Journal of Biomolecular NMR, 4 (1994) 581–586
ESCOM

J-Bio NMR 204

Preparation and heteronuclear 2D NMR spectroscopy of a DNA dodecamer containing a thymidine residue with a uniformly ^{13}C-labeled deoxyribose ring

Akira Ono[a], Shin-ichi Tate[a], Yoshiharu Ishido[b] and Masatsune Kainosho[a,*]

[a]*Department of Chemistry, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji, Tokyo 192-03, Japan*
[b]*Laboratory of Pharmaceutical Chemistry, Tokyo College of Pharmacy, 1432-1 Horinouchi, Hachioji, Tokyo 192-03, Japan*

Received 10 March 1994
Accepted 22 April 1994

Keywords: Heteronuclear 2D NMR; Isotopically labeled DNA; ^{13}C-labeled thymidine; Solid-phase DNA synthesis

SUMMARY

[$^{13}C_5$]-2-Deoxy-D-ribose, synthesized from [$^{13}C_6$]-D-glucose (98% ^{13}C), was coupled with thymine to give [1',2',3',4',5'-$^{13}C_5$]-thymidine (T) in an 18% overall yield. The thymidine was converted to the 3'-phosphoramidite derivative and was then incorporated into a dodecamer 5'-d(CGCGAATTCGCG)-3' by solid-phase DNA synthesis. Preparation of 0.24 μmole of the labeled dodecamer, which is sufficient for a single NMR sample, consumed only 25 mg of glucose. By virtue of the ^{13}C labels, all of the ^1H-^1H vicinal coupling constants in the sugar moieties were accurately determined by HCCH-E.COSY.

Conformational diversity of the sugar moieties in DNA has been considered to be an important issue in elucidation of sequence-specific DNA–protein recognition processes (Saenger, 1984). Structural information on DNA in solution, however, can only be obtained for small oligomers, because of the difficulties in analyzing the ^1H NMR spectra of larger DNAs. Heteronuclear multidimensional NMR methods, which have been successfully used for $^{13}C/^{15}N$-labeled proteins (Ikura et al., 1990; Clore and Gronenborn, 1991) and more recently for RNAs (Batey et al., 1992; Nikonowicz et al., 1992; Nikonowicz and Pardi, 1992a,b,1993; Pardi and Nikonowicz, 1992; Farmer et al., 1993,1994; Michinicka et al., 1993; Sklenar et al., 1993a,b,1994), have not been applied in NMR studies of DNA. This is obviously due to the lack of efficient methods to prepare isotopically labeled DNAs, and therefore only a few simple heteronuclear NMR experiments for DNA oligomers with naturally abundant ^{13}C (Ashcroft et al., 1989,1991; Schmieder et al., 1992), and for DNAs with ^{13}C labels at specific atomic positions (Williamson and Boxer, 1988; Kellenbach et al., 1992; Lancelot et al., 1993) have been reported so far. Neither of these approaches leads to more sophisticated heteronuclear multidimensional experiments. The situation contrasts

*To whom correspondence should be addressed.

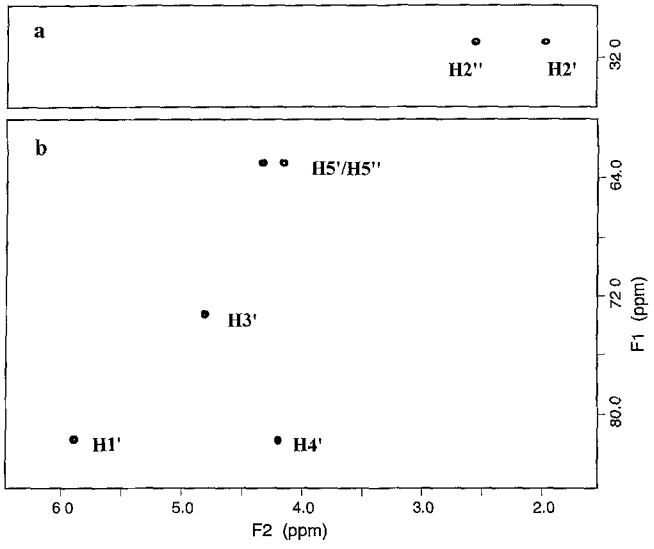

Fig. 1. The CT ^{1}H-^{13}C HSQC spectrum of a D_2O solution of 1.4 mM d(CGCGAATTCGCG)$_2$, containing 0.1 M NaCl, 0.01 M sodium phosphate and 0.1 mM EDTA. The cross peaks for $C_{2'}$-$H_{2',2''}$ are shown in (a) and the other C-H ring connectivities in (b). The solution (pH 7.0, direct meter reading) was transferred into a 5 mm o.d. Shigemi microcell (Shigemi Co. Ltd., Hachioji, Japan) to 11 mm height (effective volume of 160 µl). The spectrum was measured at 30 °C on a Bruker AMX-500 spectrometer with a total constant-time duration of 28 ms. Spectral widths in the ^{1}H and ^{13}C dimensions were 4504.5 and 7000 Hz, respectively. The ^{13}C rf was set at the center of the deoxyribose carbon signal region, i.e. 61.0 ppm. GARP decoupling with 1.56 kHz field strength was applied during the acquisition period (Shaka et al., 1985). The data were collected as a 128 (t_1) × 512 (t_2) complex matrix with 64 scans per t_1 increment, which was processed with FELIX (version 2.1, Biosym Technologies). The data were acquired over 4.8 h and processed after zero-filling in both dimensions to the final data set of 1024 (F_1) × 256 (F_2) real points.

strongly with that of RNA oligomers, which can be efficiently prepared by in vitro transcription systems employing T7 RNA polymerase (Milligan et al., 1987) and $^{13}C/^{15}N$-labeled ribomononucleotides isolated from *E. coli*, yeast, or other bacterial cells grown in isotopically labeled media (Batey et al., 1992; Nikonowicz et al., 1992; Michinicka et al., 1993). We have thus tried to develop practical methods to prepare labeled DNA in sufficient amounts for NMR studies, and at a reasonable cost. In doing so, we have successfully prepared a DNA dodecamer, 5'-d(CGCGAATTCGCG)-3', containing [1',2',3',4',5'-$^{13}C_5$]-thymidine (T) which has a fully ^{13}C-labeled deoxyribose ring (98% ^{13}C).

[$^{13}C_5$]-2-Deoxy-D-ribose was chemically synthesized from [$^{13}C_6$]-D-glucose (98% ^{13}C) in a 33% yield, as described previously for normal glucose (Hardegger, 1962), but with minor modifications. Subsequently, the labeled 2-deoxyribose was converted into α-2-deoxy-3,5-di-*O*-*p*-toluoyl-D-ribofuranosyl chloride (Hoffer, 1960) which was then coupled with silylated thymine to afford 3',5'-di-*O*-*p*-toluoylthymidine (Hubbard et al., 1984). Deprotection of the toluoyl group gave labeled thymidine in a 55% yield from the α-chlorosugar and the overall yield of [1',2',3',4',5'-$^{13}C_5$]-thymidine from the labeled glucose was 18%. The yield from the glucose was considerably higher than that of the labeled RNA monomer isolated from the microorganism cells cultured with labeled glucose, which is, as a mixture of four nucleotides, only about 4–5% glucose (Batey et al., 1992; Nikonowicz et al., 1992; Michinicka et al., 1993).

For the synthesis of DNA oligomers, labeled thymidine was converted into the corresponding nucleoside 3'-phosphoramidite, as described by Sinha et al. (1983). 5'-d-(CGCGAATTCGCG)-3' was then synthesized on a DNA synthesizer (Applied Biosystems) by the solid-phase phosphoramidite method (Beaucage and Caruthers, 1981). After deprotection and purification, the purity of the labeled dodecamer was higher than 99% according to HPLC analysis with a C-18 column. The constant-time ^{1}H-^{13}C correlation spectrum (CT-HSQC) (Santoro and King, 1992; Van de Ven and Phillippens, 1992; Vuister and Bax, 1992) of the dodecamer prepared by this procedure is shown in Fig. 1 with the assignment for the deoxyribose ring of T7, which was consistent with that reported by Hare et al. (1983). The ^{1}H-^{13}C CT-HSQC spectrum was measured over 4.8 h, but a spectrum with sufficient signal-to-noise ratio could be measured within a shorter time.

Conformational studies of $d(CGCGAATTCGCG)_2$ in solution have been carried out by ^{1}H 2D NMR (Hare et al., 1983), but the severe signal overlap hampered determination of all the coupling constants relevant to the sugar conformation. By virtue of the ^{13}C labels in the deoxyribose ring of the thymidine, many of the recently developed heteronuclear multidimensional NMR techniques can be used, and thus a detailed conformational analysis of this dodecamer is possible. In this communication, an application of the HCCH-E.COSY experiment developed by Griesinger and Eggenberger (1992) is described. The correlation peaks observed in the 2D HCCH-E.COSY spectrum, which was measured within 3.8 h using 160 µl of a 1.4 mM solution, are shown in Fig. 2. All of the vicinal coupling constants could be measured accurately from the spectrum. For example, $^{3}J_{H1',H2'}$ and $^{3}J_{H1',H2''}$ were measured from the $H_{2'}$-$C_{1'}$ and $H_{2''}$-$C_{1'}$ cross peaks, respectively, as shown in Fig. 2a. In the case of the $C_{2'}$ methylene group (Fig. 2e), two values, ($^{3}J_{H1',H2'} + {}^{3}J_{H1',H2''}$) and ($^{3}J_{H3',H2'} + {}^{3}J_{H3',H2''}$), were obtained from the two cross peaks $C_{2'}$-$H_{1'}$ and $C_{2'}$-$H_{3'}$, respectively. The accuracy of the observed coupling constants, which are listed in Table 1, could be confirmed by comparing the coupling constants observed for the $C_{2'}$ and $C_{5'}$ methylene groups with those calculated by adding the values measured independently from the correlation peaks of the relevant methines (see footnote c of Table 1).

The preferred conformation of the sugar moiety of the T7 residue in the dodecamer could be estimated by means of the relevant vicinal ^{1}H-^{1}H coupling constants, as described by Rinkel and Altona (1987). The population of the S-conformer is approximately 95%, with a pseudorotation phase angle of about 120°, indicating that the T7 deoxyribose ring adopts a nearly $C_{1'}$-exo conformation. The estimated sugar conformation for the T7 residue was slightly different from the $C_{2'}$-endo conformation typical for B-DNA, but it was similar to the corresponding sugar conformations determined by X-ray analysis, which were $O_{4'}$-endo and $C_{1'}$-exo for T7 and T19 (in the opposite strand), respectively (Dickerson and Drew, 1981; Drew et al., 1981). Detailed conformational analyses of the sugar moieties of the dodecamer will be published elsewhere.

With the recent progress in computer-assisted analysis of homonuclear ^{1}H 2D NMR spectra, most of the vicinal coupling constants in the sugar rings of small DNA oligomers can be estimated with higher precision than before (Wüthrich, 1986; Schmitz et al., 1992). However, the values of $^{3}J_{H3',H4'}$ and $^{3}J_{H4',H5'/H5''}$, which are relevant to determine the sugar–phosphate backbone conformation, cannot usually be obtained by ^{1}H NMR due to severe overlap of these resonances. The accuracy of the estimated spin coupling constants would also be limited, especially for larger oligomers, in which larger line widths make accurate measurements of the peak positions difficult, even if the assignments are firmly established. The ^{1}H-^{1}H coupling constants measured by the HCCH-E.COSY experiment, however, will be accurate, even for larger DNAs, since they are

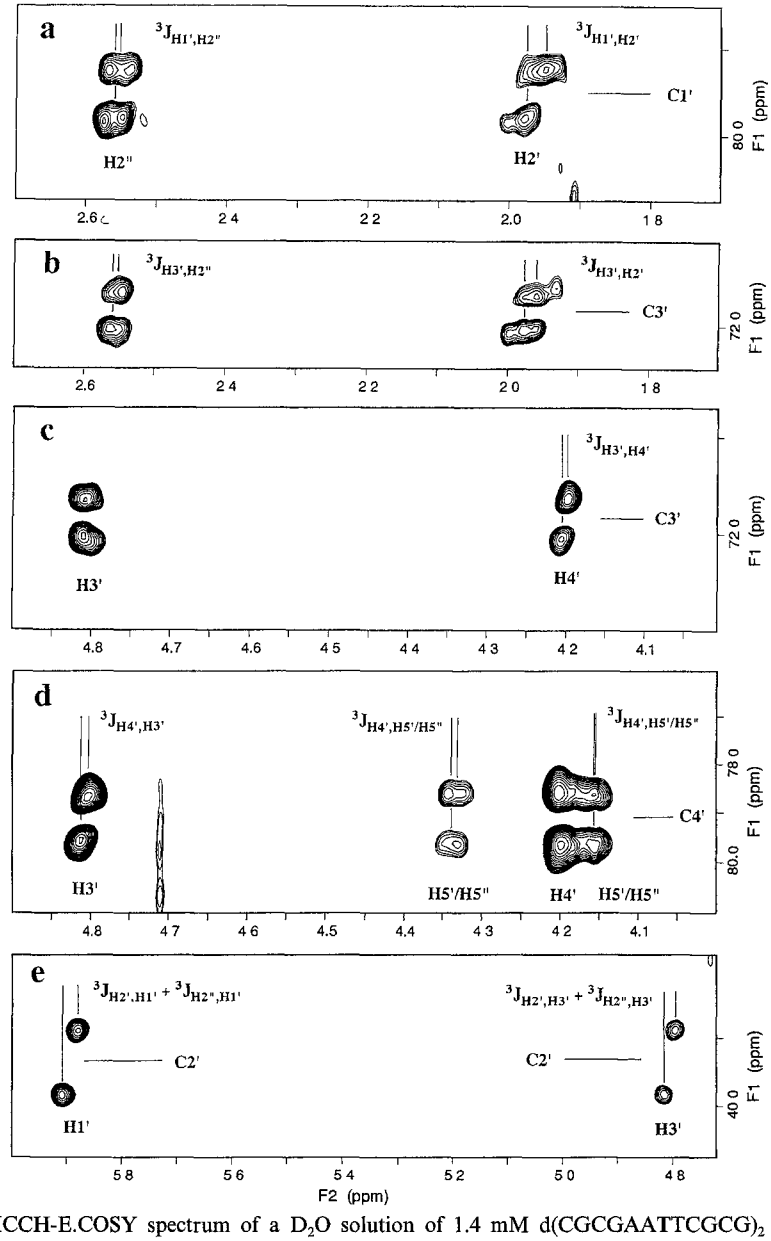

Fig. 2. The 2D HCCH-E.COSY spectrum of a D_2O solution of 1.4 mM d(CGCGAATTCGCG)$_2$ (Griesinger and Eggenberger, 1992). The sample and spectrometer conditions, unless otherwise indicated, were identical to those for the CT ^{1}H-^{13}C HSQC experiment. The ^{13}C and ^{1}H rf carrier frequencies were set at 61.0 and 4.71 ppm (water resonance), respectively. Durations of 1.5 and 6.6 ms were used for the INEPT transfer and for the total constant-time component, respectively. In order to refocus the ^{13}C-^{13}C antiphase term, a 3.3 ms delay was applied after the constant-time duration, and a 1.1 ms delay was used prior to the small flip ^{1}H pulse (30°) to measure both the methine and methylene ^{1}H signals. The acquired data matrix was 50 (t_1) × 512 (t_2) complex data points with 128 transients per t_1 increment, and the total measurement took about 3.8 h. The final data set, after zero-filling, was 512 (F_1) × 8192 (F_2) real points with a digital resolution of 0.55 Hz/point in the F_2 dimension. Data processing and measurements of the exact cross-peak positions were done by FELIX. Selected parts of the obtained spectrum are shown: (a) $C_{1'}$-$H_{2',2''}$; (b) $C_{3'}$-$H_{2',2''}$; (c) $C_{3'}$-$H_{3',4'}$; (d) $C_{4'}$-$H_{5',5''}$ and $C_{4'}$-$H_{3'}$; and (e) $C_{2'}$-$H_{1',3'}$.

TABLE 1
VICINAL ^{1}H-^{1}H SPIN COUPLING CONSTANTS OF THE T7 RESIDUE IN d(CGCGAATTCGCG)$_2$ AS DETERMINED FROM THE 2D HCCH-E.COSY SPECTRUM

$^3J_{H_iH_j}$	Observed value[a] (Hz)	Cross peak[b]	$^3J_{H_iH_j}$	Observed value[a] (Hz)	Cross peak[b]
$J_{1'2'}$	11.3	$C_{1'}$-$H_{2'}$	$(J_{3'2'} + J_{3'2''})$	9.2	$C_{2'}$-$H_{3'}$
$J_{1'2''}$	4.4	$C_{1'}$-$H_{2''}$	*$(J_{3'2'} + J_{3'2''})$*	*9.7*	
$(J_{1'2'} + J_{1'2''})$	15.2	$C_{2'}$-$H_{1'}$	$J_{3'4'}$	4.2	$C_{3'}$-$H_{4'}$
$(J_{1'2'} + J_{1'2''})$	*15.5*[c]		$J_{4'5'(5'')}$	3.0	$C_{4'}$-$H_{5'}$($H_{5''}$)
$J_{2'2''}$	14.5	$C_{2'}$-$H_{2'}$($H_{2''}$)	$J_{4'5''(5')}$	<0.6	$C_{4'}$-$H_{5''}$($H_{5'}$)
$J_{3'2'}$	7.3	$C_{3'}$-$H_{2'}$	$(J_{4'5'} + J_{4'5''})$	3.4	$C_{5'}$-$H_{4'}$
$J_{3'2''}$	2.4	$C_{3'}$-$H_{2''}$	*$(J_{4'5'} + J_{4'5''})$*	*<3.6*	

[a] Coupling constants were measured directly with the cursor in the FELIX window and contain a deviation of ± 0.6 Hz.
[b] These cross peaks were used to determine the corresponding vicinal ^{1}H-^{1}H coupling constants in the same row.
[c] The values given in italics were calculated from the spin coupling constants observed for the cross peaks between the methine carbon and the methylene protons. These values were identical, within error (± 0.6 Hz), to the observed spin coupling constants for the cross peaks between the methylene carbon and the methine proton.

measured by the displacements of the cross peaks that are split by the large ^{1}H-^{13}C single-bond couplings.

The amount of the labeled dodecamer used for the heteronuclear 2D NMR experiments described in this communication was about 0.22 μmole, which is equivalent to 1.4 mM in 160 μl solution. This amount is in fact less than that obtained by a single 1-μmole-scale solid-phase synthesis, which provides about 0.24 μmole (29 O.D. at 254 nm) after purification. Therefore, in the case of the labeled thymidine, only 25 mg of glucose (98% ^{13}C) per thymidine site was required for each NMR sample, making this approach economically very feasible.

Finally, let us point out the prospects for the synthesis of isotopically labeled deoxynucleosides other than thymidine, namely deoxycytidine, deoxyadenosine and deoxyguanosine. Uracil and adenine derivatives can be coupled with the labeled α-chlorosugar as well, although the yields for purine bases are usually low (Hubbard et al., 1984; Kawakami et al., 1989). Since deoxyuridine can be converted chemically to deoxycytidine in a good yield (Kellenbach et al., 1991), isotopically labeled deoxycytidine can be prepared in a similar manner as described for thymidine. The method, however, is not appropriate for preparing labeled deoxyadenosine and deoxyguanosine. A better alternative seems to be the chemical conversion of ribonucleosides to deoxyribonucleosides. We have recently succeeded in synthesizing highly stereoselective (2'*R*)- [2',^{2}H]-2'-deoxyribonucleosides by this method (Kawashima et al., 1993). If labeled purine ribonucleosides can be produced at reasonable cost, this procedure will be suitable for synthesizing labeled deoxyadenosine and deoxyguanosine. Some experiments aimed toward this goal will be reported shortly (Tate et al., 1994).

ACKNOWLEDGEMENTS

This work was supported in part by the special coordination fund of the Science and Technology Agency, by the Grant-in-Aid for Specially Promoted Research of the Ministry of Education (No. 05101004), and by a grant from the Human Frontier Science Program (Strasbourg, France).

We thank Mr. Hiroyuki Kyomori of Kobayashi Perfumery Ltd. for his generous help in preparing the ^{13}C-labeled thymidine.

REFERENCES

Ashcroft, J., LaPlante, S.R., Borer, P.N. and Cowburn, D. (1989) *J. Am. Chem. Soc.*, **111**, 363–365.
Ashcroft, J., Live, D.H., Patel, D.J. and Cowburn, D. (1991) *Biopolymers*, **31**, 45–55.
Batey, R.T., Inada, M., Kujawinski, E., Puglisi, J.D. and Williamson, J.R. (1992) *Nucleic Acids Res.*, **20**, 4515–4523.
Beaucage, S.L. and Caruthers, M.H. (1981) *Tetrahedron Lett.*, **22**, 1859–1862.
Clore, G.M. and Gronenborn, A.M. (1991) *Science*, **252**, 1390–1399.
Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.*, **149**, 761–786.
Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2179–2183.
Farmer II, B.T., Müller, L., Nikonowicz, E.P. and Pardi, A. (1993) *J. Am. Chem. Soc.*, **115**, 11040–11041.
Farmer II, B.T., Müller, L., Nikonowicz, E.P. and Pardi, A. (1994) *J. Biomol. NMR*, **4**, 129–133.
Griesinger, C. and Eggenberger, U. (1992) *J. Magn. Reson.*, **97**, 426–434.
Hardegger, E. (1962) *Methods Carbohydr. Chem.*, **1**, 177–179.
Hare, D.R., Wemmer, D.E., Chou, S.-H. and Drobny, G. (1983) *J. Mol. Biol.*, **171**, 319–336.
Hoffer, M. (1960) *Chem. Ber.*, **93**, 2777–2781.
Hubbard, A.J., Jones, A.S. and Walker, R.T. (1984) *Nucleic Acids Res.*, **12**, 6827–6837.
Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.
Kawakami, H., Matsushita, H., Shibagaki, M., Naoi, I., Ito, K. and Yoshikoshi, H. (1989) *Chem. Lett.*, 1365–1368.
Kawashima, E., Aoyama, Y., Sekine, T., Nakamura, E., Kainosho, M., Kyogoku, Y. and Ishido, Y. (1993) *Tetrahedron Lett.*, **34**, 1317–1320.
Kellenbach, E.R., Van den Elst, H., Boelens, R., Van der Marel, G.A., Van Boom, J.H. and Kaptein, R. (1991) *Recl. Trav Chim. Pays-Bas*, **110**, 387–388.
Kellenbach, E.R., Remerowski, M.L., Eib, D., Boelens, R., Van der Marel, G.A., Van den Elst, H., Van Boom, J.H. and Kaptein, R. (1992) *Nucleic Acids Res.*, **20**, 653–657.
Lancelot, G., Chanteloup, L., Beau, J.-M. and Thung, N.T. (1993) *J. Am. Chem. Soc.*, **115**, 1599–1600.
Michinicka, M.J., Harper, J.W. and King, G.C. (1993) *Biochemistry*, **32**, 395–400.
Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) *Nucleic Acids Res.*, **15**, 8783–8789.
Nikonowicz, E.P. and Pardi, A. (1992a) *J. Am. Chem. Soc.*, **114**, 1082–1083.
Nikonowicz, E.P. and Pardi, A. (1992b) *Nature*, **355**, 184–186.
Nikonowicz, E.P., Sirr, A., Legault, P., Jucker, F.M., Baer, L.M. and Pardi, A. (1992) *Nucleic Acids Res.*, **20**, 4507–4513.
Nikonowicz, E.P. and Pardi, A. (1993) *J. Mol. Biol.*, **232**, 1141–1156.
Pardi, A. and Nikonowicz, E.P. (1992) *J. Am. Chem. Soc.*, **114**, 9202–9203.
Rinkel, L.J. and Altona, C. (1987) *J. Biomol. Struct. Dyn.*, **4**, 621–649.
Robins, M.J., Wilson, J.S. and Hansske, F. (1983) *J. Am. Chem. Soc.*, **105**, 4059–4065.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York, NY.
Santoro, J. and King, G.C. (1992) *J. Magn. Reson.*, **97**, 202–207.
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.
Schmitz, U., Sethson, I., Egan, W.M. and James, T.L. (1992) *J. Mol. Biol.*, **227**, 510–531.
Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.
Sinha, N.D., Biernat, J. and Köster, H. (1983) Tetrahedron Lett., **24**, 5843–5846.
Sklenar, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1993a) *J. Biomol. NMR*, **3**, 721–727.
Sklenar, V., Peterson, R.D., Rejante, M.R., Wang, E. and Feigon, J. (1993b) *J. Am. Chem. Soc.*, **115**, 12181–12182.
Sklenar, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1994) *J. Biomol. NMR*, **4**, 117–122.
Tate, S., Ono, A. and Kainosho, M. (1994) *J. Am. Chem. Soc.*, **116**, in press.
Van de Ven, F.J.M. and Phillippens, M.E.P. (1992) *J. Magn. Reson.*, **97**, 637–644.
Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.
Williamson, J.R. and Boxer, S.G. (1988) *Nucleic Acids Res.*, **16**, 1529–1540.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.