Journal of Biomolecular NMR, 4 (1994) 575–580
ESCOM



J-Bio NMR 196

# Selectively $^{13}C$-enriched DNA: $^{13}C$ and $^{1}H$ assignments of a triple helix by two-dimensional relayed HMQC experiments

Olivier Bornet[a], Gérard Lancelot[a,*], Luc Chanteloup[b], Nguyen T. Thuong[a] and Jean-Marie Beau[b]

[a]*Centre de Biophysique Moléculaire, CNRS, 1A Avenue de la Recherche Scientifique, F-45071 Orléans Cedex 2, France*
[b]*Université d'Orléans, Laboratoire de Biochimie Structurale associé au CNRS, BP 6759, F-45067 Orléans Cedex 2, France*

Received 17 January 1994
Accepted 18 February 1994

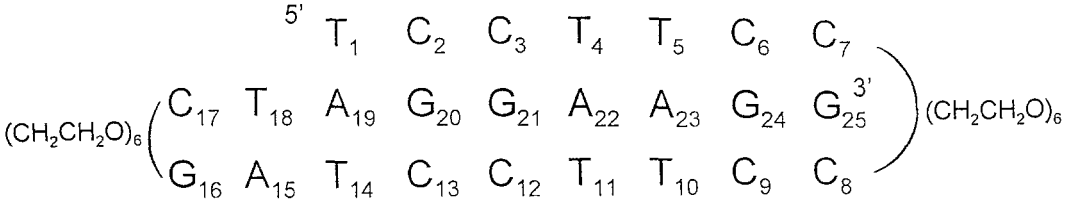
*Keywords:* $^{13}C$-labeled oligonucleotides; Triple helix

## SUMMARY

We present NMR studies of an intramolecular triple helix, the three strands of which have been linked by a hexaethylene glycol chain. To overcome the generally encountered difficulties of assignment in the homopyrimidine strands, the carbon C1' of the pyrimidines were selectively $^{13}C$-enriched. Assignments of the aromatic and sugar protons were obtained from NOESY-HMQC and TOCSY-HMQC spectra. We show that the recognition of a DNA duplex by a third strand via triplex formation is easily carried out in solution by observing the changes of the $^{1}H1'$-$^{13}C1'$ connectivities as a function of pH. Furthermore, the conformation of the sugars has been found to be C2'-endo, on the basis of the coupling constant values directly measured in an HSQC spectrum.

Besides the various double-stranded conformations, multistranded helices, such as triple and quadruple strands, have been drawing increasing attention (Cheong and Moore, 1992; Kang et al., 1992; Radhakrishnan et al., 1992a; Smith and Feigon, 1992; Wang and Patel, 1992; Gehring et al., 1993; Koshlap et al., 1993). Triple-stranded DNA structures have, for example, attracted renewed interest because of their possible occurrence in vivo (Wells et al., 1988). Moreover, the use of oligonucleotides as sequence-specific major groove binders, carrying chemical agents capable of modifying or cleaving DNA duplexes, is currently evaluated for gene regulation or development of artificial nucleases (Dervan, 1986; Thuong and Hélène, 1993). However, only a small number of structures have been elucidated in solution (Macaya et al., 1992; Radhakrishnan and Patel, 1992b,1993). The first structural characterization of triple-stranded DNA was carried out by Arnott and Selsing (1974) by means of fiber diffraction studies. The large number of bases per helical turn led these authors to suggest that poly(dT)·poly(dA)·poly(dT) forms a triple-helical structure with an A'-DNA conformation, in which all sugar residues display the C3'-endo

*To whom correspondence should be addressed.



Scheme 1. Structure of the triple helix.

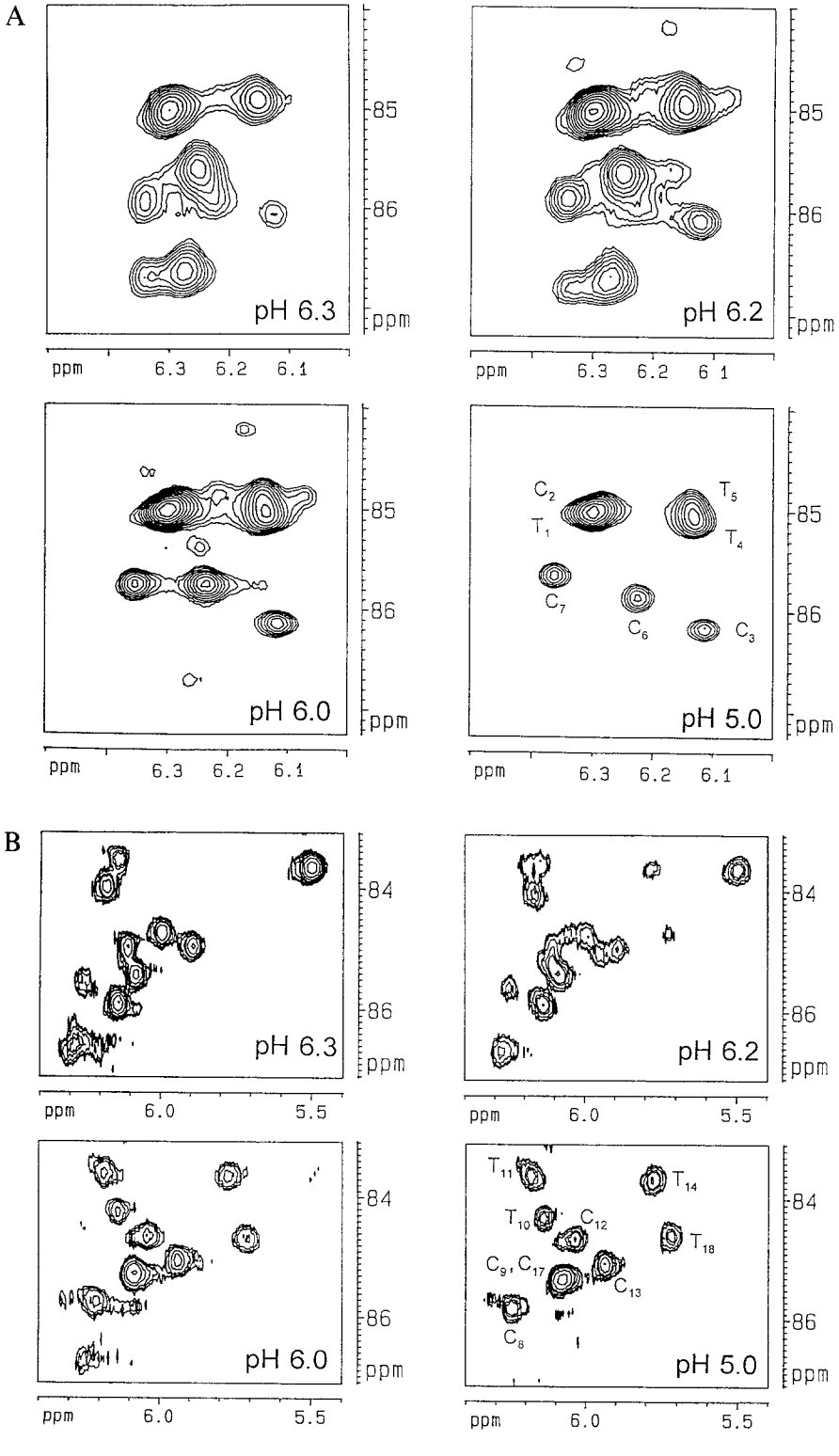
conformation. Several years later, NMR studies confirmed this C3'-endo sugar conformation (Rajagopal and Feigon, 1989a,b; Umemoto et al., 1990). More recently, vibrational spectroscopy studies demonstrated that the 2-deoxyribosyl moieties adopt an S-type conformation (Liquier et al., 1991) in poly(dT)·poly(dA)·poly(dT) and this was confirmed by NMR studies (Macaya et al., 1992). The severe overlap of the proton resonance lines occurring in triple-stranded deoxyoligonucleotides is largely responsible for these conflicting results. The same situation was found for large proteins and their solution structure determination has been greatly aided by multidimensional heteronuclear NMR, using $^{13}C$- and $^{15}N$-labeled material. Although labeled proteins are relatively easy to obtain by biochemical methods, deoxyoligonucleotides have never been prepared in this way. Only a very limited number of studies on labeled ribooligonucleotides have been published recently (Nikonowicz et al., 1992; Nikonowicz and Pardi, 1992a,b).To circumvent the overlap difficulties, we have synthesized deoxyoligonucleotides, $^{13}C$-labeled in position C1' (Chanteloup and Beau, 1992; Lancelot et al., 1993). We present here the assignment of a selectively $^{13}C$-labeled intramolecular triple-stranded helix.

In order to assign all the nonexchangeable resonance lines of the intramolecular triple helix 5'd-(TCCTTCC---CCTTCCTAG---CTAGGAAGG), where --- represents a hexaethylene glycol chain (Scheme 1), two oligonucleotides, 100% selectively $^{13}C$-enriched at position C1', were synthesized. Oligonucleotide I was $^{13}C$-labeled in pyrimidine nucleotides T1 to C7, corresponding to the third strand, and oligonucleotide II was $^{13}C$-labeled in pyrimidine nucleotides C8 to T14, C17 and T18, belonging to the duplex of the assumed triple helix.

Figure 1 shows the HMQC spectra of both oligonucleotides at different pH. The chemical shift range of the H1' and C1' resonances results in an excellent dispersion of the $^{1}H1'$-$^{13}C1'$ correlations of

→

Fig. 1. Plots of HMQC spectra of the oligonucleotides under investigation, showing the transition between the duplex and the triplex as a function of pH. Oligonucleotides I (A, see text) and II (B, see text) were initially observed at pH 7.5, and the pH was subsequently decreased by adding appropriate amounts of 0.1 M DCl. Between pH values 6.3 and 6.0, the $^{1}H$-$^{13}C$ correlation maps were strongly modified as a result of a significant slow exchange between the intramolecular duplex and the intramolecular triplex. The [1'-$^{13}C$]-dT and -dC were prepared by *N*-glycosylation (Vorbrüggen et al., 1981) (with trimethylsilyltrifluoromethanesulfonate as a promoter) of the silylated nucleobases with [1'-$^{13}C$]-phenylsulfinyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranoside, available in 77% overall yield from commercial [1'-$^{13}C$]-D-ribose (99% $^{13}C$-enriched, Centre d'Etudes Nucléaires, Saclay) (Chanteloup and Beau, 1992). The labeled nucleosides were further deoxygenated at C2' (Robins et al., 1983) and transformed to 5'-*O*-dimethoxytrityl-3'-*O*-(β-cyanoethyl-*N,N*-diisopropylphosphoramidite) building blocks (Sinha et al., 1983; Chanteloup and Beau, 1992). The required $^{13}C1'$-labeled oligodeoxynucleotides I and II were prepared on a Pharmacia automatic synthesizer via phosphoramidite chemistry (Caruther, 1987). After deprotection, the oligodeoxynucleotides were purified by anion exchange chromatography on a mono Q column (Pharmacia) and then analyzed by reversed-phase HPLC (Lichrospher RP18, Merck).


A
pH 6.3
pH 6.2
pH 6.0
pH 5.0
C2
T1
T5
T4
C7
C6
C3
ppm
6.3
6.2
6.1
85
86
B
pH 6.3
pH 6.2
pH 6.0
pH 5.0
T11
T14
T10
C12
T18
C9, C17
C13
C8
6.0
5.5
84

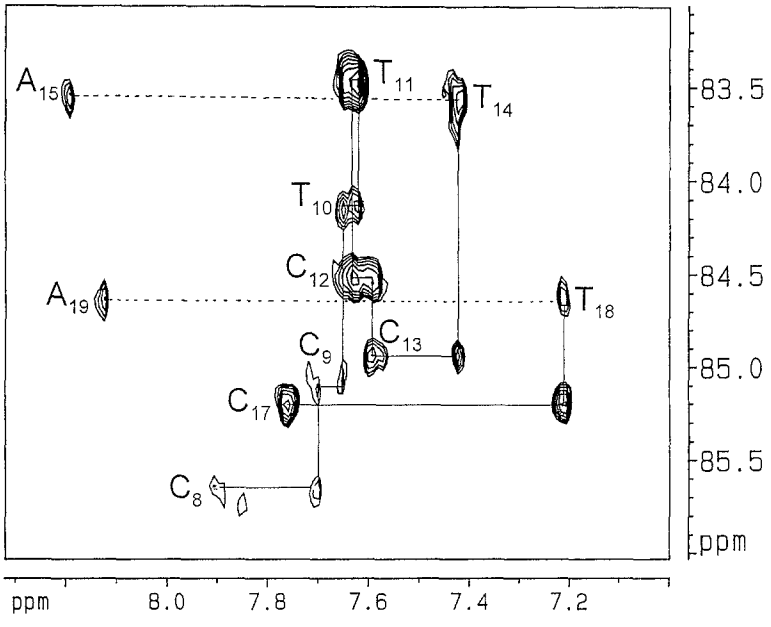



Fig. 2. Plot of the aromatic proton region of a 500 MHz 2D pure phase absorption $^{13}$C1'-$^{1}$H relayed HMQC-NOESY spectrum of the $^{13}$C1'-labeled oligonucleotide II. In a right-handed DNA helix, the H1' sugar proton of residue i is close to the H8 or H6 nucleobase protons of the 3'-neighboring residue (i − 1) and far from all sugar protons of its 5'-neighboring nucleotide (i + 1). The C1'(i)-H6(i) and C1'(i)-H6(i + 1) connectivities for the fragments 5'd-(C17–T18) and 5'-d-(C8–T14) are shown. The sequential assignment of the C1' carbons and aromatic protons is shown by solid lines. A15 and A19 are not $^{13}$C-labeled and present the connectivities C1'(T14)-H8(A15) and C1'(T18)-H8(A19). The 2D spectrum was acquired on a Bruker AMX 500 spectrometer using a 1 mM sample, 0.1 mM NaCl in $D_2O$ at 20 °C. The heteronuclear HMQC-NOESY sequence was recorded with 1024 complex data points in the $t_2$ domain, with 40 TPPI increments, 2048 scans per increment and four dummy scans. The delay 1/2J was optimized for a coupling constant of 168 Hz. GARP broad-band $^{13}$C decoupling was applied during the acquisition period in order to remove the $^{13}$C-$^{1}$H scalar coupling. The mixing time was 250 ms. The spectrum was zero-filled in $t_1$ to give a final 1024 × 40 real data matrix and an exponential multiplication was applied in the $t_2$ domain for resolution enhancement.

the two oligonucleotides, thus avoiding the tedious observation of the intrastrand NOE connectivities between adjacent nucleotides (aromatic protons to 5'-neighbor sugar H1' protons). Only a few correlations were overlapped, despite the homopyrimidine nature of the labeled strands. The pairing of the pyrimidine-containing third strand requires the protonation of the cytosine residues. Hence, the transition between the double-stranded and triple-stranded forms can be followed in the HMQC spectra (Fig. 1) as a function of pH. The $^{1}$H-$^{13}$C correlation maps remained unaltered by increasing the pH values from 6.3 to 8.0 or by decreasing them from 6.0 to 2.8. A pK value of 6.15 can thus be assumed. This NMR analysis also showed that below a pH value of 2.8, the multistranded form is destroyed, due to the protonation of the 'pyrrolic' nitrogen of the purines.

Assignments of the $^{13}$C and $^{1}$H resonances were made by relayed heteronuclear experiments, derived from the HMQC sequence. It is well known that in DNA the H1' proton of a nucleotide is close (~ 3.6 Å) to its own aromatic H8 (purine bases) or H6 (pyrimidine bases) and to the aromatic protons of its neighboring residue (3'-neighbor in a right-handed helix and 5'-neighbor in a left-handed helix). Based on this property, we used a $^{13}$C-$^{1}$H HMQC-NOESY sequence which

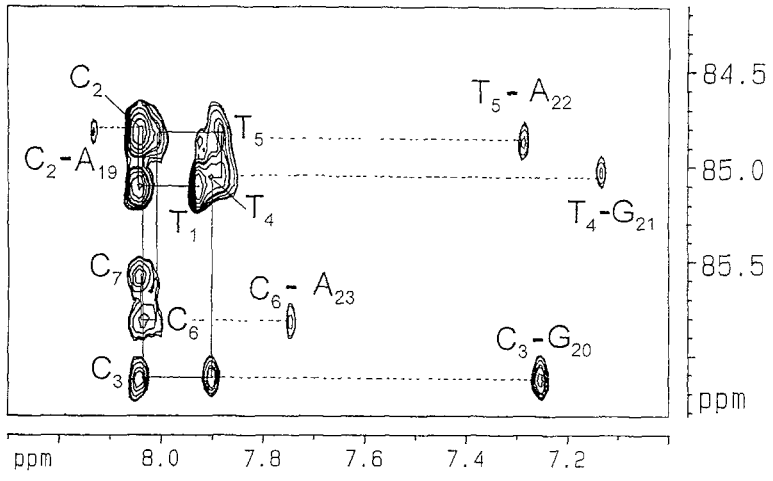



Fig. 3. Plot of the aromatic proton region of a 500 MHz 2D pure phase absorption $^{13}$C1'-$^{1}$H relayed HMQC-NOESY spectrum of the $^{13}$C1'-labeled triplex, using a 2 mM sample of oligonucleotide I, 0.1 M NaCl in $D_2O$ at 22 °C. The C1'(i)-H6(i) and C1'(i)-H6(i + 1) connectivities for the strand 5'd-(T1–C7) are shown. The sequential assignment of the C1' carbons and aromatic protons is shown by solid lines. In addition, a number of NOEs that are not observed in B-DNA are present in this triplex. These interstrand NOEs, characteristic of triple-stranded DNA, are cross peaks (shown in the figure) between purine H8 resonances (or H6 for T18) and the H1' resonance from the 3'-neighboring pyrimidine base in the third strand: C1'(C2)-H8(A19), C1'(C3)-H8(G20), C1'(T4)-H8(G21), C1'(T5)-H8(A22) and C1'(C6)-H8(A23).

transfers the magnetization of the H1' protons to the H8 or H6 protons via their $^{13}$C1' nucleus. Figures 2 and 3 show the resulting sequential assignment of the aromatic protons. For the third strand, this assignment was indeed impossible on a $^{1}$H-$^{1}$H NOESY spectrum, due to the overlap resulting from the weak dispersion of the sugar resonances (H1' as well as H2'/H2"). The triple-stranded structure was found to be a right-handed helix. Once the aromatic proton resonances of the pyrimidine strands had been assigned, examination of the $^{1}$H-$^{1}$H NOESY spectrum provided the assignment of the aromatic protons in the purine strand, whose dispersion was greater due to the larger purine ring current effects. Pairing of the third strand was confirmed by the characteristic resonances of the N3H protons of the protonated cytosines. Their assignment, as well as those of all the exchangeable protons, was obtained from their sequential connectivities observed in a $^{1}$H-$^{1}$H NOESY spectrum in water. As already noted (Macaya et al., 1992), several uncommon characteristics could be observed in the resonances of the triple helix (Figs. 2 and 3), i.e., numerous aromatic resonances of the homopurine strand are found in the 7.0–7.4 ppm range, while those of the protonated cytosines lie between 7.8 and 8.1 ppm. Each H1' resonance of the pyrimidine residues in the third strand was correlated with the aromatic protons of the 5'-neighbor of its pairing purine (T18–G25 segment). Connectivities between the imino protons of the third-strand pyrimidine residues and the H2'/H2" protons of the 5'-neighbor of the pairing purine base (T18–G25), constituting another set of interstrand NOEs characteristic of triple-helical DNA, were also observed. Once the aromatic and H1' resonances were located, examination of the through-bond J coupling correlations in the HMQC-TOCSY experiment gave the H2', H2", H3' and H4' resonances.

The sugar conformations present in a triple helix are also an important structural parameter

defining the overall geometry. Unfortunately, as often encountered in NOESY or COSY triple-helix maps, the H1'-H2'/H2" regions were too crowded to measure the coupling constants. Taking advantage of the dispersion of the H1' and C1' resonances, we circumvented this difficulty by running HSQC spectra. This experiment has the advantage over the parent HMQC sequence of producing singlets in the F1 dimension (Norwood et al., 1990). It was therefore easy to directly measure the H1'-H2' and H1'-H2" coupling constants in the F2 sections of the HSQC map. The measured coupling constant values $J_{1',2'}$ = 10.1 Hz and $J_{1',2''}$ = 5.5 Hz) show that the sugars of the third strand adopt a C2'-endo conformation.

The good dispersion of the $^{1}$H1'-$^{13}$C1' correlations is favorable for studying oligonucleotides where the overlap of the H1' resonances bars the complete assignment of, for example, the homopyrimidine strand. Complete (100%) $^{13}$C-selective enrichment offers the advantage of sensitivity, without complicating the spectrum by the $^{13}$C-$^{13}$C coupling constants and the analysis of T1 or NOE data by additional $^{13}$C-$^{13}$C relaxation pathways.

## ACKNOWLEDGEMENTS

We are grateful to the Ministère de l'Enseignement Supérieur de la Recherche (O.B.) and the Ligue Nationale Française contre le Cancer (L.C.) for predoctoral fellowships.

## REFERENCES

Arnott, S. and Selsing, E. (1974) *J. Mol. Biol.*, **88**, 509–521.
Caruther, M.H. (1987) In *Synthesis and Applications of DNA and RNA* (Ed, Narang, S.A.) Academic Press, London, pp. 47–77.
Chanteloup, L. and Beau, J.-M. (1992) *Tetrahedron Lett.*, **33**, 5347–5350.
Cheong, C. and Moore, P.B. (1992) *Biochemistry*, **31**, 8406–8414.
Dervan, P.B. (1986) *Science*, **232**, 464–471.
Gehring, K., Leroy, J.-L. and Guéron, M. (1993) *Nature*, **363**, 561–565.
Kang, C., Zhang, X., Ratliff, R., Moyzis, R. and Rich, A. (1992) *Nature*, **356**, 126–131.
Koshlap, K.M., Gillespie, P., Dervan, P.B. and Feigon, J. (1993) *J. Am. Chem. Soc.*, **115**, 7908–7909.
Lancelot, G., Chanteloup, L., Beau, J.-M. and Thuong, N.T. (1993) *J. Am. Chem. Soc.*, **115**, 1599–1600.
Liquier, J., Coffinier, P., Firon, M. and Taillandier, E. (1991) *J. Biomol. Struct. Dyn.*, **9**, 437–445.
Macaya, R., Wang, E., Schultze, P., Sklenar, V. and Feigon, J. (1992) *J. Mol. Biol.*, **225**, 755–773.
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.
Norwood, T.J., Boyd, J., Heritage, J.E., Soffe, N. and Campbell, I.D. (1990) *J. Magn. Reson* , **87**, 488–501.
Radhakrishnan, I., Patel, D.J. and Gao, X. (1992a) *Biochemistry*, **31**, 2514–2523.
Radhakrishnan, I. and Patel, D.J. (1992b) *J. Am. Chem. Soc.*, **114**, 6913–6915.
Radhakrishnan, I. and Patel, D.J. (1993) *J Am. Chem. Soc.*, **115**, 1615–1617.
Rajagopal, P. and Feigon, J. (1989a) *Biochemistry*, **28**, 7859–7870.
Rajagopal, P. and Feigon, J. (1989b) *Nature*, **339**, 637–640.
Robins, M.J., Wilson, J.S. and Hanske, F. (1983) *J. Am. Chem. Soc.*, **105**, 4059–4065.
Sinha, N.D., Biernat, J. and Köster, H. (1983) *Tetrahedron Lett.*, **24**, 5843–5846.
Smith, F.W. and Feigon, J. (1992) *Nature*, **356**, 164–168.
Thuong, N.T. and Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl.*, **32**, 666–690.
Umemoto, K., Sarma, M.H., Gupta, G., Luo, J. and Sarma, R.H. (1990) *J. Am. Chem. Soc.*, **112**, 4539–4545.
Vorbrüggen, H., Krolikiewicz, K. and Bennua, B. (1981) *Chem. Ber.*, **114**, 1235–1255.
Wang, Y. and Patel, D.J. (1992) *Biochemistry*, **31**, 8112–8119.
Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M. and Wohlrab, F. (1988) *FASEB J.*, **2**, 2939–2949.