Through-bond correlation of adenine H2 and H8 protons in unlabeled DNA fragments by HMBC spectroscopy

Maria J.P. van Dongen[a], Sybren S. Wijmenga[a,*], Ramón Eritja[b,**], Fernando Azorin[b]
and Cornelis W. Hilbers[a,***]

[a]NSR Centre for Molecular Structure, Design, and Synthesis, Laboratory of Biophysical Chemistry, University of Nijmegen,
Toernooiveld, 6525 ED Nijmegen, The Netherlands
[b]Centro de Investigación y Desarrollo-CSIC, Department of Molecular and Cellular Biology,
Jorge Girona Salgado 18-26, E-08034 Barcelona, Spain

Summary

A new application of the HMBC experiment is presented that provides a useful means to discriminate between H2 and H8 proton resonances, to assign the base proton resonances to the various residue types and, most importantly, to correlate the H2 and H8 protons for adenine or inosine residues in natural abundance ^{13}C fragments. The utility of this experiment is demonstrated for an unlabeled DNA 20-mer. Thanks to the obtained results, preliminary conclusions could be drawn regarding the molecular conformations of the non-canonical G/I–A base pairs in the hairpin formed by this fragment.

Traditionally, the sequential assignment of the proton resonances in NMR spectra of nucleic acid fragments is performed predominantly via characteristic NOE connectivities. For example, the non-exchangeable base H8/H6/H5 and sugar protons are usually assigned by sequential 5'-sugar-base-3' NOEs. The sequence-specific assignment of the H2 protons of adenine residues largely relies on the NOE contacts of these protons to imino protons (Wüthrich, 1986; Wijmenga et al., 1993). This approach becomes inadequate, however, if the strand adopts a nonhelical structure or if the bases do not form regular Watson–Crick pairs. In those cases, alternative assignment strategies, which preferably utilize through-bond correlations instead of NOEs, are required, so that the spectral interpretation becomes independent of the molecular conformation.

For ^{13}C-labeled RNA fragments, methods have been proposed to correlate the H2 protons to the H8 protons via the complex network of J-coupled ^{13}C atoms by isotropic mixing (Legault et al., 1994; Marino et al., 1994). However, isotopic labeling is not always possible or desirable. For fragments which are not labeled with ^{13}C, including the vast majority of DNA oligonucleotides used in NMR studies, one has to search for alternative methods which are applicable when the ^{13}C atoms are only present in natural abundance. This paper demonstrates that, thanks to the quality of present-day spectrometers and probes, 'classical' heteronuclear correlation methods can provide through-bond H2–H8 correlations in unlabeled oligonucleotides, if detailed knowledge on heteronuclear J-couplings is available. We report on the application of an HMBC experiment (Summers et al., 1986) to the unlabeled DNA oligomer 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3'. This fragment adopts a hairpin structure and has been designed such that, on top of a regular B-type stem, three additional G/I–A pairs may be formed. Through-bond H2–H8 correlations were found for all four adenine residues present in this fragment. As a result, preliminary conclusions regarding the molecular conformations of these noncanonical base pairs could be drawn.

In Fig. 1 portions of the HMBC spectrum obtained from 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3' are shown.
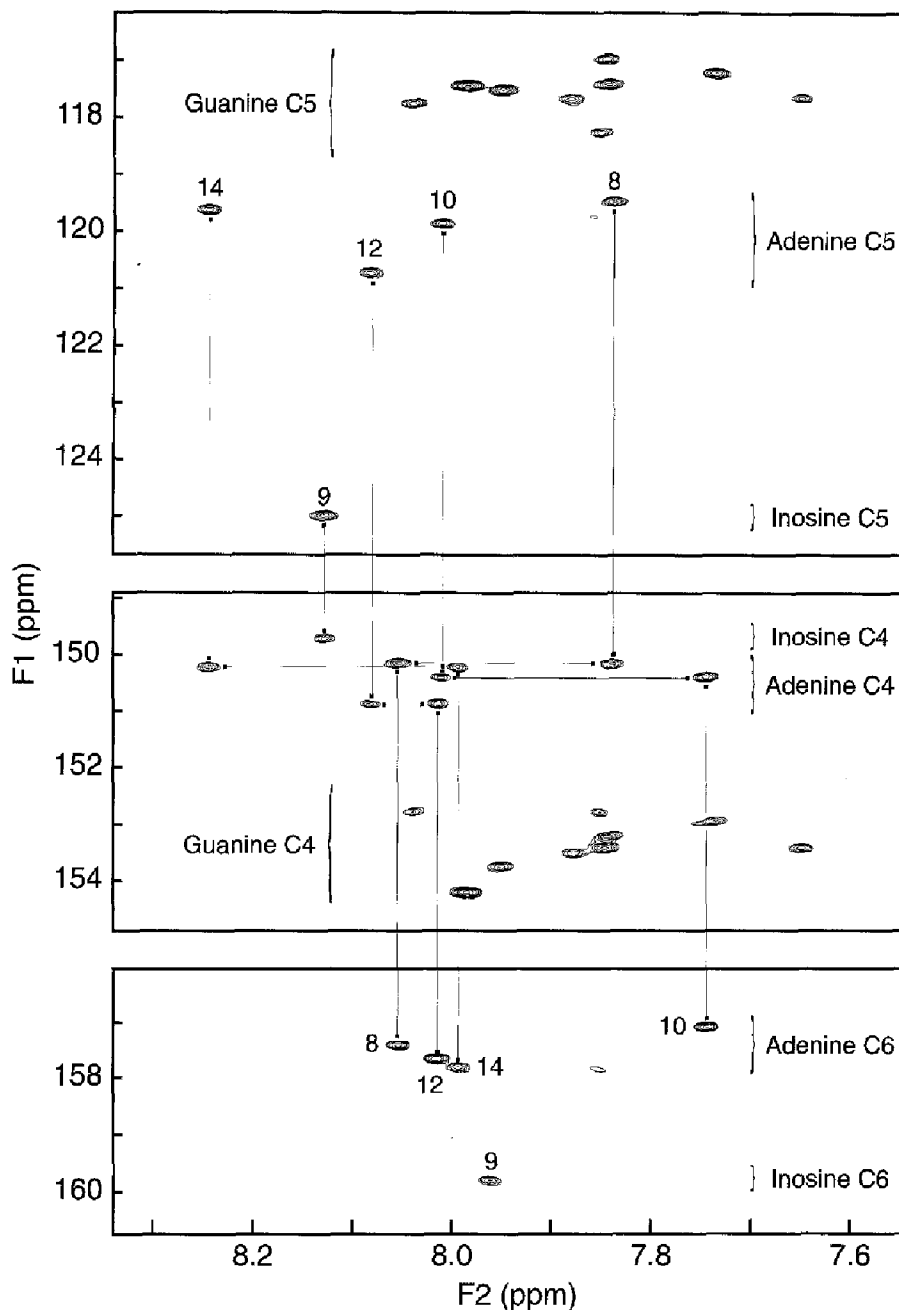

Fig. 1. Sections of the HMBC spectrum of a 1.5 mM sample of 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3' in 99.98% D_2O, pH 6.8, at 25 °C, by application of the pulse scheme shown in Fig. 2a. The spectrum was acquired in about 80 h on a Varian UNITYplus spectrometer operating at 750 MHz, although it has to be noted that the obtained signal-to-noise ratio allows the duration of the experiment to be reduced by a factor of two or three. The proton carrier was placed at the position of the water resonance and the spectral width in the F2 dimension was set to 7500 Hz. For the indirect dimension the spectral width was set to 11363 Hz, centered at 139 ppm. States-TPPI was used for phase-sensitive detection of the indirectly observed frequency (Marion et al., 1989). The spectrum was recorded with 2048 increments in t_2 and 1024 increments in t_1. Suppression of the HDO signal was achieved by presaturation during the relaxation delay. The delays δ and τ (see Fig. 2) were set to 2.5 and 60.0 ms, respectively. During detection, ^{13}C GARP decoupling was executed. The upper panel shows the C5-H8 correlations for adenine and guanine residues. In the middle panel the C4-H8 and C4-H2 cross peaks, used to correlate the adenine H8 and H2 resonances, are observed. The lower panel shows the C6-H2 correlations for adenine and inosine residues.

The spectrum shows long-range H2 to C4 and C6, as well as H8 to C4 and C5 correlations. The adenine H2-H8 correlations were obtained from this HMBC spectrum in two steps. First, the H2 and H8 resonances (both residing in the same spectral region, i.e., 7.6–8.3 ppm) were unambiguously distinguished by the chemical shift dispersion of the J-coupled carbon atoms. This was established using the fact that the H2 (and not the H8) protons are correlated with the C6 atoms, resonating between 155 and 159 ppm, whereas the H8 (and not the H2) protons are corre-

lated with the C5 atoms (115–125 ppm). Although, of course, the distinction between H2 and H8 resonances can also be obtained from the chemical shift difference of the C2 and C8 resonances in an HMQC spectrum (Müller, 1979), an HMBC spectrum may offer certain specific advantages (vide infra). In the second step, the actual through-bond H2–H8 correlations were established via a common J-coupled carbon atom. As can be seen in Fig. 1, each adenine C4 resonance (148–153 ppm) shows two cross peaks, one to an H2, the other to an H8 proton. Together, these cross peaks connect the H2 protons, via the C4 atoms, to the corresponding H8 protons, and thus provide the H2–H8 correlation.

The success of the method relies, to a large extent, on the accurate knowledge of the heteronuclear three-bond J-coupling constants involved. These values were used to calculate the optimum durations of the $^3J_{CH}$ dephasing–rephasing delays, τ, in the HMBC pulse sequence (Fig. 2a), necessary to obtain the long-range correlations, as shown in Fig. 1. For adenine, J-coupling constants were established from $^{13}C,^{15}N$-labeled AMP, as reported in Fig. 2b (Hilbers and Wijmenga, 1996; Ippel et al., 1996). Figure 2c presents the resulting C4–H2 and C4–H8 coherence transfer functions, for different values of the transverse relaxation time, T_2. It can be seen in this figure that maximum C4–H8 cross-peak intensities are obtained for
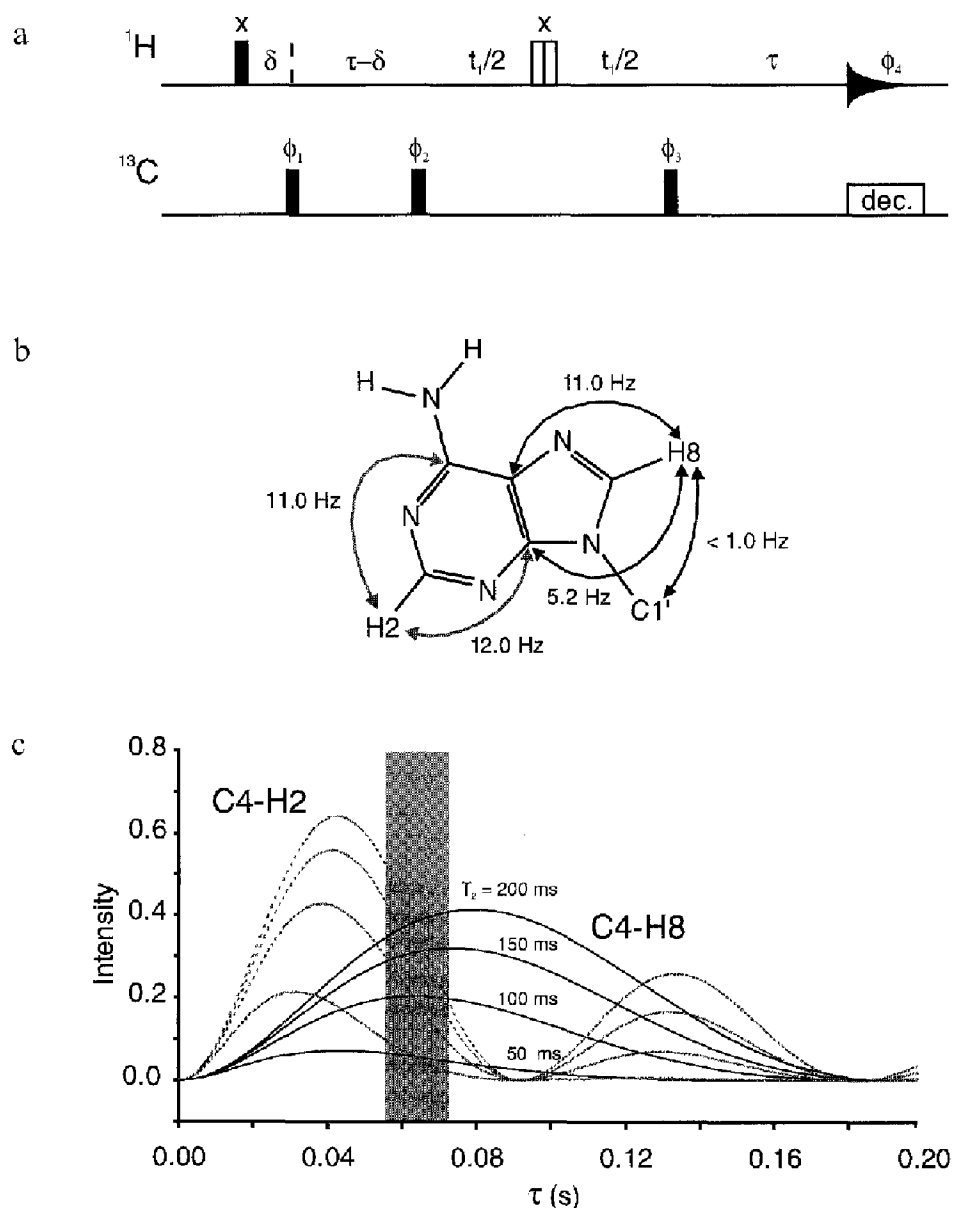

Fig. 2. (a) Pulse sequence of the HMBC experiment (Summers et al., 1986). The narrow filled bars represent 90° pulses; the double open bars represent 180° pulses. $\phi_1 = 4(x),4(-x)$; $\phi_2 = (x),(-x)$; $\phi_4 = (x),2(-x),(x)$; $\phi_5 = 2(x),2(-x)$. (b) Three-bond J-coupling constants in adenine residues (Hilbers and Wijmenga, 1996; Ippel et al., 1996). (c) Build-up curves for the adenine C4–H2 (dashed lines) and C4–H8 (solid lines) intensities in HMBC: $I_{CH} = \cos(\pi J_{CH}\delta) \cdot \sin(\pi J_{CH}(\tau-\delta)) \cdot \sin(\pi J_{CH}\tau) \cdot e^{2\tau/T_2}$, for four different values of T_2. The shaded box marks the values of τ which are suitable for H2–H8 correlations.
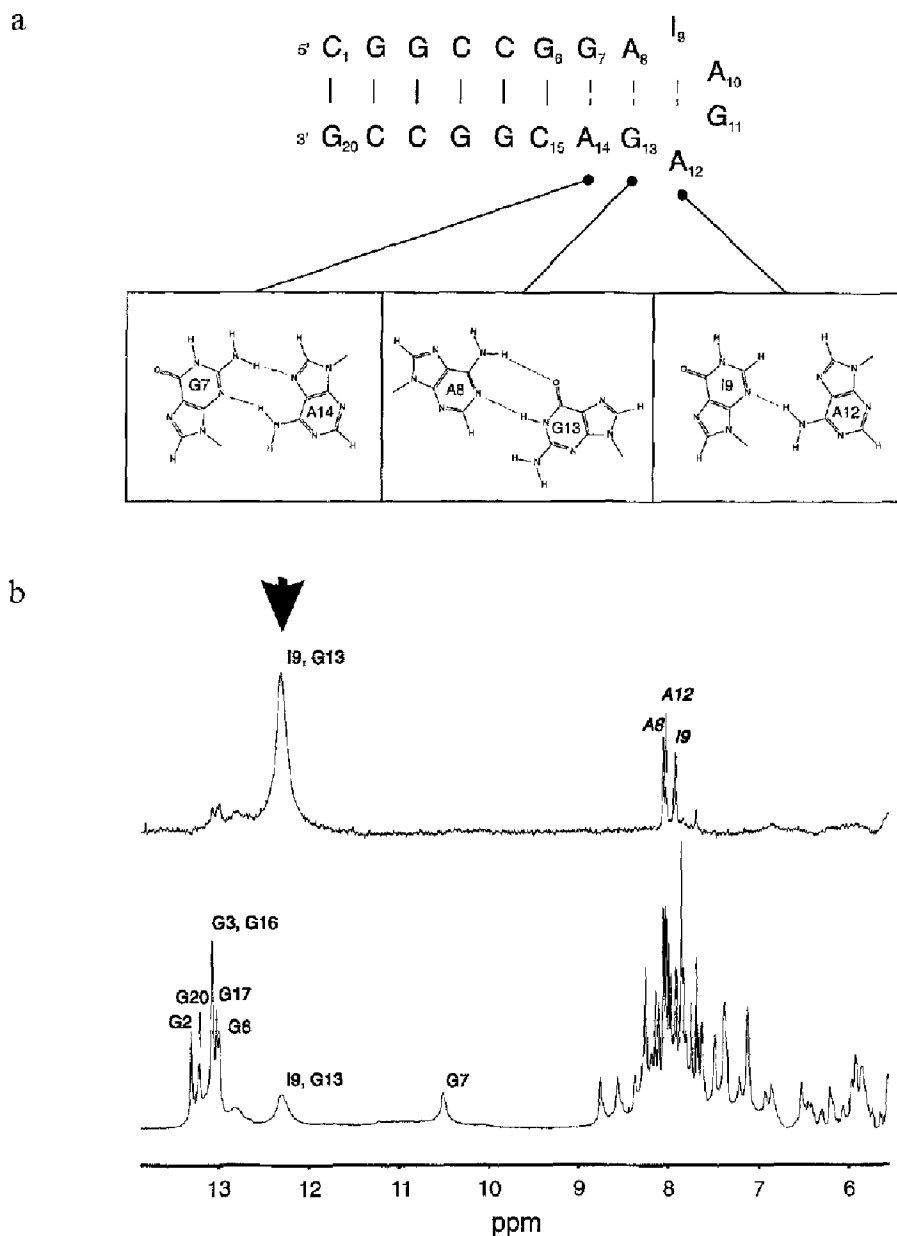
a

b

Fig. 3. (a) Proposed folding and pairing scheme, with the noncanonical base-pair geometries G7(*anti*)–A14(*anti*), A8(*anti*)–G13(*anti*), and sheared I9(*anti*)–A12(*anti*). (b) Exchangeable proton spectrum and NOE difference spectrum of the DNA 20-mer, 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3', after irradiation of the resonance at 12.3 ppm.

τ around 80 ms. The C4–H2 cross peaks show minimum intensities for this value of τ and maximum intensities for τ around 40 ms. The figure also shows that, to correlate H2 and H8, the optimum value for τ lies in the range between 60 and 65 ms (as indicated by the shaded area) and is almost independent of T_2. On the other hand, the resulting intensities are of course strongly influenced by this latter parameter. For unenriched material, as is considered here, the relaxation of the C4–H2 pair will not be strongly influenced by the size of the molecule. This is somewhat more unfavorable for the C4–H8 pair, but we estimate that the method may work for molecules consisting of up to 40 nucleotides, provided that samples with

sufficiently high concentrations can be prepared. The relevance of knowing the J-coupling constants involved is nicely illustrated by the fact that the C4–H2 cross peak for the inosine residue in the fragment could not be detected. The absence of this correlation can be understood if a different $^3J_{C4-H2}$ value is presumed for inosine compared to adenine, i.e., a value of about 15 Hz for this coupling constant would result in an intensity near zero for the value of τ used in the experiment (60 ms). Nevertheless, the identification of the inosine H2 resonance followed from the other results by elimination.

Using the H8 resonances in a conventional assignment strategy involving the sequential 5'-sugar–base-3' contacts

in a NOESY spectrum, the H2-C6-C4-C5-H8 spin systems were subsequently assigned to specific positions in the sequence. With these results available, an additional useful feature of the HMBC spectrum becomes apparent: the base protons can be assigned to the various residue types. As indicated in Fig. 1, it nicely demonstrates that within the resonance regions of the C4, C5 and C6 atoms, respectively, separate, well-defined chemical shift regions for the guanine, inosine and adenine residues can be identified (Breitmaier and Voelter, 1987; LaPlanta et al., 1988). The guanine C5 atoms, for instance, resonate between 117 and 119 ppm, the adenine C5 between 119 and 121 ppm, and the inosine C5 near 124 ppm. The correlations observed for the H6 protons of the cytosine residues of the fragment, i.e., to C2 (158–160 ppm) and C4 (167–169 ppm), even appear in completely distinct regions of the spectrum (data not shown). This distribution of carbon chemical shifts is very advantageous, compared those observed in an HMQC spectrum. Cytosine C6 resonances are usually present in the same spectral region as the C8 resonances of adenine and inosine, i.e., between 140 and 144 ppm. The C8 resonances of guanine present in an HMQC spectrum are situated somewhat shifted upfield (136–140 ppm) relative to this region. Here, the characteristic carbon chemical shifts confirm the sequential assignments obtained from the NOESY spectrum. In general, they can be used to assign the proton resonances to the various residue types.

The results obtained by the HMBC experiment were of central importance for determining the G/I–A base pair configurations in the 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3' fragment, for which intraresidue H8–H1'/H2'/H2" NOE intensities showed that all glycosidic torsion angles reside in the anti-domain (Wüthrich, 1986; Wijmenga et al., 1993). Two reasonable possibilities for these noncanonical base pairs have been postulated, i.e., 'anti–anti' and 'sheared anti anti', identified by the donating group of the guanine/inosine residue in the hydrogen bond (Kan et al., 1983; Brown et al., 1986; Li et al., 1991). In an anti–anti base pair it is the imino proton, while in a sheared anti–anti configuration it is the amino group of the guanine which serves as the hydrogen donor. This is reflected in the resonance positions of the guanine/inosine imino protons. A guanine imino proton in an anti–anti base pair resonates near 12 ppm, whereas the same proton in a sheared anti–anti base pair resonates near 10.5 ppm. For inosine these values are 15 ppm and 12 ppm, respectively (Kearns, 1977; Uesugi et al., 1987).

Inspection of the exchangeable proton spectrum of 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3' reveals that three, slightly broadened resonances are present in the region between 10.0 and 12.5 ppm, while no resonances could be observed downfield from 14 ppm (Fig. 3b). Thanks to the unambiguous resonance assignment of the H2 and H8 protons, we were able to interpret the NOEs,

observed in one-dimensional NOE difference spectra, between these imino resonances and resonances in the aromatic proton region, and to derive their sequential assignments (Fig. 3). Irradiating the two overlapping resonances at 12.3 ppm resulted in two intense NOEs at the resonance positions of H2(I9) and H2(A8). On the basis of their high intensities they were attributed to an intraresidue imino-H2 NOE for I(9) and a cross-strand imino(G13)–H2(A8) NOE, respectively. In addition, a weaker NOE at the resonance position of H2(A12) was observed, which was assigned to a sequential imino(G13)–H2(A12) NOE. The imino resonance at 10.6 ppm, attributed to imino(G7), did not produce NOEs in the aromatic proton region (data not shown). Subsequently, these data were interpreted in terms of the molecular conformations of the three noncanonical G/I–A base pairs (Fig. 3). The imino(G13)-H2(A8) NOE, as well as the resonance position of the G13 imino proton at 12.3 ppm are explained by the formation of an anti–anti base pair formed by A8 and G13. The relative up-field chemical shift of G7 imino proton (10.6 ppm), on the other hand, suggests that G7 may be involved in a sheared anti–anti G A base pair with A14. An anti–anti G7–A14 pair can certainly be ruled out, because this would have resulted in an imino(G7)–H2(A14) NOE, which is absent, and a more downfield chemical shift of the imino proton of G13. For similar reasons, an anti–anti I9–A12 base pair can be ruled out. Instead, a I9–A12 base pair involving an hydrogen bond between N3 of I9 and the amino group of A12, the same type of hydrogen bonding as in a sheared anti–anti base pair, might be present. Preliminary model building predicts an NOE between the imino proton of G13 and the H2 proton of A12, if such a base pair is formed, consistent with the observed imino(G13)–H2(A12) NOE.

The obtained results for 5'-d(CGGCCG-GAIAGAGA-CGGCCG)-3' agree with the observations for a long stretch of alternating d(GA)$_n$ folded into a hairpin analogous to the structure studied here (Huertas et al., 1993). In chemical modification studies it was found that the adenine residues in the 5' and 3' parts of the fragment showed different sensibilities to DEPC modification, indicating that the protection of the adenine N7 atoms in the two strands of the hairpin stem is dissimilar. Although the authors originally proposed a structure comprising alternating G(anti)–A(syn) and G(anti)–A(anti) base pairs on the basis of these data, they did not exclude the structure with alternating sheared G(anti)–A(anti) and G(anti)–A(anti) base pairs found here.

In conclusion, a new application of the HMBC experiment has been presented. It provides a useful means to discriminate between H2 and H8 proton resonances, to assign the base proton resonances to the various residue types and, most importantly, to correlate the H2 and H8 protons for adenine or inosine residues in natural abundance ^{13}C fragments. The method is particularly valuable

if the H2 resonance assignments cannot be made on the basis of standard NOE contacts because the strand adopts a nonhelical structure or because the bases are not involved in regular Watson Crick base pairing. Using the HMBC experiment, the H2–H8 correlations for the adenine residues of a DNA 20-mer were established, resulting in the unambiguous site-specific assignment of all H2 resonances. Thanks to these results preliminary conclusions could be drawn, regarding the molecular conformations of the noncanonical G/I–A base pairs in the hairpin formed by this fragment.

Acknowledgements

References

Breitmaier, E. and Voelter, W. (1987) In *Carbon-13 NMR Spectroscopy: High-Resolution Methods and Applications in Organic Chemistry and Biochemistry* (Ed., Ebel, H.F.), Weinheim, New York, NY, U.S.A.

Brown, T., Hunter, W.N., Kneale, G. and Kennard, O. (1986) *Proc. Natl. Acad. Sci. USA*, 83, 2402–2406.

Hilbers, C.W. and Wijmenga, S.S. (1996) In *Encyclopedia of Nuclear Magnetic Resonance* (Eds, Grant, D.M. and Harris, R.K.), Wiley, Chichester, U.K., pp. 3346–3359.

Huertas, D., Bellsolell, L., Casasnovas, J.M., Coll, M. and Azorín (1993) *EMBO J.*, 12, 4029–4038.

Ippel, J.H., Heus, H.A., Wijmenga, S.S. and Hilbers, C.W. (1996) *Magn. Reson. Chem.*, manuscript submitted for publication.

Kan, L.S., Chandrasegaran, S., Pulford, S.M. and Miller, P.S. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 4263–4265.

Kearns, D.R. (1977) *Annu. Rev. Biophys. Bioeng.*, 6, 477–523.

LaPlanta, S.R., Boudreau, E.A., Zanatta, N., Levy, G.C. and Borer, P.N. (1988) *Biochemistry*, 27, 7902–7909.

Legault, P., Farmer II, B.T., Mueller, L. and Pardi, A. (1994) *J. Am. Chem. Soc.*, 116, 2203–2204.

Li, Y., Zon, G. and Wilson, W.D. (1991) *Proc. Natl. Acad. Sci. USA*, 88, 26–30.

Marino, J.P., Prestegard, J.H. and Crothers, D.M. (1994) *J. Am. Chem. Soc.*, 116, 2205–2206.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, 85, 608–613.

Müller, L. (1979) *J. Am. Chem. Soc.*, 101, 4481–4484.

Summers, M.F., Marzilli, L.G. and Bax, A. (1986) *J. Am. Chem. Soc.*, 108, 4285–4294.

Uesugi, S., Oda, Y., Ikehara, M., Kawasa, Y. and Ohtsuka, E. (1987) *J. Biol. Chem.*, 262, 6965–6968.

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, New York, NY, U.S.A., pp. 217–288.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.