Biochem. J. (1984) **223**, 129–143
Printed in Great Britain

A computerized approach to the analysis of oligosaccharide structure by high-resolution proton n.m.r.

Elizabeth F. HOUNSELL,* Dennis J. WRIGHT,† Alastair S. R. DONALD‡ and James FEENEY§
**Applied Immunochemistry Research Group and ‡Division of Immunochemical Genetics, Clinical Research Centre, Watford Road, Harrow, Middx. HA1 3UJ, U.K., †Clinical Chemistry Department, Northwick Park Hospital, Harrow, Middx. HA1 3UJ, U.K., and §M.R.C. Biomedical Nuclear Magnetic Resonance Centre, National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.*

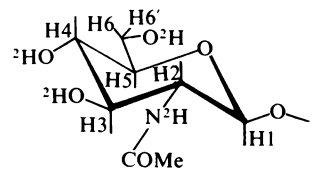
(*Received 16 April 1984/Accepted 6 June 1984*)

The 500MHz proton-n.m.r. spectra of 21 oligosaccharides, predominantly of the lacto-N and lacto-N-neo series and their derivatives containing non-reducing terminal fucose, sialic acid or *N*-acetylgalactosamine and reduced-end hexitol or hexosaminitol, were examined with ^2H_2O as solvent. The chemical-shift data obtained from analysis of the spectra were collated with data from other laboratories who have used 250–500MHz n.m.r. in the analysis of secreted and chemically synthesized oligosaccharides and of the *O*- and *N*-linked chains of glycoproteins. A referenced computer library was constructed that includes the chemical shifts of monosaccharides within oligosaccharide sequences that make up the majority of the carbohydrate structures found in mammalian glycoproteins. Examples are given of the computerized interrogation of this library for the assignment of possible structures of unknown oligosaccharides.

High-resolution proton-n.m.r. spectroscopy at high magnetic fields has emerged as a powerful technique in the structural analysis of oligosaccharides. Several groups have published data obtained at 250–500MHz for various types of oligosaccharide structure. These include the *N*-linked oligosaccharides of glycoproteins (Dorland *et al.*, 1978; Vliegenthart *et al.*, 1981; Bock *et al.*, 1982; Brisson & Carver, 1983), *O*-linked chains released from protein by base/borohydride degradation (Van Halbeek *et al.*, 1981, 1982, 1983), naturally occurring oligosaccharides (Dua & Bush, 1983; Dabrowski *et al.*, 1983), chemically synthesized glycosides and oligosaccharides (Augé *et al.*, 1979; Lemieux *et al.*, 1980; Hindsgaul *et al.*, 1982) and glycolipids (Dabrowski *et al.*, 1982; Koerner *et al.*, 1983).

Analysis of the proton-n.m.r. spectra of carbohydrate structures allows the assignment of signals arising from the several hydrogen atoms (protons) of the glycosidic rings of the constituent monosaccharides and from the acetamido methyl and glycollyl methylene groups where these are present. The protons are designated as H1, H2, H3 etc. according to the carbon atom to which they are attached in the glycosidic ring, as indicated for *N*-acetylglucosamine (structure 1). Usually the protons of the OH and NH groups are exchanged with deuterium by repeated evaporation with ^2H_2O and are not observed in the proton-n.m.r. spectrum. Each of the other protons resonates at a characteristic resonance radiofrequency, which is measured as a difference frequency from that of a reference compound. These difference frequencies are referred to as chemical shifts and are usually expressed in dimensionless units, as parts per million (p.p.m.) of the operating frequency of the instrument used (e.g. 500MHz).

Many of the glycosidic ring protons of oligosaccharides have similar chemical shifts clustered together in the 3.5–4p.p.m. range. A minority of the protons have shifts outside this range, and these easily assigned protons have been termed structural reporter groups (Vliegenthart *et al.*,



Structure 1

1981). They include the anomeric protons, the H2 proton of mannose, the H4 proton of galactose linked at C3, the H2 and H5 protons of *N*-acetylhexosaminitols, the H5 proton of fucose, the H3 axial and equatorial protons of sialic acids, the *N*-glycollyl methylene group and the *N*-acetamido and fucosyl methyl groups. Additional resonances can be assigned by using sequential spin-decoupling experiments to connect the assigned resonances of the structural reporter groups to their adjacent, 'coupled', protons, which in turn can be connected with their coupled neighbours. Two-dimensional correlated-spectroscopy experiments that provide all of this information in one experiment can also be carried out (Aué *et al.*, 1976).

The chemical shifts of the protons of a monosaccharide vary depending on the adjacent monosaccharides within an oligosaccharide sequence. As the chemical shifts are characteristic of the environment of each nucleus, the extensive amount of information available from n.m.r. analysis of different oligosaccharides forms a data base that is eminently suitable for computerization.

In the present paper we report chemical-shift data obtained from examining the 500 MHz n.m.r. of 21 oligosaccharides of the *N*-acetyl-lactosamine series, which contains structures widely distributed on *N*- and *O*-linked chains of glycoproteins, glycolipids and secreted oligosaccharides. We have combined our data and those from other groups into a library designed for a computerized approach to n.m.r. analysis of oligosaccharides. We have developed a nomenclature capable of describing sequences within all types of oligosaccharides, and used this as the basis of a method of interrogating the library to give possible assignments of oligosaccharides of unknown structures.

Materials and methods

Oligosaccharides

The following oligosaccharides obtained by chemical synthesis were gifts from Dr. Alain Veyrières (Augé *et al.*, 1979; Veyrières, 1981):

Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc
Galβ1–4GlcNAc
Galβ1–3GlcNAc

The blood-group-A-active tetrasaccharide (Yates *et al.*, 1984)

GalNAcα1–3Galβ1–4Glc
|1,2
Fucα

the blood-group-B-active tetrasaccharide (Race & Watkins, 1970)

Galα1–3Galβ1–4Glc
|1,2
Fucα

and the trisaccharides (Yates & Watkins, 1983)

GlcNAcβ1–3Galβ1–4Glc
GlcNAcβ1–3Galβ1–4GlcNAc

were prepared by enzymic methods.

The following oligosaccharides were obtained from human milk as described previously (Kobata, 1972; Anderson & Donald, 1981; Hounsell *et al.*, 1981):

Galβ1–4GlcNAcβ1–3Galβ1–4Glc (LNNT)
Galβ1–3GlcNAcβ1–3Galβ1–4Glc (LNT)
Galβ1–4GlcNAcβ1–3Galβ1–4Glc (LNFP III)
|1,3
Fucα

Galβ1–3GlcNAcβ1–3Galβ1–4Glc (LNFP II)
|1,4
Fucα

Fucα1–2Galβ1–3GlcNAcβ1–3Galβ1–4Glc (LNFP I)
Fucα1–2Galβ1–3GlcNAcβ1–3Galβ1–4Glc (LDFH I)
|1,4
Fucα

Fucα1–2Galβ1–4Glc (LDFT II)
|1,3
Fucα

Fucα1–2Galβ1–4Glc (2-fucosyl-lactose)
Galβ1–4Glc (3-fucosyl-lactose)
|1,3
Fucα

Siaα2–3Galβ1–4Glc (3′-sialyl-lactose)
Siaα2–6Galβ1–4Glc (6′-sialyl-lactose)

LNFP II and LNFP III were from a preparation of human milk kindly provided by Dr. J.-P. Prieels (Free University of Brussels, Brussels, Belgium). Lactose (Galβ1–4Glc) was obtained from Sigma Chemical Co. (Poole, Dorset, U.K.). Chitotriose (GlcNAcβ1–4GlcNAcβ1–4GlcNAc) and chitobiose were prepared by the acetolysis of chitin (Barker *et al.*, 1958). The pentasaccharide

GalNAcβ1–4Galβ1–4GlcNAcβ1–3Gal
|1,3
NeuAcα

having blood-group-Sd^a activity (Race & Sanger, 1975) was obtained from Tamm–Horsfall glycoprotein by treatment with an endo-β-galactosidase (Donald *et al.*, 1983).

Oligosaccharides were in general analysed as their alditols (designated -ol), except for 3'- and 6'-sialyl-lactose, which were examined without reduction. LDFT II and lactose were also examined after reduction with NaB^2H_4 in order to distinguish the H1 and H6 protons of the glucitol residue.

Nomenclature

The nomenclature and abbreviations used in this study are shown in Table 1. Monosaccharides are specified as GAL, FUC, NEUAC etc. when describing their chemical shifts. Adjacent monosaccharides are given 'non-specific' designations, e.g. HEX, DHEX, SA. Glycosidic linkages are defined by A or B for α- and β-configuration respectively and a number to indicate the position of linkage. Branching positions are given in square brackets. An example of the nomenclature is shown in Table 2 for the analysis of the oligosaccharide Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol. Sequences within the oligosaccharide are categorized into three groups, namely those containing the 'reducing/reduced end' residue, those specifying 'backbone' residues and those at the 'non-reducing end' of the molecule. The 'location' of each sequence and its 'designation' are those given in the data library (Table 3). Where more than one position of linkage of substituents gives the same chemical shift in the library, the possible positions of substituents are shown as consecutive numbers. For example, in sequence 31 (Table 3), β-galactose may be linked at the 4- or 3-position (HEXB43) and α-fucose either absent or linked at the 4- or 3-position (DHEXA043) to *N*-acetylglucosamine without changing the chemical shifts by more than ±0.03p.p.m. (the accuracy of the library data as described below). Where a specified monosaccharide has similar chemical shifts independent of differences in neighbouring monosaccharides, the alternative neighbouring monosaccharides in the sequence are separated by oblique lines, e.g. in sequence 43 (Table 3) the proton

Table 1. *Abbreviations used in the library*

Monosaccharides are specified as GAL, FUC, NEUAC etc. when describing their chemical shifts. Adjacent monosaccharides within an oligosaccharide sequence are given 'non-specific' abbreviations.

	Abbreviations	
Monosaccharide	'Specific'	'Non-specific'
Galactose	GAL	HEX
Mannose	MAN	HEX
Glucose	GLC	HEX
Fucose	FUC	DHEX (deoxyhexose)
N-Acetylgalactosamine	GALN	HEXN
N-Acetylglucosamine	GLCN	HEXN
Reduced derivative	-OL	-OL
N-Acetylneuraminic acid	NEUAC	SA (sialic acid)
N-Glycollylneuraminic acid	NEUGC	SA

Table 2. *Example of the nomenclature showing the sequences given in Table 3 for the analysis of the oligosaccharide Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol*

Structure and sequence abbreviations	Location	Sequence designation
Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol		
HEXNB 3HEXB 4GLCNOL	Reducing/reduced end	12
HEXNB 3GALB 4HEXNOL	Backbone	21
HEXB 4GLCNB 3HEX	Backbone	31
GALB 4HEXN	Non-reducing end	43

Table 3. *Library of chemical-shift data for H1 (α or β), H2, H3 axial (a) and H3 equatorial (e), H4 and H5 glycosidic ring protons and the two H6 protons (H6 and H6′) of hexopyranose structures, the acetamido methyl group (Ac) and the glycollyl methylene group (Gc)*

Designations 16, 17 and 42 are left vacant for additional sequences in the 'reducing/reduced end' and 'backbone' categories such as those for backbone residues in high-mannose-type glycoprotein chains and chains having bisecting *N*-acetylglucosamine residues (Brisson & Carver, 1983).

Oligosaccharide sequence		Chemical shifts (p.p.m.)									
Designation	Abbreviation	H1α	H1β	H2	H3a	H3e	H4	H5	H6	H6′	Ac/Gc
	Reducing/reduced end										
1	HEXB3GALNOL	3.74	3.67	4.40	4.07		3.51	4.18	3.69	3.65	2.051
2	HEXB3[HEXNB6]GALNOL			4.39	4.06		3.46	4.28	3.93		2.067
3	HEXNB3GALNOL			4.28	4.00		3.56	4.13	3.65		2.036
4	HEXNB3[HEXNB6]GALNOL			4.28	3.99		3.52	4.24	3.91		2.045
5	SAA6GALNOL	3.73	3.66	4.24	3.84		3.41		3.84	3.53	2.056
6	SAA6[HEXB3]GALNOL	3.86	3.75	4.38	4.06		3.53	4.24	3.86	3.49	2.048
7	HEXB4GLCNOL			4.31	3.79		3.86				2.053
8	HEXNB4GLCNOL			4.25							2.051
9	HEXB4GLCOL	3.75	3.61	3.96	3.81		3.85	3.92	3.86	3.73	
10	HEXB4[DHEXA3]GLCOL			4.16	4.06		3.91		3.65	3.55	
11	HEXB3GLCNOL			4.25				4.15			2.035
12	HEXNB3HEXB4GLCNOL			4.28							2.049
13	HEXNB63[HEXNB063]GAL	5.23	4.55								
14	HEXB4GLCN	5.22	4.71								2.04
15	HEXB4GLC	5.22	4.66	3.27	3.64						
	Backbone										
18	SAA6GALB4HEX/HEXNB			4.44	3.31						
19	SAA3GALB43HEX/HEXNB/OL			4.54	4.12		3.92				
20	HEXNA4GALB3HEXNOL			4.52	3.75		3.97				
21	HEXNB3GALB643HEXNOL			4.47	3.75	3.76	4.15		3.79		
22	DHEXA2GALB43			4.60	3.69	3.86	3.91	3.69	3.77		
23	DHEXA2[HEXA3]GALB43			4.66							
24	DHEXA2[HEXNA3]GALB43			4.72	3.90		4.25				
25	HEXNB3GALB4HEX			4.43		3.71	4.16				
26	HEXNB3GALB4HEXOL			4.48	3.59	3.72	4.15				
27	SAA3[HEXNB4]GALB4HEXN			4.53	3.35						
28	HEXNB3[HEXNB6]GALB43HEXN/OL			4.47			4.10				
29	HEX/HEXNB4GLCNB4HEXN/OL			4.60	3.74						2.07
30	HEXB43[DHEXA043]GLCNB6HEXNOL			4.56						4.00	2.06
31	HEXB43[DHEXA043]GLCNB3HEX			4.71	3.77				3.96	3.85	2.03
32	DHEXA2HEXB3[DHEXA04]GLCNB3HEX			4.66							2.06
33	HEXB4GLCNB3HEXNOL			4.63					4.02		2.08
34	HEXB3GLCNB3HEXNOL			4.65					3.95		2.09
35	HEXB4GLCNB642HEX			4.59	3.75	3.72	3.77	3.59	3.98	3.84	2.02
36	HEXA6MANB4HEXN			4.77	4.08						
37	HEXA06[HEXA3]MANB4HEXN			4.77	4.25						
38	HEXNB2MANA6HEX	4.93		4.11							
39	HEXNB2MANA3HEX	5.13		4.19							
40	HEXNB6[HEXNB2]MANA6HEX										
41	HEXNB4[HEXNB2]MANA3HEX	5.13		4.21							
	Non-reducing end										
43	GALB643HEXN/HEX/OL			4.47	3.56	3.65	3.92	3.67	3.79	3.75	
44	GLCNB6HEXNOL			4.54					3.93		2.07
45	GLCNB3HEXNOL			4.60		3.58			3.96		2.08
46	GLCNB6[HEXNB3]HEX			4.78							2.08
47	GLCNB3[HEXNB6]HEX			4.62							2.08
48	GALA3[DHEXA2]HEX	5.22			3.83	3.83	3.94	4.21	3.76	3.69	
49	GALNA3[DHEXA2]HEX	5.18		4.26	3.91		4.00	4.21	3.81	3.74	2.04
50	GLCNA4HEX	4.87						4.17			2.08
51	FUCA2HEXB4	5.31			3.83	3.92	3.82	4.25	1.23		
52	FUCA2[HEXA3]HEX	5.35			3.67	3.76	3.72	4.42	1.22		
53	FUCA2[HEXNA3]HEX	5.36			3.78	3.85	3.73	4.46	1.26		
54	FUCA3[HEXB4]HEXN	5.11			3.74	3.96	3.83	4.88	1.20		
55	FUCA4[HEXB3]HEXN	5.02			3.80	3.91	3.80	4.88	1.20		
56	FUCA3[HEXB4]HEXOL	5.08			3.83	3.68	3.91	4.27	1.22		
57	FUCA6[HEXNB4]HEXN	4.88						4.13	1.21		
58	NEUACA6HEX					1.75 2.72					2.03
59	NEUACA3HEX					1.80 2.76					2.03
60	NEUACA6HEXNOL					1.69 2.72					2.03
61	NEUGCA6HEXNOL					1.71 2.75					4.12
62	NEUACA3[HEXNB4]HEX					1.93 2.66					2.03
63	FUCA2HEXB3	5.19						4.28	1.23		
64	FUCA2HEXB4[DHEXA3]HEXOL	5.42			3.80	3.92	3.79	4.20	1.23		
65	FUCA3[DHEXA2HEXB4]HEXOL	5.07			3.83		3.81	4.20	1.23		
66	FUCA2HEXB3[DHEXA4]HEXN	5.15						4.34	1.26		
67	FUCA4[DHEXA2HEXB3]HEXN	5.03						4.13	1.27		
68	GLCNB3HEX			4.69	3.75						2.037
69	GLCNB4HEXN			4.61	3.79						2.067
70	GALNB4[SAA3]HEX			4.71							2.03

chemical shifts of the non-reducing terminal galactose are within ±0.03 p.p.m. whether the adjacent monosaccharide is HEXN, HEXNOL, HEX or HEXOL.

Data handling

Chemical shifts were measured in parts per million (p.p.m.) from the proton resonance of acetone (1 μl of 10% solution in ^2H_2O added to

each sample), and then referenced to the standard sodium 4,4-dimethyl-4-silapentane-1-sulphonate by addition of 2.225p.p.m. Chemical shifts from our data were compared with those in the literature cited in Table 4. A comparison of the data for three protons is given in Table 5. In order to render the data comparable, those of Lemieux's group (Lemieux *et al.*, 1980; Hindsgaul *et al.*, 1982), given in the literature related to acetone at 2.480p.p.m., were corrected by subtraction of 0.255p.p.m. (=2.480−2.225p.p.m.) and those of Dua & Bush (1983), given in the literature related to acetone at 2.216p.p.m., were corrected by addition of 0.009p.p.m. (=2.225−2.216). Data from other laboratories were used as published. The chemical-shift values for each proton of a specified monosaccharide within a di- or tri-saccharide sequence were averaged, and the standard deviation of ±0.03p.p.m. was used as the confidence limit of the library for the H1–H6 protons (shifts given to two decimal places in Table 3) and of ±0.005p.p.m. for the chemical shifts of the *N*-acetyl groups given to three places of decimals (Table 3). For the majority of oligosaccharides the chemical shifts were dependent only on the type of residue and linkage adjacent to the specified monosaccharide, and thus reducing/reduced end and non-reducing end sequences could be given as disaccharide and backbones as trisaccharide sequences. However, comparison of certain oligosaccharides showed that monosaccharides in sequence positions non-adjacent to the specified monosaccharide (but presumably close neighbours in the three-dimensional shape of the molecule) influenced its chemical shift by more than ±0.03p.p.m. In these cases sequences longer than di- or tri-saccharide were added to the library.

For oligosaccharides of unknown structure the observed chemical shifts were entered into the computer program and a search for matching signals in the library sequences was initiated. As shown in Tables 6–8, a print-out was obtained

Table 4. *References to the chemical-shift data used in the library*

Reference	Oligosaccharide sequence designations
Vliegenthart *et al.* (1981)	1, 5, 6, 14, 18, 19, 20, 22, 29, 35, 36, 37, 38, 39, 41, 43, 50, 51, 54, 57, 58, 60, 61
Van Halbeek *et al.* (1981)	1, 5, 6, 22, 24, 43, 61, 63
Van Halbeek *et al.* (1982)	1, 2, 3, 4, 21, 22, 30, 31, 33, 34, 43, 44, 45
Van Halbeek *et al.* (1983)	1, 2, 20, 21, 22, 28, 43, 44, 50, 51
Dorland *et al.* (1978)	14, 18, 19, 29, 35, 36, 37, 38, 39, 41, 43, 58, 59
Bock *et al.* (1982)	35, 38, 39, 43
Lemieux *et al.* (1980)	22, 23, 24, 48, 49, 51, 55
Hindsgaul *et al.* (1982)	22, 43, 51, 54
Augé *et al.* (1979)	13, 31, 35, 43, 46, 47
Dua & Bush (1983)	15, 22, 25, 31, 32, 43, 54, 55, 63, 66
Present study	7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19, 21, 22, 23, 24, 25, 26, 27, 29, 31, 32, 43, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 62, 63, 64, 65, 66, 67, 68, 69, 70

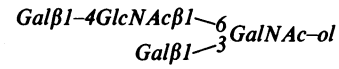
Table 5. *Comparison of the chemical-shift data for several oligosaccharide sequences reported in the cited references*
The data given by Lemieux *et al.* (1980) and Hindsgaul *et al.* (1982), shown in parentheses in the Table related to acetone at 2.480p.p.m., were corrected by subtraction of 0.255p.p.m. (2.480−2.225p.p.m.). The data given by Dua & Bush (1983), shown in parentheses in the Table related to acetone at 2.216p.p.m., were corrected by addition of 0.009p.p.m. (2.225−2.216p.p.m.). The remaining data were used as published.

	Chemical shift (p.p.m.)			
Sequence abbreviation ... (designation) ...	Galactose H1β GALB4HEXN (43)		Fucose H1α FUCA3[HEXB4]HEXN (54)	Fucose H1α FUCA2HEXB3HEXN (63)
Reference	Range	Average		
Present study	4.48–4.50	4.48	5.13	5.19
Vliegenthart *et al.* (1981)	4.45–4.48	4.46	5.13	5.21
Bock *et al.* (1982)	4.46–4.47	4.47		
Lemieux *et al.* (1980)	(4.75)	4.49	(5.39) 5.13	(5.44) 5.19
Dua & Bush (1983)	(4.42–4.44)	4.44	(5.12) 5.13	(5.17) 5.18
Augé *et al.* (1979)	4.44–4.50	4.47	–	
Value used	4.47 (±0.03)		5.13	5.19 (±0.02)

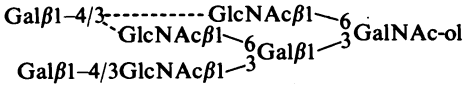
Table 6. *Structural interpretation of the chemical shifts of the disaccharide Galβ1–4GlcNAc-ol*
Underlined sequences give two possible structural assignments Galβ1–4GlcNAc-ol and Galβ1–4Glc-ol. From the number and relative intensity of the anomeric protons given in the spectrum (Fig. 1*a*) it can be seen that the latter structure is not present as a contaminant. The chemical shifts corresponding to the Glc-ol in this latter sequence coincidentally have chemical shifts present in the spectrum of Galβ1–4GlcNAc-ol. This is also true of the chemical shifts of sequences 12 and 20. Presence of these sequences in the compound analysed can be disregarded because there is no complementary full match for either β- or α-HexNAc or →3HexNAc-ol.

Abbreviation and designation		Available shifts in library	No. of matching shifts	
HEXB3GALNOL	1	9	5	
HEXB3[HEXNB6]GALNOL	2	6	2	
HEXNB3GALNOL	3	6	3	
HEXNB3[HEXNB6]GALNOL	4	6	3	
SAA6GALNOL	5	8	6	
SAA6[HEXB3]GALNOL	6	9	5	
HEXB4GLCNOL	7	4	4	FULL MATCH
HEXNB4GLCNOL	8	2	1	
HEXB4GLCOL	9	8	8	FULL MATCH
HEXB4[DHEXA3]GLCOL	10	5	3	
HEXB3GLCNOL	11	3	0	
HEXNB3HEXB4GLCNOL	12	2	2	FULL MATCH
HEXNB63[HEXNB063]GAL	13	2	0	
HEXB4GLCN	14	3	1	
HEXB4GLC	15	4	1	
	16	0	0	
	17	0	0	
SAA6GALB4HEX/HEXNB	18	2	0	
SAA3GALB43HEX/HEXNB/OL	19	3	1	
HEXNA4GALB3HEXNOL	20	3	3	FULL MATCH
HEXNB3GALB643HEXNOL	21	5	4	
DHEXA2GALB43	22	6	5	
DHEXA2[HEXA3]GALB43	23	1	0	
DHEXA2[HEXNA3]GALB43	24	3	1	
HEXNB3GALB4HEX	25	3	1	
HEXNB3GALB4HEXOL	26	4	3	
SAA3[HEXNB4]GALB4HEXN	27	2	1	
HEXNB3[HEXNB6]GALB43HEXN/OL	28	2	1	
HEX/HEXNB4GLCNB4HEXN/OL	29	3	2	
HEXB43[DHEXA043]GLCNB6HEXNOL	30	3	1	
HEXB43[DHEXA043]GLCNB3HEX	31	5	4	
DHEXA2HEXB3[DHEXA04]GLCNB3HEX	32	2	1	
HEXB4GLCNB3HEXNOL	33	3	1	
HEXB3GLCNB3HEXNOL	34	3	1	
HEXB4GLCNB642HEX	35	8	5	
HEXA6MANB4HEXN	36	2	0	
HEXA06[HEXA3]MANB4HEXN	37	2	0	
HEXNB2MANA6HEX	38	2	0	
HEXNB2MANA3HEX	39	2	0	
HEXNB6[HEXNB2]MANA6HEX	40	0	0	
HEXNB4[HEXNB2]MANA3HEX	41	2	0	
	42	0	0	
GALB643HEXN/HEX/OL	43	7	7	FULL MATCH
GLCNB6HEXNOL	44	3	2	
GLCNB3HEXNOL	45	4	3	
GLCNB6[HEXNB3]HEX	46	2	1	
GLCNB3[HEXNB6]HEX	47	2	1	
GALA3[DHEXA2]HEX	48	7	5	
GALNA3[DHEXA2]HEX	49	8	4	
GLCNA4HEX	50	3	1	
FUCA2HEXB4	51	6	3	
FUCA2[HEXA3]HEX	52	6	3	
FUCA2[HEXNA3]HEX	53	6	3	
FUCA3[HEXB4]HEXN	54	6	3	
FUCA4[HEXB3]HEXN	55	6	3	
FUCA3[HEXB4]HEXOL	56	6	4	
FUCA6[HEXNB4]HEXN	57	3	0	
NEUACA6HEX	58	3	1	
NEUACA3HEX	59	3	1	
NEUACA6HEXNOL	60	3	1	
NEUGLA6HEXNOL	61	3	0	
NEUACA3[HEXNB4]HEX	62	3	1	
FUCA2HEXB3	63	3	1	
FUCA2HEXB4[DHEXA3]HEXOL	64	6	3	
FUCA3[DHEXA2HEXB4]HEXOL	65	5	2	
FUCA2HEXB3[DHEXA4]HEXN	66	3	1	
FUCA4[DHEXA2HEXB3]HEXN	67	3	0	
GLCNB3HEX	68	3	1	
GLCNB4HEXN	69	3	1	
GALNB4[SAA3]HEX	70	2	1	

Table 7. *Structural interpretation of the chemical shifts of the tetrasaccharide*



The underlined 'full-match' sequences are compatible with this structure. The following larger structure including sequences 28 and 44 can also be proposed, having one or two non-reducing-end Gal residues (sequence 43) and one or two non-reducing end GlcNAcβ1–6 residues (sequence 44):



The presence of this structure can readily be ruled out from the number of anomeric protons given in the spectrum.

Abbreviation and designation		Available shifts in library	No. of matching shifts	
HEXB3GALNOL	1	9	7	
HEXB3[HEXNB6]GALNOL	2	6	6	FULL MATCH
HEXNB3GALNOL	3	6	4	
HEXNB3[HEXNB6]GALNOL	4	6	4	
SAA6GALNOL	5	8	3	
SAA6[HEXB3]GALNOL	6	9	5	
HEXB4GLCNOL	7	4	2	
HEXNB4GLCNOL	8	2	1	
HEXB4GLCOL	9	8	5	
HEXB4[DHEXA3]GLCOL	10	5	4	
HEXB3GLCNOL	11	3	1	
HEXNB3HEXB4GLCNOL	12	2	1	
HEXNB63[HEXNB063]GAL	13	2	1	
HEXB4GLCN	14	3	1	
HEXB4GLC	15	4	1	
	16	0	0	
	17	0	0	
SAA6GALB4HEX/HEXNB	18	2	1	
SAA3GALB43HEX/HEXNB/OL	19	3	2	
HEXNA4GALB3HEXNOL	20	3	2	
HEXNB3GALB643HEXNOL	21	5	4	
DHEXA2GALB43	22	6	4	
DHEXA2[HEXA3]GALB43	23	1	0	
DHEXA2[HEXNA3]GALB43	24	3	2	
HEXNB3GALB4HEX	25	3	2	
HEXNB3GALB4HEXOL	26	4	3	
SAA3[HEXNB4]GALB4HEXN	27	2	1	
HEXNB3[HEXNB6]GALB43HEXN/OL	28	2	2	FULL MATCH
HEX/HEXNB4GLCNB4HEXN/OL	29	3	2	
HEXB43[DHEXA043]GLCNB6HEXNOL	30	3	3	FULL MATCH
HEXB43[DHEXA043]GLCNB3HEX	31	5	2	
DHEXA2HEXB3[DHEXA04]GLCNB3HEX	32	2	1	
HEXB4GLCNB3HEXNOL	33	3	2	
HEXB3GLCNB3HEXNOL	34	3	2	
HEXB4GLCNB642HEX	35	8	6	
HEXA6MANB4HEXN	36	2	1	
HEXA06[HEXA3]MANB4HEXN	37	2	1	
HEXNB2MANA6HEX	38	2	0	
HEXNB2MANA3HEX	39	2	0	
HEXNB6[HEXNB2]MANA6HEX	40	0	0	
HEXNB4[HEXNB2]MANA3HEX	41	2	0	
	42	0	0	
GALB643HEXN/HEX/OL	43	7	7	FULL MATCH
GLCNB6HEXNOL	44	3	3	FULL MATCH
GLCNB3HEXNOL	45	4	3	
GLCNB6[HEXNB3]HEX	46	2	1	
GLCNB3[HEXNB6]HEX	47	2	1	
GALA3[DHEXA2]HEX	48	7	3	
GALNA3[DHEXA2]HEX	49	8	6	
GLCNA4HEX	50	3	1	
FUCA2HEXB4	51	6	3	
FUCA2[HEXA3]HEX	52	6	4	
FUCA2[HEXNA3]HEX	53	6	3	
FUCA3[HEXB4]HEXN	54	6	2	
FUCA4[HEXB3]HEXN	55	6	3	
FUCA3[HEXB4]HEXOL	56	6	3	
FUCA6[HEXNB4]HEXN	57	3	0	
NEUACA6HEX	58	3	0	
NEUACA3HEX	59	3	0	
NEUACA6HEXNOL	60	3	0	
NEUGLA6HEXNOL	61	3	0	
NEUACA3[HEXNB4]HEX	62	3	0	
FUCA2HEXB3	63	3	1	
FUCA2HEXB4[DHEXA3]HEXOL	64	6	3	
FUCA3[DHEXA2HEXB4]HEXOL	65	5	1	
FUCA2HEXB3[DHEXA4]HEXN	66	3	0	
FUCA4[DHEXA2HEXB3]HEXN	67	3	0	
GLCNB3HEX	68	3	1	
GLCNB4HEXN	69	3	2	
GALNB4[SAA3]HEX	70	2	0	

Table 8. *Structural interpretation of the chemical shifts of the pentasaccharide*

Galβ1–4GlcNAcβ1–3Galβ1–4Glc-ol
|1,3
Fucα

The underlined 'full-match' sequences are compatible with this structure. Sequences 20 and 50 and 21 and 70 allow the possibility of the alternative structures GlcNAcα1–4Galβ1–4/3GlcNAcβ1–3Galβ1–4Glc-ol and GalNAcβ1–4/3GlcNAcβ1–3Galβ1–4Glc-ol. These structures also both include sequences 9, 26 and 31. Branched structures, which could additionally include sequences 43 and 54, and the inclusion of sequences 25 into longer backbones, can be ruled out, as the resulting oligosaccharides would of necessity be larger than the pentasaccharide shown to be present from the number of anomeric protons in the spectrum.

Abbreviation and designation		Available shifts in library	No. of matching shifts	
HEXB3GALNOL	1	9	6	
HEXB3[HEXNB6]GALNOL	2	6	2	
HEXNB3GALNOL	3	6	4	
HEXNB3[HEXNB6]GALNOL	4	6	3	
SAA6GALNOL	5	8	6	
SAA6[HEXB3]GALNOL	6	9	5	
HEXB4GLCNOL	7	4	2	
HEXNB4GLCNOL	8	2	0	
HEXB4GLCOL	9	8	8	FULL MATCH
HEXB4[DHEXA3]GLCOL	10	5	4	
HEXB3GLCNOL	11	3	1	
HEXNB3HEXB4GLCNOL	12	2	0	
HEXNB63[HEXNB063]GAL	13	2	0	
HEXB4GLCN	14	3	2	
HEXB4GLC	15	4	1	
	16	0	0	
	17	0	0	
SAA6GALB4HEX/HEXNB	18	2	1	
SAA3GALB43HEX/HEXNB/OL	19	3	2	
HEXNA4GALB3HEXNOL	20	3	3	FULL MATCH
HEXNB3GALB643HEXNOL	21	5	5	FULL MATCH
DHEXA2GALB43	22	6	5	
DHEXA2[HEXA3]GALB43	23	1	0	
DHEXA2[HEXNA3]GALB43	24	3	2	
HEXNB3GALB4HEX	25	3	3	FULL MATCH
HEXNB3GALB4HEXOL	26	4	4	FULL MATCH
SAA3[HEXNB4]GALB4HEXN	27	2	0	
HEXNB3[HEXNB6]GALB43HEXN/OL	28	2	1	
HEX/HEXNB4GLCNB4HEXN/OL	29	3	2	
HEXB43[DHEXA043]GLCNB6HEXNOL	30	3	2	
HEXB43[DHEXA043]GLCNB3HEX	31	5	5	FULL MATCH
DHEXA2HEXB3[DHEXA04]GLCNB3HEX	32	2	1	
HEXB4GLCNB3HEXNOL	33	3	1	
HEXB3GLCNB3HEXNOL	34	3	2	
HEXB4GLCNB642HEX	35	8	7	
HEXA6MANB4HEXN	36	2	0	
HEXA06[HEXA3]MANB4HEXN	37	2	0	
HEXNB2MANA6HEX	38	2	0	
HEXNB2MANA3HEX	39	2	1	
HEXNB6[HEXNB2]MANA6HEX	40	0	0	
HEXNB4[HEXNB2]MANA3HEX	41	2	1	
	42	0	0	
GALB643HEXN/HEX/OL	43	7	7	FULL MATCH
GLCNB6HEXNOL	44	3	2	
GLCNB3HEXNOL	45	4	3	
GLCNB6[HEXNB3]HEX	46	2	1	
GLCNB3[HEXNB6]HEX	47	2	1	
GALA3[DHEXA2]HEX	48	7	5	
GALNA3[DHEXA2]HEX	49	8	5	
GLCNA4HEX	50	3	3	FULL MATCH
FUCA2HEXB4	51	6	3	
FUCA2[HEXA3]HEX	52	6	3	
FUCA2[HEXNA3]HEX	53	6	4	
FUCA3[HEXB4]HEXN	54	6	6	FULL MATCH
FUCA4[HEXB3]HEXN	55	6	5	
FUCA3[HEXB4]HEXOL	56	6	3	
FUCA6[HEXNB4]HEXN	57	3	2	
NEUACA6HEX	58	3	1	
NEUACA3HEX	59	3	1	
NEUACA6HEXNOL	60	3	1	
NEUGLA6HEXNOL	61	3	1	
NEUACA3[HEXNB4]HEX	62	3	1	
FUCA2HEXB3	63	3	0	
FUCA2HEXB4[DHEXA3]HEXOL	64	6	3	
FUCA3[DHEXA2HEXB4]HEXOL	65	5	2	
FUCA2HEXB3[DHEXA4]HEXN	66	3	1	
FUCA4[DHEXA2HEXB3]HEXN	67	3	1	
GLCNB3HEX	68	3	2	
GLCNB4HEXN	69	3	1	
GALNB4[SAA3]HEX	70	2	2	FULL MATCH

indicating how many proton chemical shifts were available in the library for each sequence and how many matching chemical shifts were found in the spectrum under study.

N.m.r. spectroscopy

The 500MHz proton-n.m.r. spectra were obtained with a Bruker AM500 spectrometer operating in the Fourier-transform mode and equipped with an Aspect 2000 computer. Spectra were obtained by using quadrature detection with a spectral width of 6000Hz, and the data were collected using 32000 data points (giving 0.37Hz/point digital resolution). Usually 300–500 transients were accumulated. Before Fourier transformation the free-induction decay was multiplied by an exponential function giving a line broadening of 0.5Hz to improve the signal-to-noise ratio. In several cases resolution-enhanced spectra were obtained by using a Gaussian multiplication technique (Ferrige & Lindon, 1979).

Lacitol was examined by two-dimensional correlated spectroscopy. In these experiments the digital resolution was 1.17Hz/point (2400Hz spectral width and 4000 data points in F_2 dimension).

The probe temperature was 295°C in all the experiments.

Results and discussion

The library of available chemical shifts for oligosaccharide sequences in mammalian glycoproteins is given in Table 3. This library of data is a composite of literature values from several groups who have used ^2H_2O as solvent together with our results obtained from analysis of the 500MHz spectra of the oligosaccharides listed in the Materials and methods section. References for the studies on the various oligosaccharide sequences are given in Table 4.

Original data

Oligosaccharides with reduced-end N-acetylglucosaminitol residues and/or non-reducing-end N-acetylglucosamine. N.m.r. data on oligosaccharides having reduced-end glucosaminitol and non-reducing-end *N*-acetylglucosamine were not available in the literature, and we have therefore studied the 500MHz spectra of the following oligosaccharide alditols:

Galβ1–4GlcNAc-ol
Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol
GlcNAcβ1–4GlcNAcβ1–4GlcNAc-ol
GlcNAcβ1–4GlcNAc-ol
GlcNAcβ1–3Galβ1–4GlcNAc-ol
Galβ1–3GlcNAc-ol

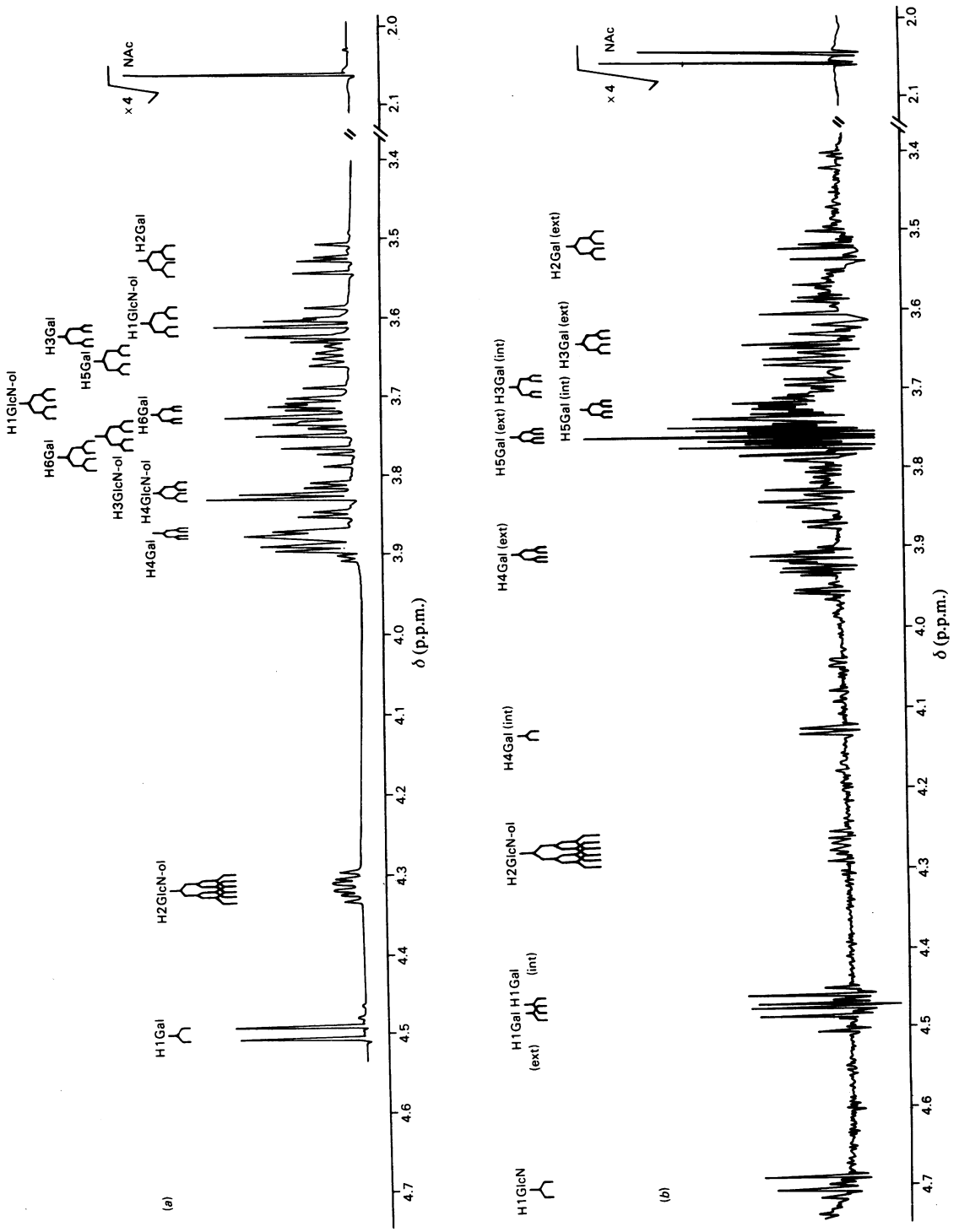
The spectra of the first three of these compounds are shown in Fig. 1. Their chemical shifts were assigned by spin-decoupling experiments and by comparison of the n.m.r. spectra with each other and with those of the trisaccharide GlcNAcβ1–3Galβ1–4Glc-ol and lactitol shown in Fig. 2.

The chemical shifts of the monosaccharides within these seven alditols are given in the library (Table 3) with the sequence designations as follows: reduced-end *N*-acetylglucosaminitol, sequences 7, 8, 11 and 12; reduced-end glucitol, sequence 9; backbone galactose, sequences 21 and 26; backbone *N*-acetylglucosamine, sequences 29 and 31; non-reducing-end galactose, sequence 43; non-reducing-end *N*-acetylglucosamine, sequences 68 and 69.

Assignment of the chemical shifts for the tetrasaccharide Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol was aided by comparison of those for LNNT (Dua & Bush, 1983) and our studies on LNNT-ol. This comparison enabled the chemical shifts of the external, non-reducing-end, galactose (sequence 43) and the internal galactose (sequence 21) to be distinguished and also those of the acetamido methyl protons of the internal *N*-acetylglucosamine residue. The H1 chemical shifts of the galactose residues were 4.489 and 4.474p.p.m. respectively and that for the internal acetamido methyl group 2.034p.p.m. (the corresponding chemical shifts of LNNT-ol were 4.494, 4.478 and 2.033p.p.m. respectively). The acetamido methyl protons of the *N*-acetylglucosaminitol could thus be assigned the chemical shift of 2.049p.p.m. This assignment was corroborated by analysis of the trisaccharides GlcNAcβ1–3Galβ1–4Glc-ol and GlcNAcβ1–3Galβ1–4GlcNAc-ol (the acetamido methyl-group protons had chemical shifts of 2.037, 2.037 and 2.049p.p.m. respectively).

The acetamido methyl groups of the disaccharides Galβ1–4GlcNAc-ol and Galβ1–3GlcNAc-ol had chemical shifts of 2.053 and 2.035p.p.m. respectively, and GlcNAcβ1–4GlcNAc-ol 2.066 and 2.052p.p.m. Those of chitotri-itol (GlcNAcβ1–4GlcNAcβ1–4GlcNAc-ol) at 2.067, 2.057 and 2.051p.p.m. were therefore assigned to the non-reducing-end, the internal and the reduced residue respectively. The data of Dorland *et al.* (1978) and Vliegenthart *et al.* (1981) give the average chemical shift of the internal GlcNAc of the sequence Manβ1–4GlcNAcβ1–4GlcNAc attached to asparagine as 2.075p.p.m. in the absence of, and 2.095p.p.m. in the presence of, a fucose residue attached to the adjacent GlcNAc–Asn. Thus internal GlcNAc in sequence 29 has been assigned the chemical shift of 2.07 ± 0.03p.p.m.

The chemical shifts of the H1 protons of the two non-reducing *N*-acetylglucosamine residues of chi-


(a)
H1Gal
H2GlcN-ol
H4Gal
H4GlcN-ol
H6Gal
H3GlcN-ol
H6Gal
H1GlcN-ol
H5Gal
H3Gal
H1GlcN-ol
H2Gal
NAc
×4
δ (p.p.m.)
(b)
H1GlcN
H1Gal (ext)
H1Gal (int)
H2GlcN-ol
H4Gal (int)
H4Gal (ext)
H5Gal (ext)
H3Gal (int)
H5Gal (int)
H3Gal (ext)
H2Gal (ext)
NAc
×4
δ (p.p.m.)

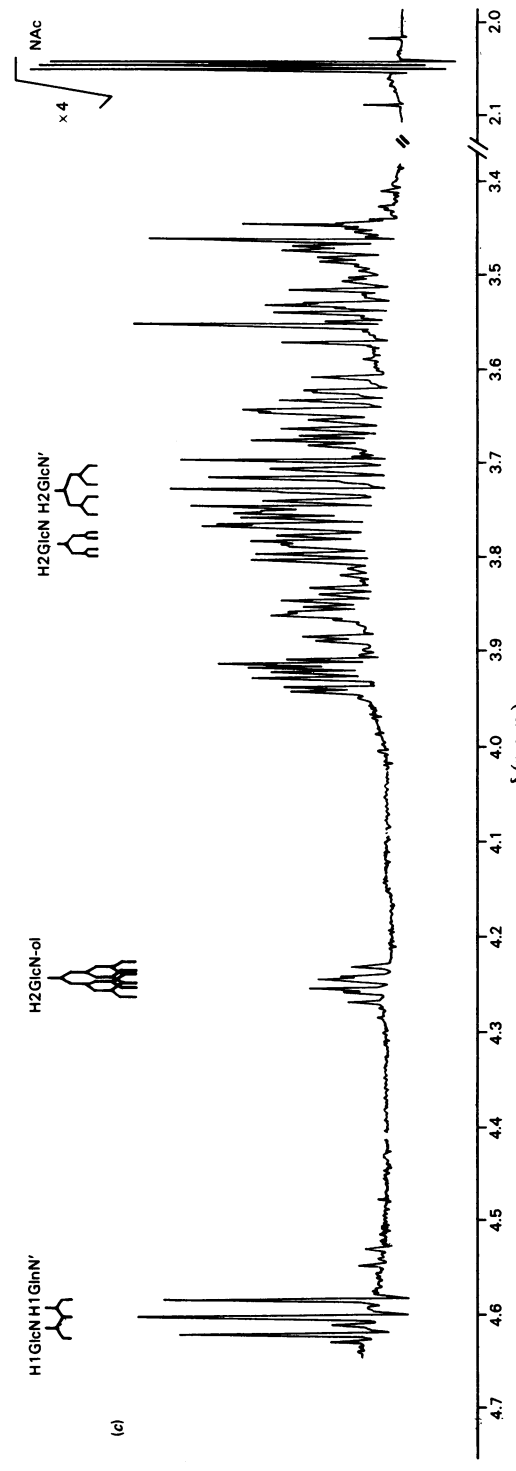


Fig. 1. *The 500 MHz proton-n.m.r. spectra of the following oligosaccharides:*

(*a*) *Galβ1–4GlcNAc-ol*
(*b*) *Galβ1–4GlcNAcβ1–3Galβ1–4GlcNAc-ol*
(*c*) *GlcNAcβ1–4GlcNAcβ1–4GlcNAc-ol*

Proton chemical shifts were assigned by spin-decoupling experiments and by comparison with other n.m.r. spectra. In (*b*) the non-reducing-end galactose residue, Gal (ext) and the internal galactose residue Gal (int) were distinguished by comparison with the spectrum of LNNT (Dua & Bush, 1983) and LNNT-ol (our studies). The two non-reducing *N*-acetylglucosamine residues in (*c*) designated as GlcNAc and GlcNAc′ are suggested, as described in the text, to arise from the non-reduced-end and internal GlcNAc residues respectively.

totri-itol were 4.611 and 4.592p.p.m., and the corresponding H2 chemical shifts (obtained by spin-decoupling and measurement of coupling constants, which were 9.5 and 8.5Hz respectively) were 3.794 and 3.741 p.p.m. These were assigned to the external (non-reducing) and internal residues respectively by comparison with chitobi-itol (GlcNAcβ1–4GlcNAc-ol; H1β 4.615p.p.m.).

Oligosaccharides of the fucosyl-lactitol series. Our studies on the alditols of LNFP I, LNFP II, LNFP III and LDFH I confirm evidence previously available in the literature that the chemical shifts of the protons of fucose residues show a great variability from one oligosaccharide to the next. To assist in characterizing such protons we have compared the n.m.r. spectra of the following fucosylated lactitol oligosaccharides with each other and that of lactitol.

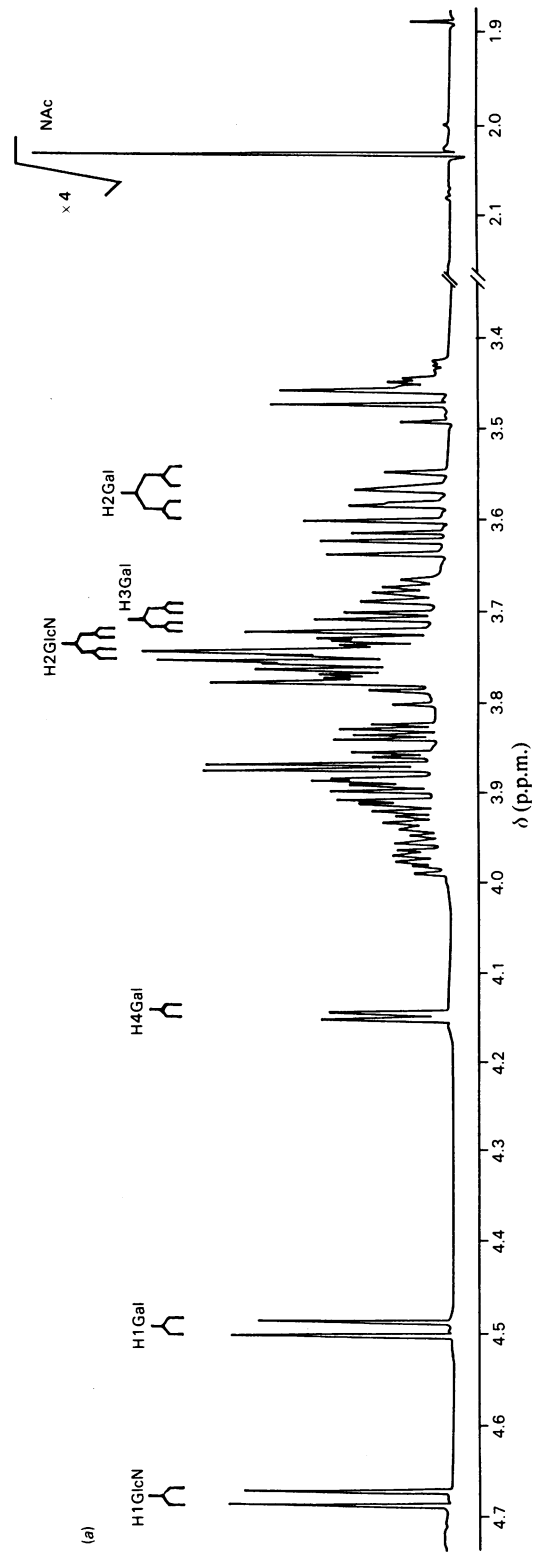
Fucα1–2Galβ1–4Glc-ol
|1,3
Fucα

Fucα1–2Galβ1–4Glc-ol
Galβ1–4Glc–ol
|1,3
Fucα

The chemical shift data for these oligosaccharide alditols are given in the sequences designated 9, 10, 22, 56, 64 and 65 (Table 3), and also contribute to the values of sequences 43 and 51. In particular, these studies have shown the substantial changes in the chemical shifts of fucose residues in the presence of (*a*) differing linkages of an adjacent galactose to the next monosaccharide along the sequence and (*b*) a second fucose residue on a non-adjacent monosaccharide as in the following examples:

Sequence designation in library	Example of sequence	Fucα1–2 H1 shift
51	{ Fucα1–2Galβ1–4Glc-ol Fucα1–2Galβ1–4GlcNAc }	5.31
64	Fucα1–2Galβ1–4Glc-ol |1,3 Fucα	5.42
51	Fucα1–2Galβ1–4GlcNAc |1,3 Fucα	5.29
66	Fucα1–2Galβ1–3GlcNAc |1,4 Fucα	5.15

Our studies also corroborate the data from other laboratories cited in Table 4 on the changes in the chemical shifts of fucose linked to different



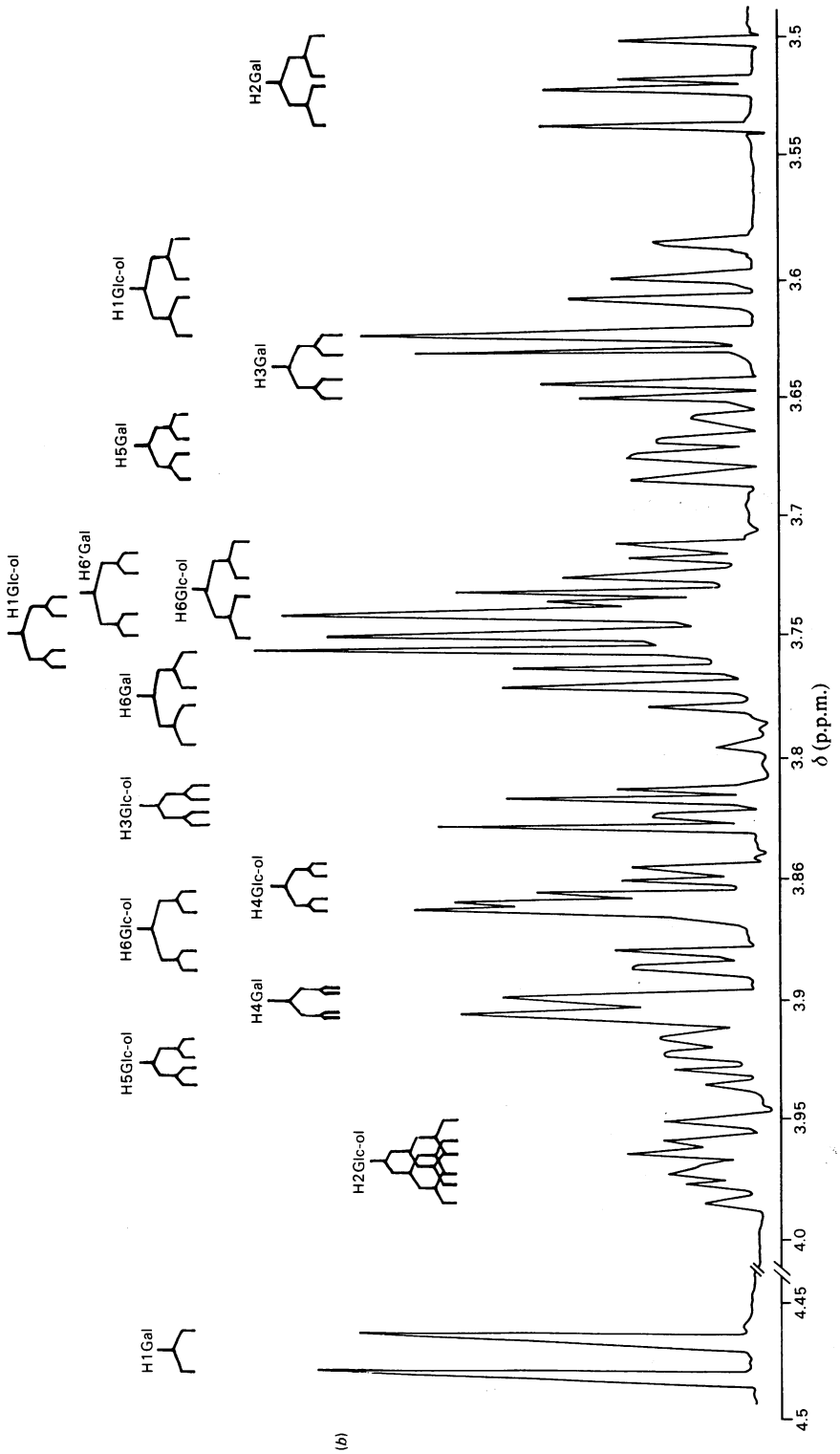


Fig. 2. *The 500 MHz n.m.r. spectra of (a) the trisaccharide GlcNAcβ1–3Galβ1–4Glc-ol and (b) lactitol (Galβ1–4Glc-ol)*

positions as in the following examples:

Sequence designation in library	Example of sequence	Fuc H1 shift
65	Fucα1–3Glc-ol	5.07
54	Fucα1–3GlcNAc	5.11
55	Fucα1–4GlcNAc	5.02
57	Fucα1–6GlcNAc	4.88

Oligosaccharides with blood-group-A, -B and -Sd^a activities. Our data on the blood-group-A and -B tetrasaccharide structures

GalNAcα1–3Galβ1–4Glc-ol (Blood group A)
|1,2
Fucα

Galα1–3Galβ1–4Glc-ol (Blood group B)
|1,2
Fucα

given in library sequence numbers 9, 24, 49, 53 and 9, 23, 48, 52 respectively (reported in full elsewhere, Yates *et al.*, 1984) are in accord with those of Lemieux *et al.* (1980). The blood-group-Sd^a determinant, like the blood-group-A, -B and -H antigens, is a genetic marker of erythrocytes and is also found in various secreted glycoproteins. Analyses of several oligosaccharides isolated from an Sd^a-active Tamm–Horsfall glycoprotein of urine by methylation analysis and sequential exoglycosidase digestion (Donald *et al.*, 1983) have established the structure of a pentasaccharide with Sd^a activity to be as follows:

GalNAcβ1–4Galβ1–4GlcNAcβ1–3Gal
|2,3
NeuAcα

Chemical shifts for the sialic acid, the *N*-acetylgalactosamine and the galactose to which they are attached, given in sequences 62, 70 and 27 respectively, are in good agreement with those reported for the structurally related Cad determinant (Blanchard *et al.*, 1983).

Interrogation of the library for oligosaccharides of unknown structure

As described in the Materials and methods section, the proton resonances of unknown oligosaccharides are compared with those in the library and a print-out is obtained of possible sequences present. As an example Tables 6–8 show the print-out for the following oligosaccharides chosen because of their different sizes and reduced-end residues:

Galβ1–4GlcNAc-ol
Galβ1–4GlcNAcβ1–6
3GalNAc-ol
Galβ1

Galβ1–4GlcNAcβ1–3Galβ1–4Glc-ol
|1,3
Fucα

In each case the library of sequences is given together with their designation, as well as the number of chemical shifts recorded in the up-to-date library for each sequence and the number of matching chemical shifts found in the 'unknown' oligosaccharide. Some of the chemical shifts of the disaccharide (Table 6) had already been entered in the library, and therefore finding the correct assignment of sequences 7 and 43 is not surprising. On the other hand, the tetrasaccharide was an oligosaccharide of unknown structure obtained from alkaline-borohydride degradation of a glycoprotein preparation of human meconium (Gooi & Feizi, 1982). The 'full-match' sequences make up two possible oligosaccharides, as shown in Table 7. Presence of the larger structure can be ruled out by inspection of the number and relative intensity of the anomeric protons in the spectrum. These together would readily show the presence of minor contaminants and the size of the major oligosaccharide component. The tetrasaccharide structure deduced in Table 7 was in accord with the number of anomeric protons. This structure was confirmed by methylation analysis and fast-atom-bombardment and electron-impact mass spectrometry (E. F. Hounsell & A. M. Lawson, unpublished work).

The pentasaccharide shown in Table 8 was one of several oligosaccharides purified (by reduction, acetylation and preparative t.l.c. as described by Hounsell *et al.*, 1981) from a tetra- to heptasaccharide fraction of human milk provided by Dr. Prieels. Structural assignment was achieved by analysis of the 'full matches' obtained. Data for these sequences were from various sources as follows: those of sequence 9 from lactitol (Fig. 2*b*), those of sequence 25 from the analysis of unreduced LNF III (Dua & Bush, 1983), those of sequence 26 from the present study on reduced LNF III, those of sequence 43 as shown in Table 5 and those of sequence 54 from the work of Vliegenthart *et al.* (1981), Hindsgaul *et al.* (1982) and Dua & Bush (1983).

Conclusion

These studies have shown that data from high-resolution n.m.r. can be compared from one

laboratory to another and from sequences within several categories of oligosaccharide structure, i.e. reducing oligosaccharides, alditols, glycosides and *N*- and *O*-linked carbohydrate chains of glycoproteins. Further, it has been shown that the library with the level of accuracy described can distinguish between oligosaccharides and provide structural assignments.

The majority of the known trisaccharide sequences found in oligosaccharides formed by mammalian biosynthesis are included in the library. New chemical-shift data, from more-extensive n.m.r. analyses of the known oligosaccharides and from new oligosaccharides with additional sequences, can be readily incorporated into the library. The program can also be extended to incorporate coupling-constant data to give additional information for proton identification.

In addition to providing an approach for the structural assignment of unknown oligosaccharides, the library of data has aided comparisons of the chemical shifts of the same monosaccharide in different oligosaccharides. Of particular interest are the variety of proton chemical shifts of fucose residues, as these residues are important structures in several known antigenic determinants, and the signals that constitute an n.m.r. spectrum reflect the solution conformations of oligosaccharides that are recognized by antibody combining sites.

We thank Dr. Ten Feizi for helpful discussions and Mrs. S. Schwarz for the preparation of this manuscript.

References

Anderson, A. & Donald, A. S. R. (1981) *J. Chromatogr.* **211**, 170–174

Aué, W. P., Karhan, J. & Ernst, R. R. (1976) *J. Chem. Phys.* **64**, 2229–2246

Augé, C., David, S. & Veyrières, A. (1979) *Nouv. J. Chim.* **3**, 491–497

Barker, S. A., Foster, A. B., Stacey, M. & Webber, J. (1958) *J. Chem. Soc.* 2218–2226

Blanchard, D., Cartron, J.-P., Fournet, B., Montreuil, J., Van Halbeek, H. & Vliegenthart, J. F. G. (1983) *J. Biol. Chem.* **258**, 7691–7695

Bock, K., Arnarp, J. & Lonngren, J. (1982) *Eur. J. Biochem.* **129**, 171–178

Brisson, J.-R. & Carver, J. P. (1983) *Biochemistry* **22**, 3673–3680

Dabrowski, J., Hanfland, P. & Egge, H. (1982) *Methods Enzymol.* **83D**, 69–88

Dabrowski, U., Egge, H. & Dabrowski, J. (1983) *Arch. Biochem. Biophys.* **224**, 254–260

Donald, A. S. R., Yates, A. D., Soh, C. P. C., Morgan, W. T. J. & Watkins, W. M. (1983) *Biochem. Biophys. Res. Commun.* **115**, 625–631

Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J.-C., Fournet, B., Spik, G. & Montreuil, J. (1978) *Eur. J. Biochem.* **87**, 323–329

Dua, V. K. & Bush, C. A. (1983) *Anal. Biochem.* **133**, 1–8

Ferrige, A. G. & Lindon, J. C. (1979) *J. Magn. Reson.* **31**, 337–340

Gooi, H. C. & Feizi, T. (1982) *Biochem. Biophys. Res. Commun.* **106**, 539–545

Hindsgaul, O., Norberg, T., Le Pendu, J. & Lemieux, R. U. (1982) *Carbohydr. Res.* **109**, 109–142

Hounsell, E. F., Gooi, H. C. & Feizi, T. (1981) *FEBS Lett.* **131**, 279–282

Kobata, A. (1972) *Methods Enzymol.* **28**, 262–271

Koerner, T. A. W., Prestegard, J. H., Demou, P. C. & Yu, R. K. (1983) *Biochemistry* **22**, 2676–2687

Lemieux, R. U., Bock, K., Delbaere, L. T. J., Koto, S. & Rao, V. S. (1980) *Can. J. Chem.* **58**, 631–653

Race, R. R. & Sanger, R. (1975) *Blood Groups in Man*, 6th edn., pp. 395–405, Blackwell, Oxford

Race, C. & Watkins, W. M. (1970) *FEBS Lett.* **19**, 279–283

Van Halbook, H., Dorland, L., Haverkamp, J., Veldink, G. A., Vliegenthart, J. F. G., Fournet, B., Ricart, G., Montreuil, J., Gathman, W. P. & Aminoff, D. (1981) *Eur. J. Biochem.* **118**, 487–495

Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Hull, W. E., Lamblin, G., Lhermitte, M., Boersma, A. & Roussel, P. (1982) *Eur. J. Biochem.* **127**, 7–20

Van Halbeek, H., Gerwig, G. J., Vliegenthart, J. F. G., Smits, H. L., Van Kerkoff, P. J. M. & Kramer, M. F. (1983) *Biochim. Biophys. Acta* **747**, 107–116

Veyrières, A. (1981) *J. Chem. Soc. Perkin Trans.* **1**, 1626–1629

Vliegenthart, J. F. G., Van Halbeek, H. & Dorland, L. (1981) *Pure Appl. Chem.* **53**, 45–77

Yates, A. D. & Watkins, W. M. (1983) *Carbohydr. Res.* **120**, 251–268

Yates, A. D., Feeney, J., Donald, A. S. R. & Watkins, W. M. (1984) *Carbohydr. Res.* **130**, 251–260