/(2J_{long-range})$, ^{13}C-^{1}H rephasing still occurs during the acquisition time, unless ^{13}C is decoupled. If so and if homonuclear J-couplings are negligible, as in the case here, a pure absorption spectrum can be obtained.

Figure 2B shows the JRHMBC spectrum of [d(CGCGAATTCGCG)]$_2$. The imino protons are correlated to guanosine H8 and thymidine CH_3 protons via ^{13}C5 (Figure 2A). The J-couplings of H8 and CH_3 to ^{13}C5 have been reported earlier (Schmieder et al., 1992). The present measurements show that the imino proton-^{13}C5 J-coupling is in the range of 5 to 9 Hz. More precise measurements are in progress.

The sequence d(CCCTAA5mCCCTAACCCUAACCCT), modified from a fragment of the cytidine-rich strand of the human telomere, folds into an intramolecular i-motif (Phan et al., 2000). The chemical shifts of the uridine and thymidine imino protons (11.3 to 10.5 ppm) indicate H-bonding to oxygen of the oligonucleotide or of water. The imino proton of T4 was identified by the NOE with its own methyl protons,

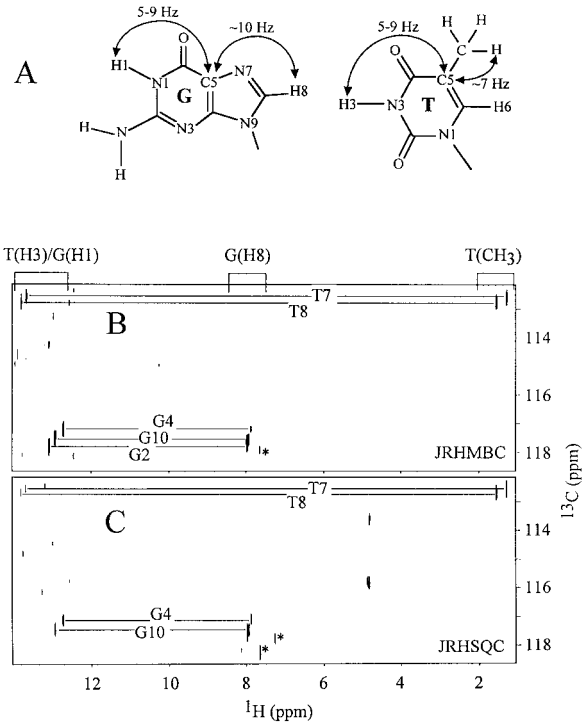

Figure 2. (A) The bases of guanosine and thymidine. Arrows indicate the coherence transfer pathways used in the experiments. (B–C) Spectra recorded by the sequences in Figure 1 for the DNA duplex [d(CGCGAATTCGCG)]$_2$ (strand concentration 1.9 mM; pH 7; temperature 25 °C; 90% H_2O, 10% D_2O) on a 500 MHz Varian Unity Inova spectrometer. The spectra were recorded with spectral widths of 12 kHz and 2 kHz, centered at 4.8 ppm and 113.3 ppm, respectively in the ^{1}H and ^{13}C dimensions, 36 complex t_1 points, 1024 complex t_2 points, repetition delay of 1 s. For the JRHMBC spectrum (B), 1024 scans per FID, $\tau = 30$ ms, total measurement time of 24 h. For the JRHSQC spectrum (C), 2048 scans per FID, $\tau = 50$ ms, total measurement time of 48 h. The spectrum is anti-phase in the proton dimension. Only positive contours are shown. In both cases, the delay Δ of the JR sequence was set for maximum sensitivity at 13.5 ppm, the residual water signal was reduced further by a digital shift correction (Roth et al., 1980). The cross peaks of long-range correlation within adenosines (^{1}H2-^{13}C4), which are detected due to spectral folding, are labeled with an asterisk.

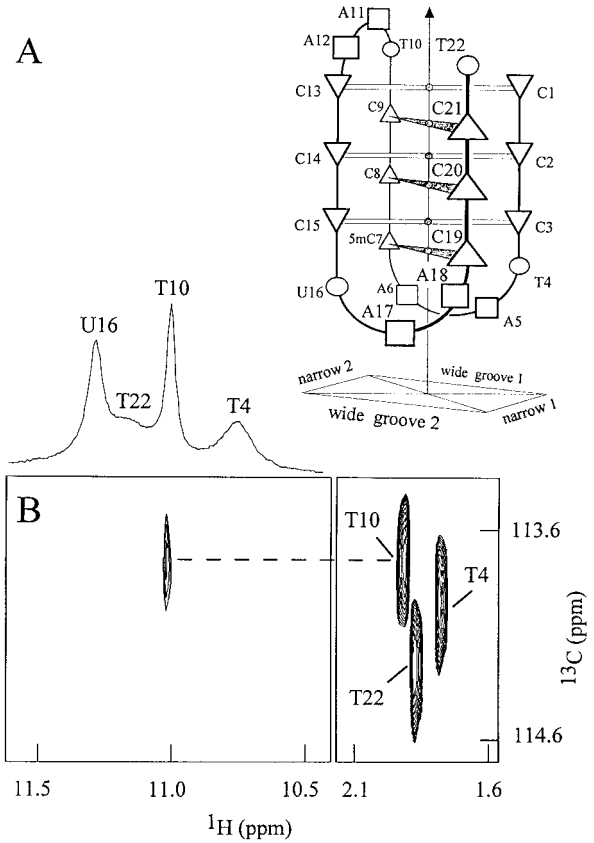

Figure 3. (A) Schematic structure and (B) JRHMBC spectrum of d(CCCTAA5mCCCTAACCCUAACCCT) (strand concentration 2.4 mM; pH 5; temperature 12 °C; 90% H_2O, 10% D_2O). The maximum sensitivity of the JR sequence was set at 11 ppm, the residual water signal was reduced further by a digital shift correction. Experimental conditions are the same as for JRHMBC in Figure 2, except for: $\tau = 17$ ms; 8192 scans per FID; 21 complex t_1 points; total measurement time 109 h.

and that of U16 by the absence of any such NOE. But the imino protons of T22 and T10 could not be thus assigned because both are connected to the two methyl groups by NOE cross peaks of comparable intensities (scheme in Figure 3A). We therefore resorted to JRHMBC for the assignment of these imino protons (Figure 3B). The T10 imino proton is assigned via the correlation to its identified CH_3, obtained even though this proton exchanges with water as fast as in isolated thymidine (∼50 ms at 12 °C). The transfer delay τ was 17 ms, much shorter than $1/(2J_{\mathrm{long-range}})$ which is ∼70 ms for methyl proton-^{13}C5 and imino proton-^{13}C5, as explained above. The total measurement time was 109 h, although it may be noted that the spectrum could have been assigned with only 5 complex t_1 points (not shown), i.e. a total measurement time of 26 h. Due to exchange broadening, the imino proton-^{13}C5 cross peaks are not seen for the other thymidines and for uridine.

Another sequence for long-range correlation is JRHSQC (Figure 1B). It combines a modified non-refocused HSQC (Sklenár et al., 1994) which is twice as short as a standard HSQC (Bodenhausen and Ruben, 1980), with JR water suppression. The water flip-back (Grzesiek and Bax, 1993) takes place during the INEPT (Morris and Freeman, 1979) period (between points *a* and *b*), so as to minimize the duration

of the excitation sequence. Prior to the application of the water-selective pulses, the water magnetization is defocused by gradient G_1, so that radiation damping is ineffective. This enables a good inversion by the 180° water-selective pulses. The first of these compensates for the subsequent 180° hard pulse. The second one, phase-shifted by 45° and in combination with the two G_1 gradients, rotates the water magnetization by 90° in the transverse plane, so that the next 90° hard pulse brings it along −Oz. This orientation is maintained against radiation damping with the help of gradients G_2 and G_3 (Otting, 1994), until the 180° decoupling pulse flips it back along +Oz. Finally, the signal is detected by a JR sequence, which is adjusted as described in Phan et al. (2000). The water flip-back and the suppression of radiation damping were evaluated by comparing the water proton magnetization detected at points *a*, *b* and *c*. Typically, more than 80% is found at point *b* as compared to *a*, and more than 95% is found at point *c* as compared to *b*.

The JRHSQC signal is proportional to a factor given by:

$$B = \sin(\omega\Delta).\sin(\pi J_{long-range}\tau).\exp(-\tau/T_2^{\#}) \quad (3)$$

Note that $B^2\sin(\omega\Delta)$ is equal to A if one ignores the short delay δ. Both factors are maximum for:

$$\tau = (1/\pi J_{long-range}).\arctan(\pi J_{long-range}T_2^{\#}) \quad (4)$$

The $\sin(\omega\Delta)$ excitation profile is preferred over the $\sin^3(\omega\Delta)$ of JRHMBC. On the other hand, a disadvantage of JRHSQC is that the signal is anti-phase in the proton dimension. This causes partial signal loss in case of poor spectral resolution.

Figure 2C shows the JRHSQC spectrum of $[d(CGCGAATTCGCG)]_2$. The narrow non-exchangeable protons are detected with good sensitivity. But the broad imino proton peaks are weaker than in JRHMBC (Figure 2B).

The technique, thus demonstrated for unlabeled DNA of moderate concentration (about 2 mM) and relatively rapidly exchanging protons, would also be useful for RNA and for other bases than guanosine and thymidine. In uridine and protonated cytidine, the H5 and imino protons can be correlated to $^{13}C5$ by one-bond and long-range experiments, respectively. Other long-range natural abundance heteronuclear correlations in nucleotides are under investigation.

Acknowledgements

The author thanks Maurice Guéron for encouragement during this work and for suggestions in the preparation of the manuscript, Jean-Louis Leroy and Frederic Dardel for helpful discussions, and the French Ministère des Affaires Etrangères for financial support. The 500 MHz Varian Unity Inova spectrometer was procured with the help of the Région Ile de France.

References

Bax, A. and Summers, M.F. (1986) *J. Am. Chem. Soc.*, **108**, 2093–2094.
Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.
Fiala, R., Jiang, F. and Patel, D.J. (1996) *J. Am. Chem. Soc.*, **118**, 689–690.
Gruschus, J.M. and Ferretti, J. (1998) *J. Magn. Reson.*, **135**, 87–92.
Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.
Morris, G.A. and Freeman, R. (1978) *J. Magn. Reson.*, **29**, 433–462.
Morris, G.A. and Freeman, R. (1979) *J. Am. Chem. Soc.*, **101**, 760–762.
Otting, G. (1994) *J. Magn. Reson.*, **B103**, 288–291.
Phan, A.T., Guéron, M. and Leroy, J.L. (2000) submitted.
Plateau, P. and Guéron, M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
Roth, K., Kimber, B.J. and Feeney, J. (1980) *J. Magn. Reson.*, **41**, 302–309.
Roy, S., Papastavros, M.Z., Sanchez, V. and Redfield, A.G. (1984) *Biochemistry*, **23**, 4395–4400.
Schmieder, P., Ippel, J.H., van den Elst, H., van der Marel, G.A., van Boom, J.H., Altona, C. and Kessler, H. (1992) *Nucleic Acids Res.*, **20**, 4747–4751.
Simorre, J.P., Zimmermann, G.R., Mueller, L. and Pardi, A. (1996) *J. Biomol. NMR*, **7**, 153–156.
Simorre, J.P., Zimmermann, G.R., Pardi, A., Farmer II, B.T. and Mueller, L. (1995) *J. Biomol. NMR*, **6**, 427–432.
Sklenár, V. and Bax, A. (1987) *J. Magn. Reson.*, **74**, 469–479.
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.
States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) *J. Magn. Reson.*, **48**, 286–292.
Szewczak, A.A., Kellogg, G.W. and Moore, P.B. (1993) *FEBS Lett.*, **327**, 261–264.
van Dongen, M.J.P., Wijmenga, S.S., Eritja, R., Azorin, F. and Hilbers, C.W. (1996) *J. Biomol. NMR*, **8**, 207–212.
Wijmenga, S.S., Mooren, M.M.W. and Hilbers, C.W. (1993) In *NMR of Macromolecules, a Practical Approach* (Ed., Roberts, G.C.K.), Oxford University Press, Oxford, pp. 217–288.
Wöhnert, J., Ramachandran, R., Görlach, M. and Brown, L.R. (1999) *J. Magn. Reson.*, **139**, 430–433.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.
Zhu, G. and Bax, A. (1993) *J. Magn. Reson.*, **A104**, 353–357.