*Carbohydrate Research*, 152 (1986) 63–72
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

# COMPLETE $^1$H-N.M.R. ASSIGNMENTS FOR TWO CORE-REGION OLIGOSACCHARIDES OF HUMAN MECONIUM GLYCOPROTEINS, USING 1D AND 2D METHODS AT 500 MHz

JAMES FEENEY, THOMAS A. FRENKIEL,

*MRC Biomedical NMR Centre, National Institute for Medical Research, Mill Hill, London, NW7 1AA (Great Britain)*

AND ELIZABETH F. HOUNSELL

*Applied Immunochemistry Research Group, MRC Clinical Research Centre, Watford Road, Harrow, Middlesex, HA1 3UJ (Great Britain)*

(Received October 5th, 1985; accepted for publication, November 15th, 1985)

## ABSTRACT

The complete $^1$H-n.m.r. assignments for α-D-GalNAc-(1→3)-D-GalNAc-ol and β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-D-GalNAc-ol have been made using a combination of 2D correlation experiments (COSY, RELAYED-COSY, and $F_1$-decoupled) and an analysis of the high-resolution 1D-n.m.r. spectra at 500 MHz.

## INTRODUCTION

Several glycoproteins from human meconium express differentiation and tumour-associated antigens recognised by natural and hybridoma-derived monoclonal antibodies[1–6]. We have recently reported the structural analysis of the *O*-glycosylically linked core-region oligosaccharides of human meconium glycoproteins which express oncofoetal antigens[7]. The structural analysis was achieved by a combination of techniques which included positive and negative ion f.a.b-m.s. of the oligosaccharides and their methylated derivatives[8–10], e.i.-m.s. of methylated oligosaccharides[11–13], capillary g.l.c.–m.s. of partially methylated alditol acetates[14], g.l.c. analysis, and 500-MHz n.m.r. spectroscopy[15]. Two oligosaccharides, α-D-GalNAc-(1→3)-D-GalNAc-ol and β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-D-GalNAc-ol, for which n.m.r. data had not previously been published, were characterised and partial assignments of their n.m.r. spectra were given. We have now completely assigned both spectra by using a combination of 1D and 2D $^1$H-n.m.r. techniques, namely, homonuclear correlation spectroscopy (COSY)[16–20] (in a double-quantum-filtered-[21]phase-sensitive form[19,20]), $F_1$-decoupled correlation spectroscopy[18,22], and homonuclear relayed correlation spectroscopy[23–26] (RELAYED-COSY). Several workers have already pointed out the value of such techniques for unravelling complex $^1$H-n.m.r. spectra of oligosaccharides[26–29].

The development of n.m.r. spectroscopy as a structural tool for carbohydrate

0008-6215/86/$ 03.50 

studies has relied heavily on measuring the characteristic $^1H$ chemical shifts of the signals of so-called "reporter" groups which give well resolved signals in the spectrum and for the protons to which they are directly coupled[15,30,31]. Additional chemical-shift information from resonances which overlap in the region 3.5–4.0 p.p.m. is now becoming increasingly available, using modern 2D-n.m.r. methods, and this larger data base, together with computerised methods of searching the data[32], will extend the usefulness of the n.m.r. technique as a structural tool. The badly overlapping signals of H-5,6,6′ in Gal residues and of H-3,4 in GlcNAc residues are particularly difficult to assign and a strategy for these assignments has now been developed.

## EXPERIMENTAL

*Isolation and purification of oligosaccharides.* — The isolation and purification of the oligosaccharides has been described[7]. Glycoproteins were extracted from meconium samples of blood-group O neonates of secretor type by digestion with pronase followed by precipitation in aqueous 67% ethanol, and the fraction enriched in Ii antigen was isolated by affinity chromatography. Mild hydrolysis with acid was used to remove the sialic acid and fucose residues from this fraction, and the oligosaccharides were then released by base–borohydride degradation, and purified by gel-permeation chromatography on Bio-Gel P4 and h.p.l.c. on octadecylsilyl and aminopropylsilyl adsorbents[7].

From the combined m.s., n.m.r., and g.l.c. studies, the major oligosaccharides from the meconium glycoproteins were characterised. Two of these, namely, α-D-GalNAc-(1→3)-D-GalNAc-ol and β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-D-GalNAc-ol proved to have novel structures[7] for which n.m.r. data had not previously been reported.

*N.m.r. methods.* — A solution of the oligosaccharides in $D_2O$ was freeze-dried and the process was repeated a further three times in order to replace exchangeable hydrogen with deuterium. The samples were made up as 22 and 6mM solutions for the di- and tri-saccharides, respectively, in $D_2O$ with 2 μM acetone as the internal reference. 500-MHz $^1H$-N.m.r. spectra were obtained using a Bruker AM500 spectrometer, a pulse interval of 8.19 s, and a flip-angle of 65°. In the highest resolution 1D-experiments, a spectral width of 2 kHz was used with 32 K data points, giving a digital resolution of 0.12 Hz/pt. Up to 500 transients were averaged and the free induction decays were multiplied by a Gaussian resolution-enhancement function. Chemical shifts were measured in p.p.m. from the signal for internal acetone and given with reference to DSS (4,4-dimethyl-4-silapentane-1-sulphonate) taken as 2.225 p.p.m. from acetone at 295 K. Spectral simulations were carried out using the PANIC programme which is part of the standard Bruker software.

The 2D experiments were all variants of the original two-pulse correlation experiment (COSY) which allows pairs of spins that have a resolved scalar coupling

to be identified. Two types of signal are important in the experiments involving pairs of coupled spins (cross-peaks) and transitions within a single multiplet (diagonal peaks). Only the first convey useful correlation information. Three variants of the basic experiment were employed. (*a*) A "double-quantum filter"[21] was used together with phase-cycling and data processing to give a phase-sensitive spectrum[19,20]. The combination of these two techniques has the advantage that it gives pure absorption-phase line-shapes for *all* the peaks in the 2D spectrum. (*b*) A simplified correlation spectrum can be obtained by means of a pulse sequence which eliminates all splittings due to proton–proton coupling from one of the dimensions of the 2D spectrum ($F_1$-decoupling)[18,22]; this makes it easier to extract information from crowded spectral regions. (*c*) Introduction of an additional coherence transfer step gives the relayed correlation experiment[23–25]. Extra information is available from this RELAYED-COSY experiment, as cross-peaks can occur between pairs of spins which are not coupled to each other but which are both coupled to a third spin.

## RESULTS AND DISCUSSION

*α-D-GalNAc-(1→3)-D-GalNAc-ol.* — F.a.b.-m.s., e.i.-m.s., and g.l.c.–m.s. had shown[7] this oligosaccharide to be a (1→3)-linked disaccharide containing GalNAc and GalNAc-ol. The 500-MHz $^1$H-n.m.r. spectrum for the ring protons of the disaccharide is given in Fig. 1. The starting point for the assignment of the GalNAc residue is the low-field signal (d, *J* 4.07 Hz) at 5.103 p.p.m. which can be assigned to H-1 of α-GalNAc on the basis of its chemical shift and coupling constant[15]. The signals for H-2,3,4 were then assigned by using sequential spin-decoupling experiments and also from observing connectivities in the 2D-COSY and $F_1$-decoupled COSY spectra. The signal for H-5 (m, 4.073 p.p.m.) was assigned on the basis of its common, small coupling constant to H-4 ($J_{4,5}$ 1.2 Hz) and this then allowed the signals of H-6,6′ to be assigned *via* their connectivities to H-5 seen in the 2D-COSY spectrum. For the GalNAc-ol residue, H-2 and H-5 (4.395 and 3.749 p.p.m., respectively) gave characteristic multiplets resulting from coupling with three vicinal protons. Van Halbeek *et al.*[33] have assigned unambiguously the lower field signal to H-2 by selective deuteration at C-1 (borodeuteride reduction of oligosaccharides released from glycoprotein). The 2D-COSY experiments provide the connectivities of H-2 to the vicinally coupled protons (H-1,1′ and H-3), and these can be assigned by noting that the H-1,1′ pair of nuclei give rise to multiplets featuring a large geminal coupling ($J_{1,1'}$ 11.45 Hz). By following the sequential connectivities involving the assigned H-3 multiplet and those from H-4,5,6,6′ in the 2D-COSY spectrum, all the assignments could be made. The chemical shifts of H-6,6′ are accidentally equivalent and this gives rise to a deceptively simple spectrum in which the doublet splitting for the H-6,6′ signal is 1/2 ($J_{5,6} + J_{5,6'}$); the $J_{6,6'}$ geminal coupling does not feature in such a spectrum.

The assigned multiplets were analysed as first-order spectra (except for the

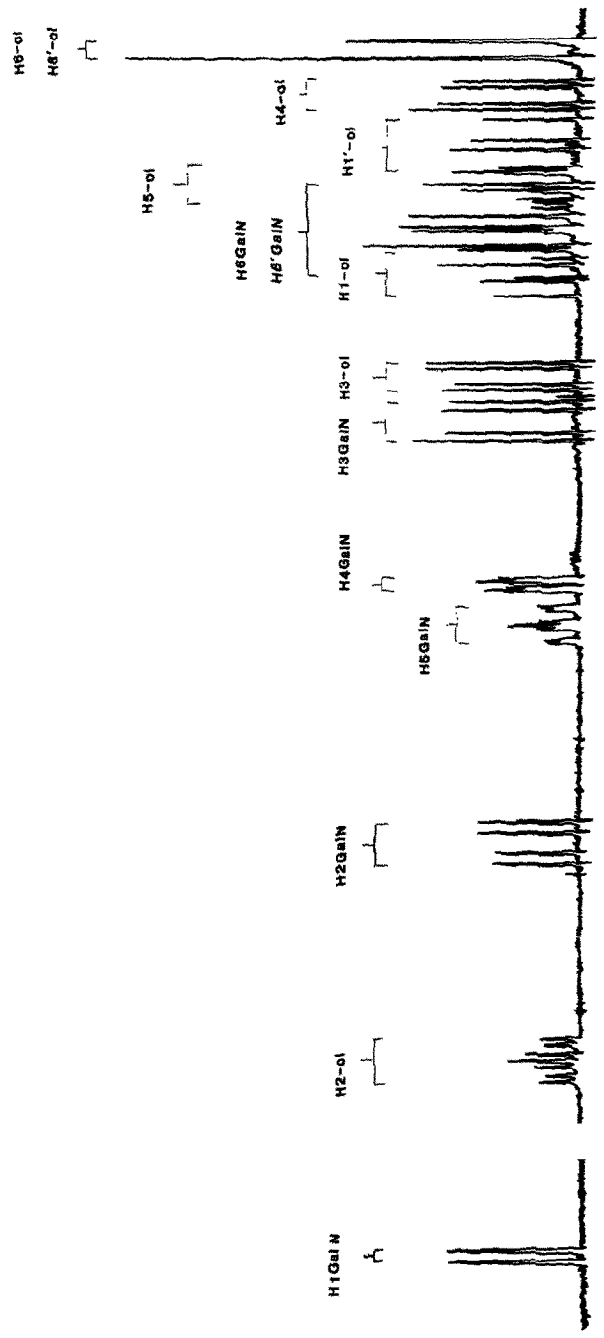


Fig. 1. The resolution-enhanced 500-MHz $^1$H-n.m.r. spectrum of α-GalNAc-(1→3)-GalNAc-ol.

TABLE I

$^1$H-N.M.R. CHEMICAL SHIFTS (p.p.m.) AND COUPLING CONSTANTS (Hz) FOR α-D-GalNAc-(1→3)-D-GalNAc-ol (**1**) AND β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-D-GalNAc-ol (**2**) FROM MECONIUM GLYCOPROTEINS[a]

| | **1** | | **2** | | |
|---|---|---|---|---|---|
| *Chemical shifts* | *GalNAc* | *GalNAc-ol* | *Gal* | *GlcNAc* | *GalNAc-ol* |
| H-1 | 5.103 | 3.810 | 4.471 | 4.576 | 3.728 |
| H-1′ | — | 3.718 | — | — | 3.672 |
| H-2 | 4.235 | 4.395 | 3.538 | 3.772 | 4.247 |
| H-3 | 3.921 | 3.888 | 3.667 | 3.744 | 3.843 |
| H-4 | 4.043 | 3.680 | 3.924 | 3.717 | 3.375 |
| H-5 | 4.073 | 3.749 | 3.715 | 3.600 | 4.027 |
| H-6 | 3.791 | 3.647 | 3.78[b] | 3.994 | 3.932 |
| H-6′ | 3.768 | 3.647 | 3.746 | 3.831 | 3.707 |
| *Coupling constants* | | | | | |
| $J_{1,1'}$ | — | 11.45 | — | — | 11.42 |
| $J_{1,2}$ | 4.07 | 5.60 | 7.80 | 8.22 | 6.13 |
| $J_{1',2}$ | — | 8.22 | — | — | 7.90 |
| $J_{2,3}$ | 11.32 | 2.32 | 9.98 | 10.66 | 1.50 |
| $J_{3,4}$ | 3.11 | 8.00 | 3.42 | — | 9.63 |
| $J_{4,5}$ | 1.20 | 2.08 | 0.89 | 7.5 | 1.42 |
| $J_{5,6}$ | 5.36 | [c] | — | 2.3 | 5.00 |
| $J_{5,6'}$ | 7.00 | [c] | — | 5.2 | 7.50 |
| $J_{6,6'}$ | 11.74 | — | — | 12.27 | 10.65 |

[a]Errors: $J$ ±0.1 Hz; $\delta$ ±0.001. [b]Error: ±0.01 p.p.m. [c]1/2 ($J_{5,6}$ + $J_{5,6'}$) = 6.28.

strongly coupled pairs of geminal protons), and the coupling constants and chemical shifts are given in Table I.

*β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-D-GalNAc-ol.* — F.a.b.-m.s., e.i.-m.s., and g.l.c.–m.s. had shown[7] this fragment to be a trisaccharide containing a terminal non-reducing Gal, 4-linked GlcNAc, and 6-linked GalNAc-ol. In the $^1$H-n.m.r. spectrum of this trisaccharide (Fig. 2), the low-field doublets for H-1 of Gal and GlcNAc and the multiplets of H-2 and H-5 of GalNAc-ol can be assigned readily on the basis of their multiplicities and chemical shifts[15,33]. The phase-sensitive, double-quantum-filtered COSY spectrum (Fig. 3) allowed these signals to be connected with those from their directly coupled proton neighbours, leading immediately to assignments for Gal H-2, GlcNAc H-2, and all the GalNAc-ol protons.

For the Gal residue, the presence of cross-peaks (Fig. 3) correlating Gal H-2,3 and H-3,4 allowed Gal H-3,4 to be assigned; no cross-peak between H-4 and H-5 could be detected. Examination of the multiplet for Gal H-4 in the resolution-enhanced 1D spectrum (Fig. 2) reveals $J_{4,5}$ to be very small (0.89 Hz), and a search for the presence of this splitting in the region 3.64–3.82 p.p.m. revealed the

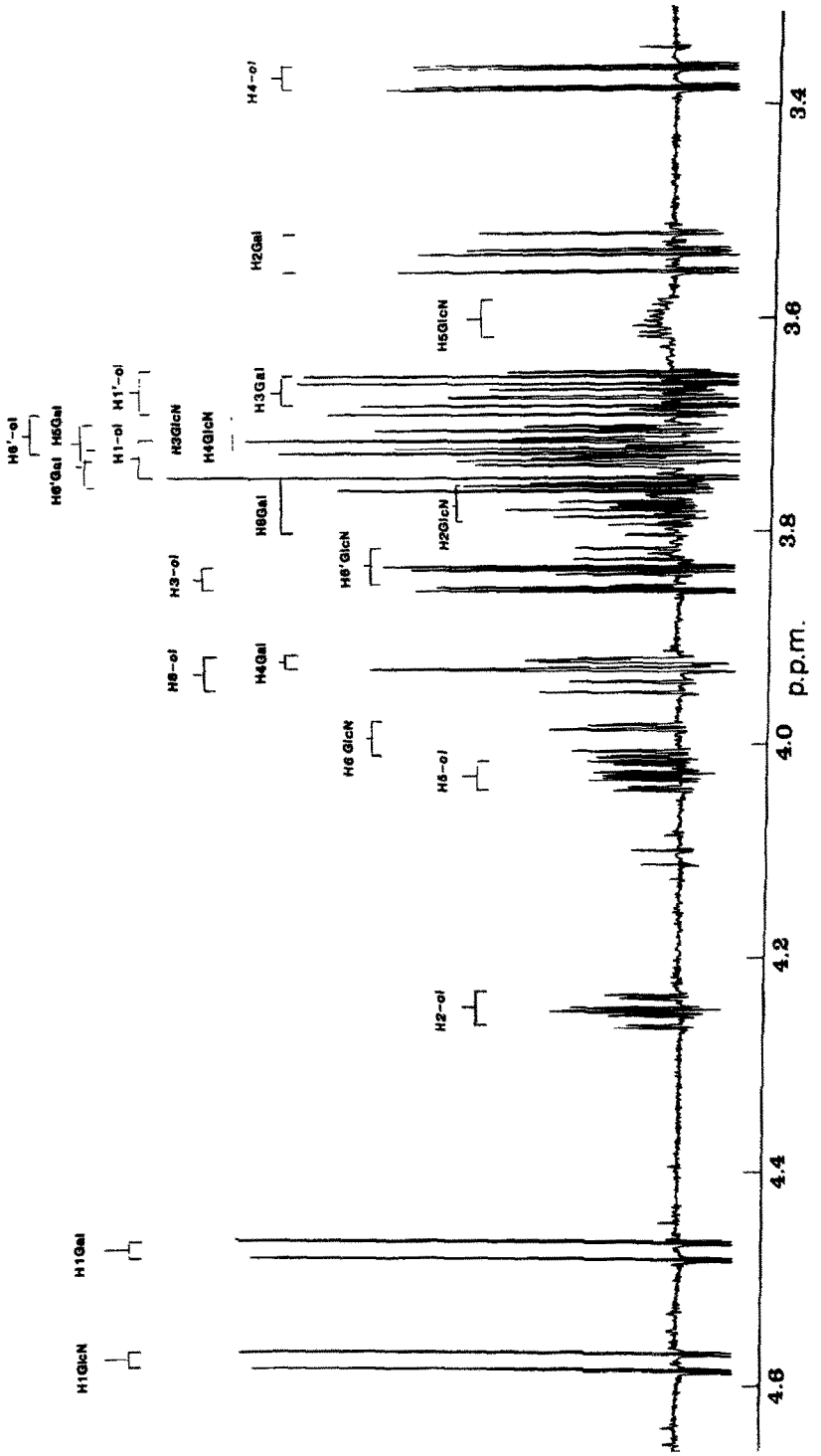


Fig. 2. The resolution-enhanced 500-MHz $^1$H-n.m.r. spectrum of *β*-Gal-(1→4)-*β*-GlcNAc-(1→6)-GalNAc-ol.

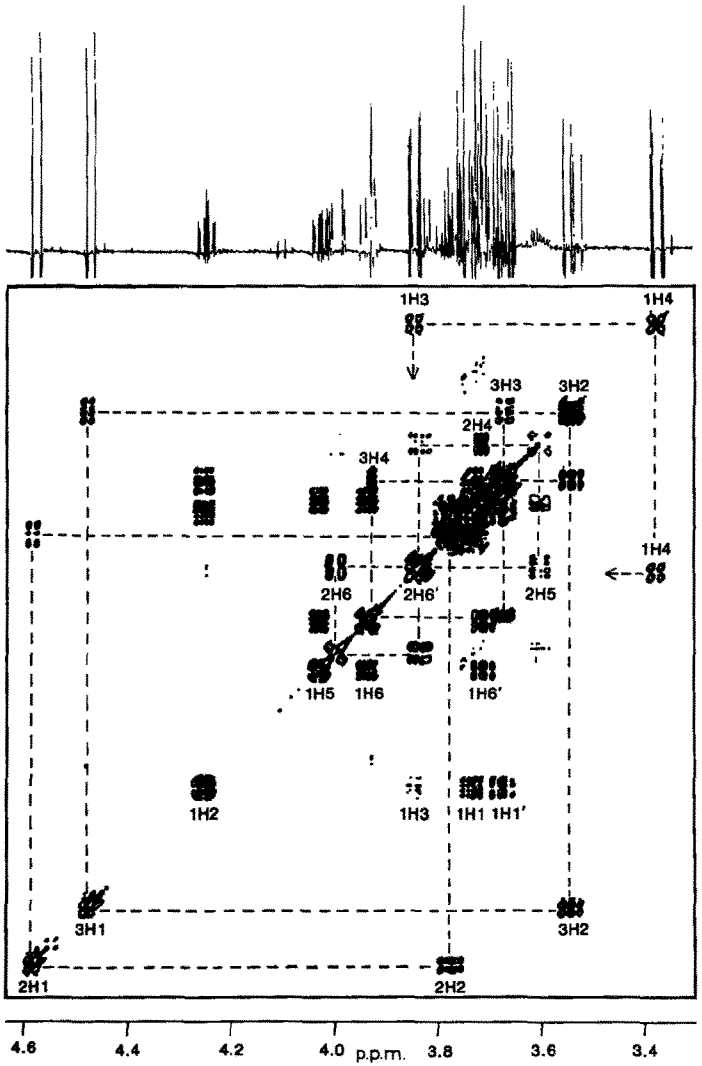


Fig. 3. Phase-sensitive double-quantum-filtered $^1$H COSY spectrum (3.30–4.64 p.p.m.) of β-Gal-(1→4)-β-GlcNAc-(1→6)-GalNAc-ol obtained at 500 MHz; the *N*-acetyl methyl protons resonate at higher field. A frequency scale has been attached to the $F_2$ axis, and dashed lines have been added to emphasise some of the correlations. Key: $F_2$ frequencies of some of the multiplets have been labelled using an abbreviated form iHj where i denotes the residue (numbered from the right) and j indicates the position within the residue, *e.g.*, 3H2 indicates H-2 of Gal.

multiplet for Gal H-5 at 3.715 p.p.m. It was not possible to assign the signals for Gal H-6,6′ from the COSY spectrum shown in Fig. 3, although they were known to be in the region 3.64–3.82 p.p.m.

For the GlcNAc residue, the H-5 resonance can be assigned readily to the multiplet at 3.600 p.p.m. on the basis of its chemical shift and multiplicity (ddd).

The COSY spectrum (Fig. 3) can then be used to assign GlcNAc H-4,6,6′. At this stage, the multiplet for GlcNAc H-3 could not be assigned although it is known to be in the region 3.64–3.82 p.p.m.; its coupled protons (H-2,4) also resonated in this narrow region and, thus, the cross-peaks in the COSY spectrum were near the diagonal and very difficult to detect. This problem can be overcome by using a RELAYED-COSY experiment[23–25] which gives the connectivities not only between directly coupled protons but also between non-directly coupled protons

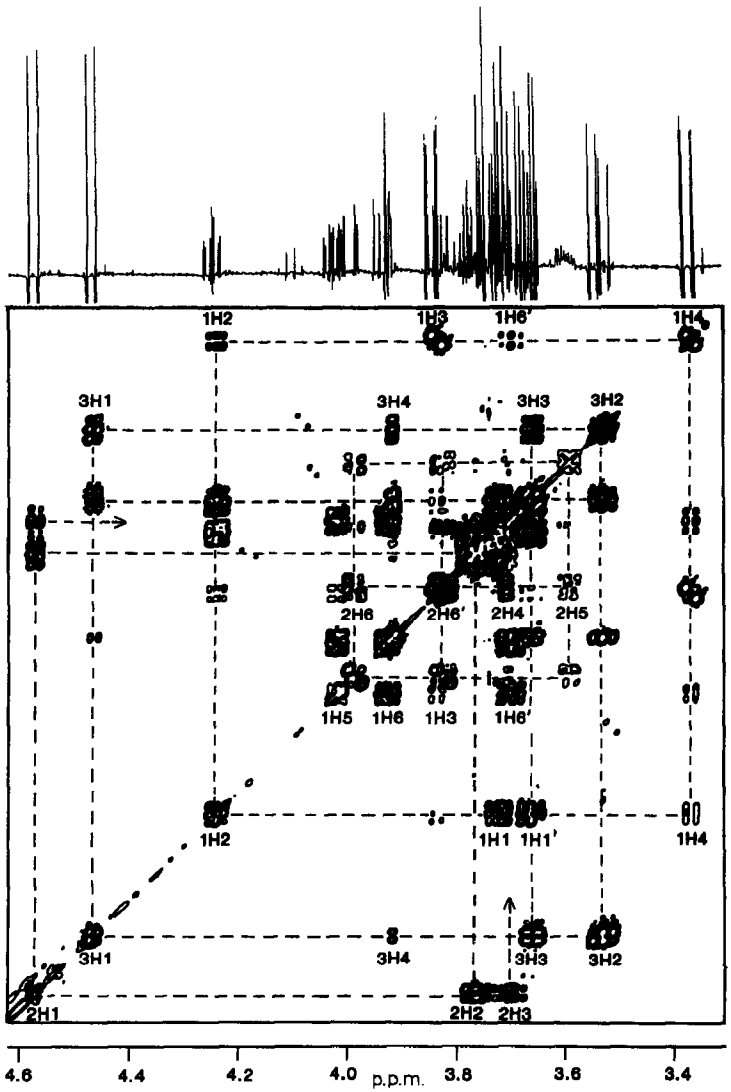


Fig. 4. Absolute-value RELAYED-COSY spectrum of β-Gal-(1→4)-β-GlcNAc-(1→6)-GalNAc-ol. The labelling scheme is the same as that used in Fig. 3, and a similar region of the spectrum has been plotted. A comparison with Fig. 3 reveals that additional cross-peaks have appeared, such as those between the multiplets labelled 2H1 and 2H3, 3H1 and 3H3, and 2H6′, and 2H4.

which are coupled to some other common nucleus. Thus, the RELAYED-COSY spectrum of the trisaccharide (Fig. 4) leads immediately to the assignment of GlcNAc H-3 from the appearance of a cross-peak between the signals for H-1 and H-3 of GlcNAc in a clear region of the 2D contour plot. Unfortunately, no connectivities between Gal H-6,6′ and Gal H-4,5 were detected.

In order to analyse the spectrum fully, it is necessary to assign the individual lines in the high-resolution 1D spectrum using the connectivities determined in the lower resolution 2D spectrum. This process can be assisted by examining the 1D spectrum for repeating splittings which provide an excellent correlation between connected multiplets in first-order spectra. Using this approach, all the assignments for multiplets in the 2D spectrum could be transferred to the higher resolution 1D spectrum. This approach revealed the remaining unassigned lines in the region 3.64–3.82 p.p.m. and these could then be assigned to Gal H-6,6′. The $^1$H chemical shifts and coupling constants extracted from an analysis of the spectrum of the trisaccharide are given in Table I.

Using this combination of 1D and 2D methods, the $^1$H-n.m.r. spectra of both oligosaccharides have been completely assigned. In each case, a final confirmation of the results was obtained by noting the excellent agreement between the observed spectral frequencies and those calculated using the parameters given in Table I. The 2D correlation experiments carried out did not by themselves provide the complete assignments. It was necessary to combine the results of the 2D experiments with an analysis of a highly resolved 1D spectrum, and this procedure provided an effective and efficient means of making the assignments.

ACKNOWLEDGMENTS

We thank Dr. H. van Halbeek for assistance in assigning some of the GalNAc-ol $^1$H signals.

REFERENCES

1 E. Wood, J. Lecomte, R. A. Childs, and T. Feizi, *Mol. Immunol.*, 16 (1979) 813–819.
2 H. C. Gooi, T. Feizi, A. Kapadia, B. B. Knowles, D. Solter, and M. J. Evans, *Nature (London)*, 292 (1981) 156–158.
3 T. Feizi, *Ciba Found. Symp.*, 96 (1983) 216–221.
4 H. C. Gooi, S. J. Thorpe, E. F. Hounsell, H. Rumpold, D. Kraft, O. Forster, and T. Feizi, *Eur. J. Immunol.*, 13 (1983) 306–312.
5 H. C. Gooi, L. K. Williams, K. Uemura, E. F. Hounsell, R. A. J. McIlhinney, and T. Feizi, *Mol. Immunol.*, 30 (1983) 607–613.
6 H. C. Gooi, K. Uemura, P. A. W. Edwards, C. S. Foster, N. Pickering, and T. Feizi, *Carbohydr. Res.*, 120 (1983) 293–302.
7 E. F. Hounsell, A. M. Lawson, J. Feeney, H. C. Gooi, N. J. Pickering, M. S. Stoll, S. C. Lui, and T. Feizi, *Eur. J. Biochem.*, 148 (1985) 367–377.
8 L. S. Forsberg, A. Dell, D. J. Walton, and C. E. Ballou, *J. Biol. Chem.*, 257 (1982) 3555–3563.
9 J. P. Kamerling, W. Heerma, J. F. G. Vliegenthart, B. N. Green, I. A. S. Lewis, G. Strecker, and G. Spik, *Biomed. Mass Spectrom.*, 1 (1983) 49–56.
10 E. F. Hounsell, M. J. Madigan, and A. M. Lawson, *Biochem. J.*, 219 (1984) 947–952.
11 E. F. Hounsell, E. Wood, T. Feizi, M. Fukuda, M. E. Powell, and S. Hakomori, *Carbohydr. Res.*, 90 (1981) 282–307.

12 K. A. KARLSSON, in L. A. WITTIG (Ed.), *Glycolipid Methodology*, American Oil Chemist's Society, Champaign, IL, 1976 pp. 97–122.
13 P. HALLGREN AND A. LUNDBLAD, *J. Biol. Chem.*, 252 (1977) 1023–1033.
14 A. M. LAWSON, E. F. HOUNSELL, AND T. FEIZI, *Int. J. Mass. Spectrom. Ion Phys.*, 48 (1983) 149–152.
15 J. F. G. VLIEGENTHART, H. VAN HALBEEK, AND L. DORLAND, *Pure Appl. Chem.*, 53 (1981) 45–77.
16 W. P. AUE, E. BARTHOLDI, AND R. R. ERNST, *J. Chem. Phys.*, 64 (1976) 2229–2246.
17 K. NAGAYAMA, A. KUMAR, K. WUTHRICH, AND R. R. ERNST, *J. Magn. Reson.*, 40 (1980) 321–334.
18 A. BAX AND R. FREEMAN, *J. Magn. Reson.*, 44 (1981) 542–561.
19 D. MARION AND K. WUTHRICH, *Biochem. Biophys. Res. Commun.*, 113 (1983) 967–974.
20 D. J. STATES, R. A. HABERKORN, AND D. J. RUBEN, *J. Magn. Reson.*, 48 (1982) 286–292.
21 U. PIANTINI, O. W. SORENSEN, AND R. R. ERNST, *J. Am. Chem. Soc.*, 104 (1982) 6800–6801.
22 M. RANCE, G. WAGNER, O. W. SORENSON, K. WUTHRICH, AND R. R. ERNST, *J. Magn. Reson.*, 59 (1984) 250–261.
23 G. EICH, G. BODENHAUSEN, AND R. R. ERNST, *J. Am. Chem. Soc.*, 104 (1982) 3731–3732.
24 G. WAGNER, *J. Magn. Reson.*, 55 (1983) 151–156.
25 G. KING AND P. E. WRIGHT, *J. Magn. Reson.*, 54 (1983) 328–332.
26 S. W. HOMANS, R. A. DWEK, D. L. FERNANDES, AND T. W. RADEMACHER, *Proc. Natl. Acad. Sci. U.S.A.*, 81 (1984) 6286–6289.
27 L. D. HALL, G. A. MORRIS, AND S. SUKUMAR, *J. Am. Chem. Soc.*, 102 (1980) 1745–1747.
28 S. W. HOMANS, R. A. DWEK, D. L. FERNANDES, AND T. W. RADEMACHER, *FEBS Lett.*, 164 (1983) 231–235.
29 S. W. HOMANS, R. A. DWEK, D. L. FERNANDES, AND T. W. RADEMACHER, *Biochim. Biophys. Acta*, 760 (1983) 256–261.
30 R. U. LEMIEUX, K. BOCK, L. T. J. DELBAERE, S. KOTO, AND V. S. RAO, *Can. J. Chem.*, 58 (1980) 631–653.
31 J. P. CARVER AND A. A. GREY, *Biochemistry*, 20 (1981) 6607–6616.
32 E. F. HOUNSELL, D. J. WRIGHT, A. S. R. DONALD, AND J. FEENEY, *Biochem. J.*, 223 (1984) 129–143.
33 H. VAN HALBEEK, L. DORLAND, J. HAVERKAMP, G. A. VELDINK, J. F. G. VLIEGENTHART, B. FOURNET, G. RICART, J. MONTREUIL, W. P. GATHMAN, AND D. AMINOFF, *Eur. J. Biochem.*, 118 (1981) 487–495.