Journal of Biomolecular NMR, **23:** 257–262, 2002.
KLUWER/ESCOM

Differentiation between unlabeled and very-low-level fully ^{15}N,^{13}C-labeled nucleotides for resonance assignments in nucleic acids

Anh Tuân Phan* & Dinshaw J. Patel
Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, U.S.A.

Received 12 March 2002; Accepted 26 June 2002

Key words: low isotopic abundance, low-level fully labeled nucleotides, model-independent assignments, nucleic acids, site-specific enrichment

Abstract

Based on different characteristics between unlabeled and fully ^{15}N,^{13}C-labeled nucleotides, we develop a method for unambiguous resonance assignments in nucleic acids following site-specific fully ^{15}N,^{13}C isotope incorporation at very low levels[1]. The J-couplings between heteronuclei provide for distinction between the NMR signals of the fully labeled nucleotides and those of the natural abundance nucleotides. The method is demonstrated for DNA oligonucleotides[2], in the dimeric G-quadruplex [d(GGGTTCAGG)]$_2$ and in the 22-nucleotide human telomeric fragment d[AG$_3$(TTAG$_3$)$_3$]. We expect this approach to be useful for selective monitoring of important functional domains and of their interactions in large nucleic acids.

Introduction

Resonance assignments are critical for structure determination of nucleic acids by NMR. NOE and through-bond correlation techniques are commonly used to achieve this goal (reviewed in Wijmenga and van Buuren, 1998). However, unambiguous assignments cannot always be obtained solely by these approaches, especially for large macromolecules. Both through-space and through-bond approaches require some reliable resonance identifications as a starting point. The former is often model-dependent, while the latter is sometimes insensitive or lacking in resolution. The site-specific 100% isotope labeling approach (Jiang et al., 1997) is a straightforward but costly alternative. We have recently proposed a simple and cost-effective approach (Phan, 1999; Phan et al., 2001) to unambiguously assign resonances originating from a few-percent-(^{15}N/^{13}C)-labeled nucleotide (as low as 1%) in a DNA oligonucleotide (Phan and Patel, 2002). In this approach, one nucleotide (or a small number of nucleotides) of the DNA sequence is incorporated using a phosphoramidite mix which includes a small fraction of fully ^{15}N/^{13}C-labeled phosphoramidite**,***. The resonances of the labeled nucleotide are distinguished from unlabeled ones primarily by their heteronuclear correlation peak intensity, which is increased as compared to the natural abundance reference spectrum. However, intensity comparisons are difficult in crowded regions, even when much higher labeling levels are used.

In a previous paper (Phan and Patel, 2002) we pointed out that the fully labeled nucleotide contains ^{15}N and ^{13}C at every position, in contrast to random distribution of ^{15}N and ^{13}C at natural abundance. This makes C1′-H1′ and C5′-H5′/H5″ correlation peaks disappear in the CT-HSQC spectrum (constant delay

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

**A 100%-^{15}N,^{13}C-labeled nucleotide, which contains ^{15}N and ^{13}C at every position, is referred to as a fully labeled nucleotide. Dilution of such a nucleotide with its unlabeled counterpart leads to a low-level (or diluted) fully labeled nucleotide.

***DNA oligonucleotides were synthesized by the β-cyanoethyl phosphoramidite method. Each fully labeled phosphoramidite, purchased from Martek, was diluted with its unlabeled counterpart to obtain the desired labeling levels, typically 0.5 to 5%. Only one nucleotide at a time is labeled in the DNA oligonucleotides.

T = 25 ms) of a 1% fully labeled sample (Phan and Patel, 2002). In the present work, we show how this property can be used to differentiate between unlabeled and fully labeled nucleotides (even at very low levels of labeling) for unambiguous resonance assignments. We demonstrate the method in the dimeric G-quadruplex $[d(GGGTTCAGG)]_2$, which contains G•G•G•G tetrads sandwiched between G•(C-A) triads (Kettani et al., 2000), and in the intramolecular monomeric G-quadruplex $d[AG_3(TTAG_3)_3]$ from the G-rich strand of the human telomere (Wang and Patel, 1993).

Results and discussion

Sugar protons usually resonate in the most crowded region of the proton NMR spectrum of nucleic acids. In many NMR studies, sugar protons such as H4′, H5′, H5″ could not be assigned and used for structural characterization. A correlation experiment involving ^{13}C-^{13}C J-couplings (e.g. a HCCH experiment) would directly select correlations from fully labeled sugars, because a labeled sugar ^{13}C has ^{13}C neighbors in contrast to a natural abundance sugar ^{13}C whose neighbors are ^{12}C for 99% of them. In the literature, ^{13}C-^{13}C filters have been previously proposed for different purposes, such as selection of different isotopomers (Wand et al., 1995) and base-type-selective spectral editing (Brutscher et al., 2001).

To select protons such as H1′, whose attached $^{13}C1'$ is well separated in chemical shift from the neighboring C2′, we used a simple selective CT-HSQC (sCT-HSQC) pulse sequence (Figure 1A). This scheme differs from the standard CT-HSQC (Santoro and King, 1992; van de Ven and Philippens, 1992; Vuister and Bax, 1992) by the addition of two 90° C2′-selective pulses in the center of the constant-time period (between times *a* and *b*). These two pulses act as a 0° pulse for $\Phi_3 = -x$ and as a 180° C2′-selective pulse for $\Phi_3 = x$. The 180° C2′-selective pulse would decouple the effect of $J_{C1'C2'}$ during the constant-time period. The C1′-H1′ correlation originating from the labeled nucleotide is therefore sensitive to Φ_3, while the natural abundance C1′-H1′ correlations are not. If the receiver is phase-cycled according to Φ_3, the C1′-H1′ correlation from the labeled nucleotide is observed, while those from unlabeled nucleotides cancel out because they are detected with opposite sign for $\Phi_3 = -x$ and x. Alternatively, one can perform separate experiments with $\Phi_3 = -x$ and $\Phi_3 = x$; the spectra from fully labeled and from natural abundance moieties are then obtained by addition and subtraction of sub-spectra; a single sub-spectra recorded with $\Phi_3 = -x$ (standard CT-HSQC) may be used for different band-selective experiment with $\Phi_3 = x$. Note that ^{15}N decoupling may be applied during the constant-time period to avoid signal loss due to $^1J_{CN}$.

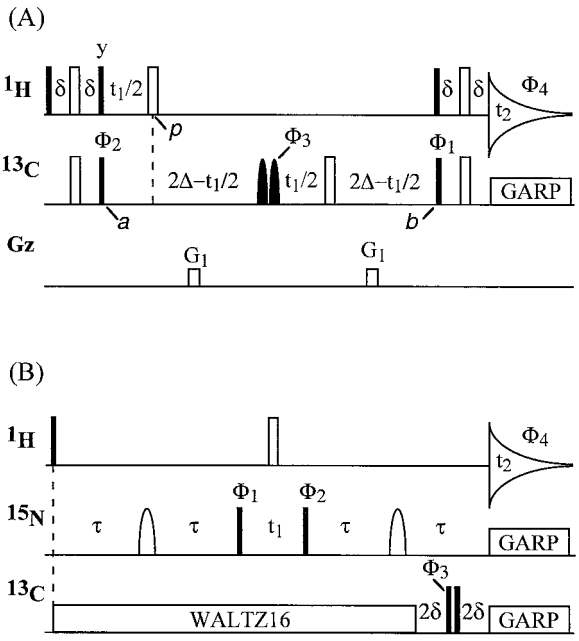

Figure 1. Pulse sequences used in this work for selection of fully ^{15}N,^{13}C-labeled nucleotides. Rectangles represent 180° (open) and 90° (filled) hard pulses; ovals represent 180° (open) and 90° (filled) selective pulses. The selective pulses were generated by a frequency-shifted DANTE sequence (Morris and Freeman, 1978). Unless indicated otherwise, pulses are applied along the x-axis. Phase cycling: Φ_1 = x, −x; Φ_2 = 2(x), 2(−x); Φ_3 = 4(x), 4(−x); Φ_4 = x, −x, −x, x, −x, x, x, −x. In addition, Φ_1 is phase-cycled in hypercomplex mode. (A) sCT-HSQC pulse sequence used to select protons of fully labeled nucleotides using the J-couplings between vicinal ^{13}C neighbors. (B) HN(C) pulse sequence used to select the N7-H8 correlation in fully labeled purine bases. Delays: $\delta \approx 1/(4J_{CH})$; $\tau \approx 1/(4J_{NH})$; $\Delta = 1/(4J_{CC})$. Gradients: G_1 = (0.5 ms, 9 G/cm).

Figure 2C shows the C1′-H1′ correlation of G2 (fully labeled in 1% of the strands) in $[d(GGGTTCAGG)]_2$, selected in a 1D sCT-HSQC spectrum. The C3′-H3′ correlation of G2 is also observed. The one-bond sugar-proton correlation peak is negative for C1′ which has an odd number (one) of carbon neighbors, and positive for C3′ which has an even number (two) of carbon neighbors. With appropriate band-selective ^{13}C pulses, this sCT-HSQC experiment can be applied to other sugar protons. For example, C1′-selective pulses provide for obser-

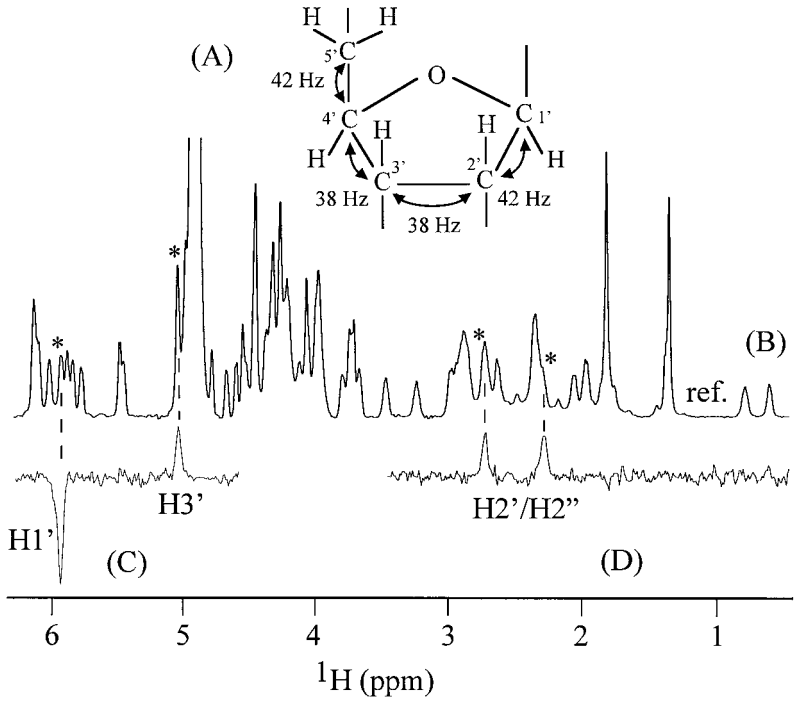

Figure 2. (A) Scheme of ribose sugar with arrows indicating $^{1}J_{CC}$ couplings. (B-D) Sugar proton spectra of a $[d(GGGTTCAGG)]_2$ sample, where G2 is fully $^{15}N,^{13}C$-labeled in 1% of the strands (strand concentration 3.5 mM; 100 mM NaCl; 2 mM phosphate; pH 6.6; 10 °C; 99.9% D_2O). The 600 MHz spectra were obtained with a Varian Unity Inova spectrometer. (B) 1D reference spectrum. (C, D) 1D sCT-HSQC spectra recorded by sequence 1A: $t_1 = 0$; pulse *p* omitted; 8192 scans; $\delta = 1.5$ ms; $\Delta = 6.25$ ms; repetition delay, 1 s; total measurement time, 2.5 h. The selective pulses were set at the chemical shift of C2′ (39 ppm) for (C) and of C1′ (84 ppm) for (D); the 90° selective pulse lengths were 0.6 ms (C) and 1.3 ms (D) respectively. Asterisks indicate sugar protons assigned to G2.

vation of the C2′-H2′/H2″ correlations of the labeled nucleotide (Figure 2D).

The sensitivity of the sCT-HSQC method is tested for the 22-nucleotide human telomeric fragment $d[AG_3(TTAG_3)_3]$ at very low concentration (0.5 mM). Figure 3B shows the C1′-H1′ and C3′-H3′ correlations of G2 (fully labeled in 2% of the strands) recorded when the selective pulses are set on C2′. Figure 3C shows the C5′-H5′/H5″ correlations of G2 recorded when the selective pulses are set on C4′. The C3′-H3′ correlation is also observed in this experiment (not shown). Figure 3D shows the C4′-H4′ correlation of G2 recorded when the selective pulses are set on C5′. These results verify and confirm the previous assignments (Wang and Patel, 1993).

The sCT-HSQC method can also be used for pyrimidine aromatic protons. The sensitivity is even higher than for sugar protons thanks to larger $^{1}J_{CC}$ couplings (∼55–68 Hz). This is demonstrated in the Figure 4 for a cytidine of $[d(GGGTTCAGG)]_2$. The H6 proton of cytidine 6 (fully labeled in 0.4% of the strands) is distinguished from other aromatic protons by a 1D experiment with the selective pulses set on the $^{13}C5$ frequency. The H5 proton could be selected by setting the selective pulses at the $^{13}C6$ frequency (not shown). The method can be applied to other pyrimidine bases such as thymidine (T) and uridine (U).

For purine bases, the sCT-HSQC experiment is not applicable to C8-H8 or C2-H2, since neither C8 nor C2 has a carbon neighbor. A HCN experiment (Sklenár et al., 1993; Marino et al. 1997; Fiala et al., 2000; Riek et al., 2001) can be used to select fully $^{15}N,^{13}C$-labeled base protons, but consecutive magnetization transfer by $^{1}J_{CH}$ and $^{1}J_{CN}$, is not very efficient due to the small $^{1}J_{CN}$ coupling (∼11 Hz) (reviewed in Wijmenga and van Buuren, 1998). Alternatively, one can use the $^{2}J_{NH}$ coupling (∼11 Hz) for magnetization transfer and the $^{1}J_{CH}$ coupling for resolving the final spectrum (not for a transfer step). A selective HMBC experiment, which is derived from the HMQC and HMBC sequences (Summers et al., 1986) by the addition of 180° N9-selective pulses (to eliminate the effect of $^{2}J_{N9H8}$) (Sklenár et al., 1994; Phan

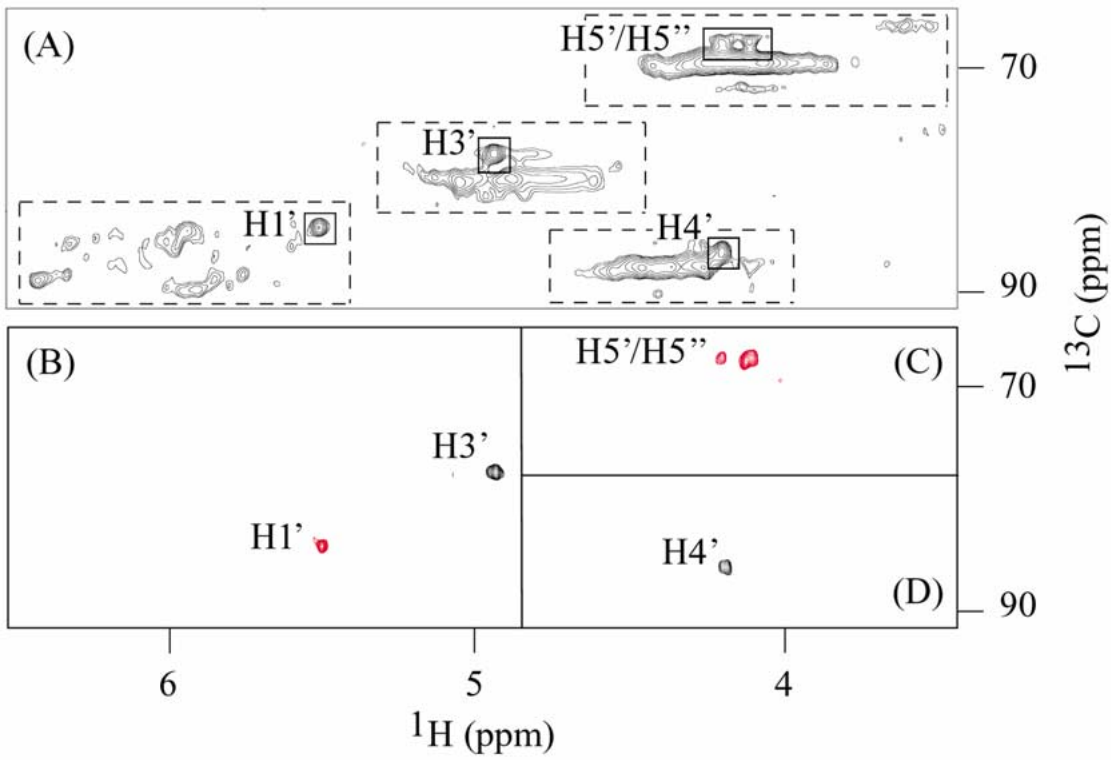

Figure 3. Assignment of G2 sugar protons of d[$AG_3(TTAG_3)_3$] in a sample where G2 is fully ^{15}N,^{13}C-labeled in 2% of strands (strand concentration 0.5 mM; 100 mM NaCl; pH 7; 25 °C; 99.9% D_2O). Spectral width, 6 kHz in both ^{1}H and ^{13}C dimensions; 16 complex t_1 points; 1024 complex t_2 points; repetition delay, 1 s. (A) Standard HSQC (Bodenhausen and Ruben, 1980): 1024 scans per FID; total measurement time, 9.5 h. (B-D) sCT-HSQC: 2048 scans per FID; $\delta = 1.5$ ms; $\Delta = 6.25$ ms, total measurement time, 19 h. The selective pulses were set at the chemical shift of C2′ (39 ppm) for (B), of C4′ (87 ppm) for (C), and of C5′ (66 ppm) for (D); the 90° selective pulse lengths were 0.6 ms (B), 0.9 ms (C), and 0.9 ms (D) respectively. Negative peaks are shown in red.

and Patel, 2002) and ^{13}C decoupling during excitation (to suppress unwanted pathways by all ^{13}C), correlates N7 to H8. In the selective HMBC spectrum of $[d(GGGTTCAGG)]_2$ (Figure 5B), the H8 proton of G2 (fully labeled in 1% of the strands), simultaneously coupled with ^{15}N7 and ^{13}C8 in fully labeled bases, exhibits a N7-H8 correlation peak with a splitting $^{1}J_{C8H8}$ ($\sim$215 Hz) in the proton dimension, when ^{13}C is not decoupled during acquisition. When it is decoupled, the splitting disappears and the intensity of the central peak increases (Figure 5C). This specific pattern distinguishes fully labeled bases from unlabeled counterparts.

This method may be inapplicable in case of spectral overlap. If so, one may instead eliminate the signals from the natural abundance correlations with a ^{13}C-filter (Otting and Wüthrich, 1990). Figure 1B presents an efficient and sensitive way to do that without any additional delays. It could be called a HN(C) experiment. Figure 5D shows a HN(C) spectrum, where the N7-H8 correlation of G2 is selected while those of other nucleotides are filtered out.

Conclusion

Our approach has been to prepare DNA oligonucleotide samples at millimolar concentrations (e.g., 1 μmol of DNA in 0.5 ml of water) where the majority of strands are unlabeled and a minority of strands (ca. 1%) contains one fully labeled nucleotide. For a given quantity of labeled phosphoramidite (hence for a given cost), an alternate approach would be to prepare a small sample (e.g., 0.01 μmol) where all the strands (100%) contain one fully labeled nucleotide. The DNA could then be diluted in the standard volume of water or put in a small volume such as a capillary. Some practical considerations argue in favor of the former approach. Thus, there is less loss of the labeled material during work up of phosphoramidite oligonu-

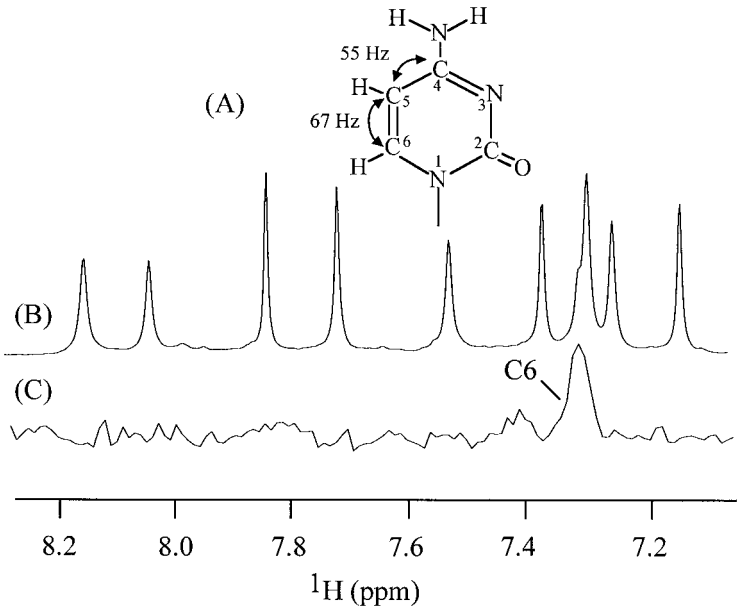

Figure 4. (A) The $^{1}J_{CC}$ couplings in the cytidine. (B, C) Assignment of C6(H6) in $[d(GGGTTCAGG)]_2$ using a sample where C6 is fully $^{15}N,^{13}C$-labeled in 0.4% of the strands (same condition as in Figure 2). (B) 1D reference spectrum. (C) 1D sCT-HSQC spectrum recorded by sequence 1A: $t_1 = 0$; selective pulses on $^{13}C5$ (97 ppm); 90° selective pulse length, 0.6 ms; pulse *p* omitted; 8192 scans; $\delta = 1.3$ ms; $\Delta = 3.7$ ms; repetition delay, 1 s; total measurement time, 2.5 h.

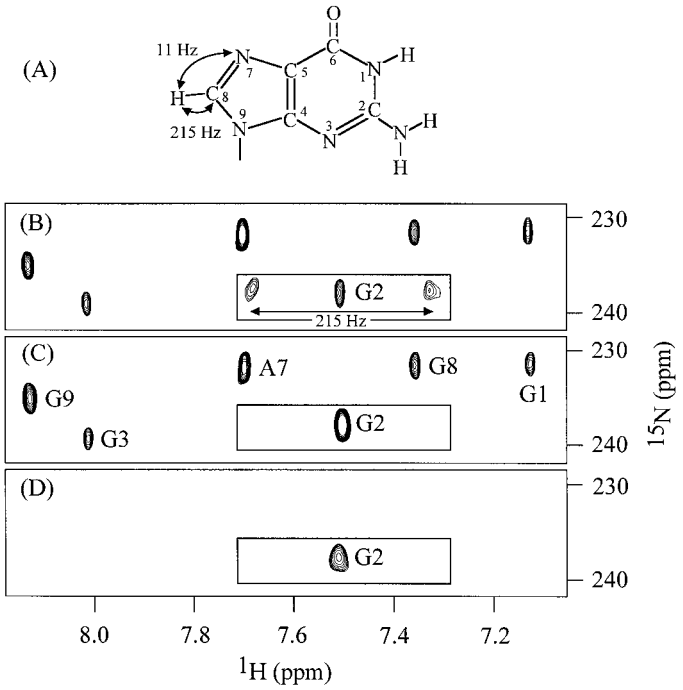

Figure 5. (A) The $^{1}J_{CH}$ and $^{2}J_{NH}$ couplings of guanosine used in the experiment. (B-D) Assignment of G2(H8) in $[d(GGGTTCAGG)]_2$ (same sample and condition as in Figure 2). Selective HMBC spectra recorded without (B) and with (C) ^{13}C decoupling during acquisition. (D) HN(C) spectrum recorded by sequence 1B. Spectral width, 6 kHz and 8 kHz respectively in the ^{1}H and ^{15}N dimensions; 32 complex t_1 points; 1024 complex t_2 points; 512 scans per FID; $\delta = 1.25$ ms; $\tau = 15.5$ ms; repetition delay, 1.1 s; total measurement time, 11 h. The N7-H8 correlation patterns for G2 are framed.

cleotide synthesis at the 1 μmol scale, compared with synthesis at the smaller 0.01 μmol scale. Furthermore, heteronuclear correlations of ^{13}C or ^{15}N in natural abundance, and proton-proton NOESY are optimally sensitive only at millimolar concentrations. In addition, one can also consider intermediate labeling situations (e.g., 10%) with the former approach.

In the present work, we have demonstrated how the distinct characteristics associated with unlabeled and labeled moieties can be used (e.g., by using ^{13}C-^{13}C or ^{15}N-^{1}H-(^{13}C) correlations) to distinguish between the signals of a small fraction of labeled nucleotides and those of unlabeled ones, providing means for unambiguous resonance assignments in a cost-effective site-specific low-level labeling approach. This opens opportunities for NMR studies of larger nucleic acids and their complexes. In particular, one can unambiguously assign some key resonances without assigning resonances of the entire molecule and subsequently obtain local structural information from NOE, through-bond (scalar and H-bond) J-couplings, as well as global structural information from residual dipolar coupling studies. Such an approach should prove useful for selective monitoring of functional domains and of their interactions in large nucleic acids. Note that larger molecules may require a TROSY variant of the pulse sequence and higher labeling levels than those used in this paper.

The approach will apply to RNA when labeled ribo-phosphoramidites become available at moderate cost. Presently, residue-type low-level labeling may be carried out by enzymatic synthesis. The approach would be also applicable to other biopolymers such as peptides and proteins.

Acknowledgements

We thank Maurice Guéron for helpful comments. The manuscript has been improved thanks to many suggestions of the reviewers. This research was supported by NIH GM-34504 to DJP.

References

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Brutscher, B., Boisbouvier, J., Kupce, E., Tisné, C., Dardel, F., Marion, D. and Simorre, J.P. (2001) *J. Biomol. NMR*, **19**, 141–151.

Fiala, R., Czernek, J. and Sklenár, V. (2000) *J. Biomol. NMR*, **16**, 291–302.

Jiang, F., Patel, D.J., Zhang, X., Zhao, H. and Jones, R.A. (1997) *J. Biomol. NMR*, **9**, 55–62.

Kettani, A., Basu, G., Gorin, A., Majumdar, A., Skripkin, E. and Patel, D.J. (2000) *J. Mol. Biol.*, **301**, 129–146.

Marino, J.P., Diener, J.L., Moore, P.B. and Griesinger C. (1997) *J. Am. Chem. Soc.*, **119**, 7361–7366.

Morris G.A. and Freeman, R. (1978) *J. Magn. Reson.*, **29**, 433–462.

Otting, G. and Wüthrich, K. (1990) *Quart. Rev. Biophys.*, **23**, 39–96.

Phan, A.T. (1999) Ph.D. Thesis, Ecole Polytechnique, Palaiseau, France.

Phan, A.T. and Patel, D.J. (2002) *J. Am. Chem. Soc.*, **124**, 1160–1161.

Phan, A.T., Guéron, M. and Leroy, J.L. (2001) *Meth. Enzymol.*, **338**, 341–371.

Riek, R., Pervushin, K., Fernandez, C., Kainosho, M. and Wüthrich, K. (2001) *J. Am. Chem. Soc.*, **123**, 658–664.

Santoro, J. and King, G. (1992) *J. Magn. Reson.*, **97**, 202–207.

Sklenár, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1993) *J. Am. Chem. Soc.*, **115**, 12181–12182.

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

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

van de Ven, F.J.M. and Philippens, M.E.P. (1992) *J. Magn. Reson.*, **97**, 637–644.

Vuister, G. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.

Wand, A.J., Bieber, R.J., Urbauer, J.L., McEvoy, R.P. and Gan, Z. (1995) *J. Magn. Reson., Ser. B*, **108**, 173–175.

Wang, Y. and Patel, D.J. (1993) *Structure*, **1**, 263–282.

Wijmenga, S.S. and van Buuren, B.N.M. (1998) *Progr. NMR Spectroscopy*, **32**, 287–387.