Eur. J. Biochem. *127*, 21–29 (1982)

Characterization of the Primary Structure and the Microheterogeneity of the Carbohydrate Chains of Porcine Blood-Group H Substance by 500-MHz ^{1}H-NMR Spectroscopy

Herman VAN HALBEEK, Lambertus DORLAND, Johannes F. G. VLIEGENTHART, Nicolai K. KOCHETKOV, Nicolai P. ARBATSKY, and Varvara A. DEREVITSKAYA

Department of Bio-Organic Chemistry, University of Utrecht; and N. D. Zelinsky Institute of Organic Chemistry, Academy of Sciences, Moscow

(Received May 10, 1982)

In blood-group H substance from pig stomach cell linings, carbohydrate structures are known to occur which are *O*-glycosidically linked via GalNAc to Ser or Thr of the polypeptide backbone. They have in common the Gal(β1→3)GalNAc core unit, which bears one or more *N*-acetyllactosamine branches. The latter are terminated either by Fuc in (α1→2) linkage or by GlcNAc in (α1→4) linkage to Gal. Both terminal sequences are considered to be porcine blood-group H antigenic determinants.

Employing 500-MHz ^{1}H-NMR spectroscopy, the spectral features of the oligosaccharide-alditols corresponding to these chains were established. Insight could be gained into the microheterogeneity displayed by some of the alditol fractions as to the distribution of the terminating Fuc and GlcNAc residues over the various branches. Such information can hardly be obtained using other approaches.

500-MHz ^{1}H-NMR spectroscopy is one of the most powerful methods in the primary structural analysis of the carbohydrate chains of glycoproteins. So far, this technique was most successfully applied for the examination of the *N*-glycosidic types of glycoprotein oligosaccharide chains [1–6]. Recently, it was employed also for the characterization of *O*-glycosidic carbohydrate chains of the mucin type, like acidic oligosaccharide units from cow milk κ-casein [7,8], neutral oligosaccharides from human bronchial mucin [9], and neutral and acidic carbohydrate chains from porcine submaxillary-mucin glycoproteins exhibiting blood-group A or H activity [7,10]. In this paper we report the 500-MHz ^{1}H-NMR spectral parameters of the carbohydrate units of blood-group H substance from pig stomach cell linings. As has been established for *N*-glycosidic carbohydrate chains [7,11–13], the ^{1}H-NMR approach at 500 MHz can disclose microheterogeneity in mixtures of closely related components which is almost untraceable using other analytical routes. It will be shown that this also holds for *O*-glycosidic, mucin-type chains.

MATERIALS AND METHODS

Oligosaccharide-alditols derived from blood-group H substance were obtained from pig stomach cell linings as described previously [14–17]. Their structures were established by combination of methylation analysis, periodate oxidation, Smith degradation, chromium trioxide oxidation, and digestion with various exoglycosidases [15,16]. In addition, for a few of them ^{13}C-NMR spectroscopic analysis was conducted [18].

Six of the oligosaccharide-alditol fractions isolated were subjected to 500-MHz ^{1}H-NMR spectroscopy. After dissolution of the deuterium-exchanged samples in ^2H_2O, the spectra were recorded using a Bruker WM-500 spectrometer operating at a probe temperature of 300 K. Experimental details were described previously [6,7].

RESULTS AND DISCUSSION

500-MHz ^{1}H-NMR spectroscopy was performed for six oligosaccharide-alditol fractions derived from porcine blood-group H substance. The structures of the alditols are given in Fig. 1, arranged in order of increasing complexity. They are representative of the structural divergency known to occur in the carbohydrate moieties of the blood-group H active glycoproteins in pig stomach cell linings [16,17]. The spectra of fractions 11.1, 14.5, and 16.2 are depicted in Fig. 2, 3, and 4, respectively; that of the trisaccharide-alditol OS3 has been published previously [7]. The chemical shifts of the structural-reporter groups [7] bearing the essential information on the primary structure of the oligosaccharide-alditols from blood-group H substance are compiled in Table 1. For the spectral interpretation, advantage was taken from the ^{1}H-NMR data established for a series of reference compounds. The structures of the latter are listed in Fig. 5; their spectral data are listed in Table 2.

The common presence of the Gal(β1→3)GalNAc-ol unit in the oligosaccharide-alditols from blood-group H substance, as well as in reference compounds 1–5, is reflected by the chemical shift for H-2 of GalNAc-ol: $\delta \approx 4.39$ ppm (see Tables 1 and 2). According to our proposal [7–10], this parameter is indeed hardly sensitive to any elongation of

Abbreviations. Fuc, L-fucose; Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; GalNAc-ol, *N*-acetyl-D-galactosaminitol; GlcNAc, *N*-acetyl-D-glucosamine; NMR, nuclear magnetic resonance; Ser, L-serine; Thr, L-threonine.

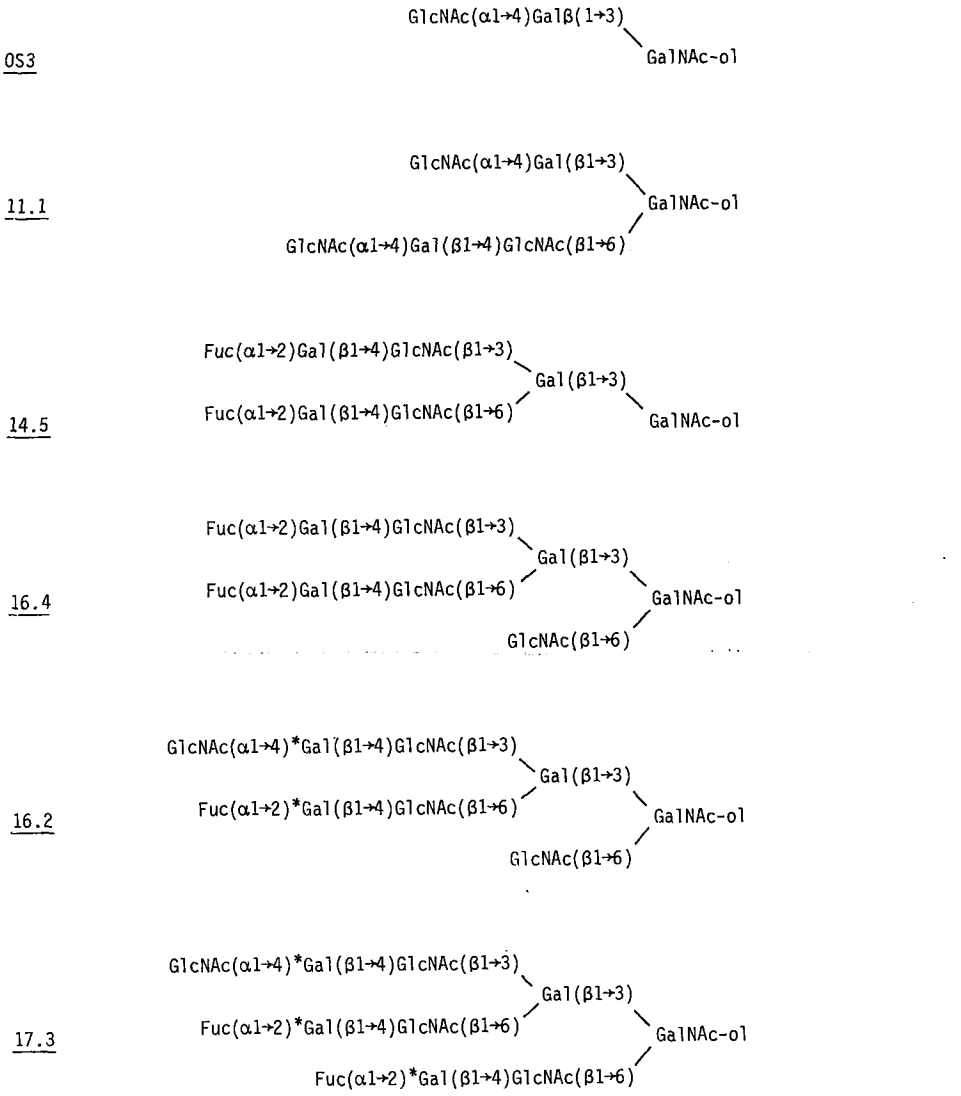


Fig. 1. *Structures of oligosaccharide-alditols obtained from porcine blood-group H substance which were used for 500-MHz ^{1}H-NMR studies.* For reasons of consistency, the fractions are designated according to the system previously introduced [14–16]. For fractions 16.2 and 17.3, the present study revealed microheterogeneity as to the distribution of the terminal residues over the branches. Therefore, residues marked by an asterisk may be interchanged

the Gal(β1→3)GalNAc-ol portion, either at Gal or at GalNAc-ol.

From the spectrum of the trisaccharide-alditol OS3, the ^{1}H-NMR spectral features typical of the GlcNAc(α1→4)Gal(β1→·) sequence can be readily deduced. The terminal GlcNAc residue is characterized by four structural-reporter groups: its H-1, H-5, and H-4 atoms, and its *N*-acetyl methyl protons. The chemical shift of H-1 ($\delta \approx 4.87$ ppm), in conjunction with its coupling constant ($J_{1,2} = 4.0$ Hz), is indicative of an α-glycosidic linkage between GlcNAc and Gal. This is in accord with conclusions from ^{13}C-NMR studies [16, 18]. The GlcNAc H-5 resonance at $\delta \approx 4.18$ ppm is obscured, because of partial overlap of this multiplet with the GalNAc-ol H-5 signal ($\delta \approx 4.19$ ppm). The observation of its pattern is not hampered if GalNAc-ol bears a substituent at C-6 (*vide infra*). The chemical shift of GlcNAc H-4 is practically identical with that of GalNAc-ol H-4; however, the resonance patterns from these two protons are quite distinct, owing to the significant differences in their vicinal coupling constants (for GlcNAc, $J_{3,4} = 8.9$ Hz; $J_{4,5} = 10.3$ Hz; for GalNAc-ol, $J_{3,4} = 8.8$ Hz; $J_{4,5} = 1.5$ Hz). The H-4 resonances together give rise to a symmetrical assembly of eight lines. It should be noted that the assignments of the GlcNAc skeleton proton signals were proved by double-resonance, spin-tickling experiments. The *N*-acetyl singlet of GlcNAc is found at rather low-field position ($\delta \approx 2.09$ ppm). The introduction of GlcNAc in (α1→4) linkage to Gal causes downfield shifts of the Gal structural-reporter groups (compare with alditol 1, Table 2): for H-1, the shift increment ($\Delta\delta$) ≈ 0.05 ppm, and for H-4, $\Delta\delta \approx 0.07$ ppm. The chemical shifts of GalNAc-ol H-3 and H-4 and of the *N*-acetyl protons are slightly but significantly affected (compare also compounds 2 and 4, see Table 2).

For the spectral interpretation, compound 11.1 can be conceived as an extension of alditol 3 with two terminating GlcNAc residues in (α1→4) linkage to Gal. The presence of GlcNAc in (β1→6) linkage to GalNAc-ol in fraction 11.1 is reflected by the differences in the chemical shift of GalNAc-ol

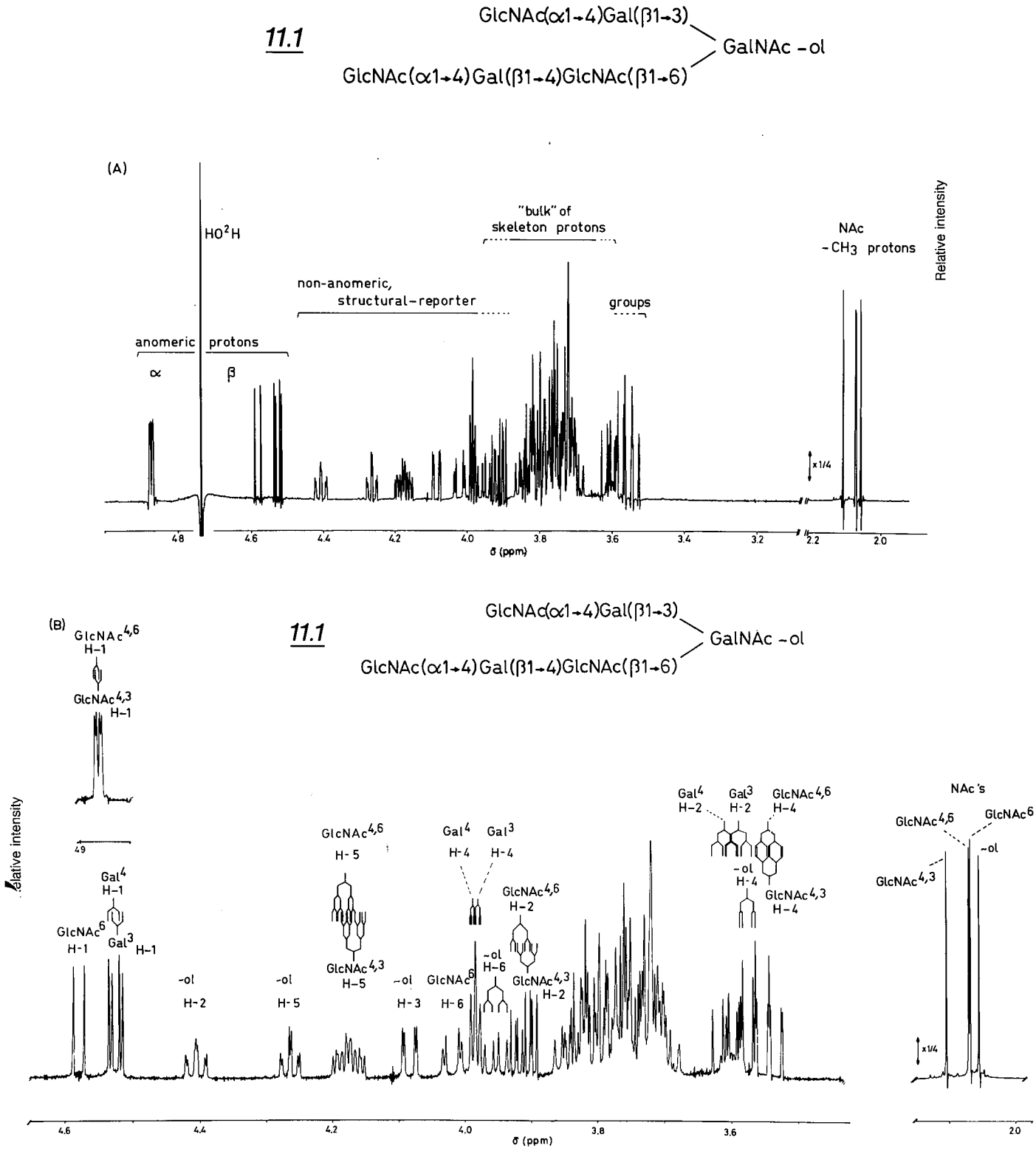


Fig. 2. *Resolution-enhanced 500-MHz ^{1}H-NMR spectrum of hexasaccharide-alditol 11.1 from porcine blood-group H substance, in ^2H_2O at 300 K.* (A) Overall spectrum. (B) Expanded, structural-reporter-group regions of the spectrum. For designation of the residues, see Table 1. The relative-intensity scale of the *N*-acetyl proton region deviates from that of the other parts of the spectrum, as indicated

H-5 and H-6, as compared to the corresponding protons in fraction OS3 and alditol 1 [8,9]. The influences of introduction of the α-linked GlcNAc residues upon the chemical shifts of the Gal^3 and Gal^4 reporter groups are independent of each other. (For designation of residues, see Tables 1 and 2.) This is illustrated by the difference in chemical shift between H-1 of Gal^3 and Gal^4 ($\Delta\delta = -0.005$ ppm) being maintained as compared to compound 3. Moreover, the effects of attachment of α-linked GlcNAc are completely analogous to those described for the step from compound 1 to oligosaccharide

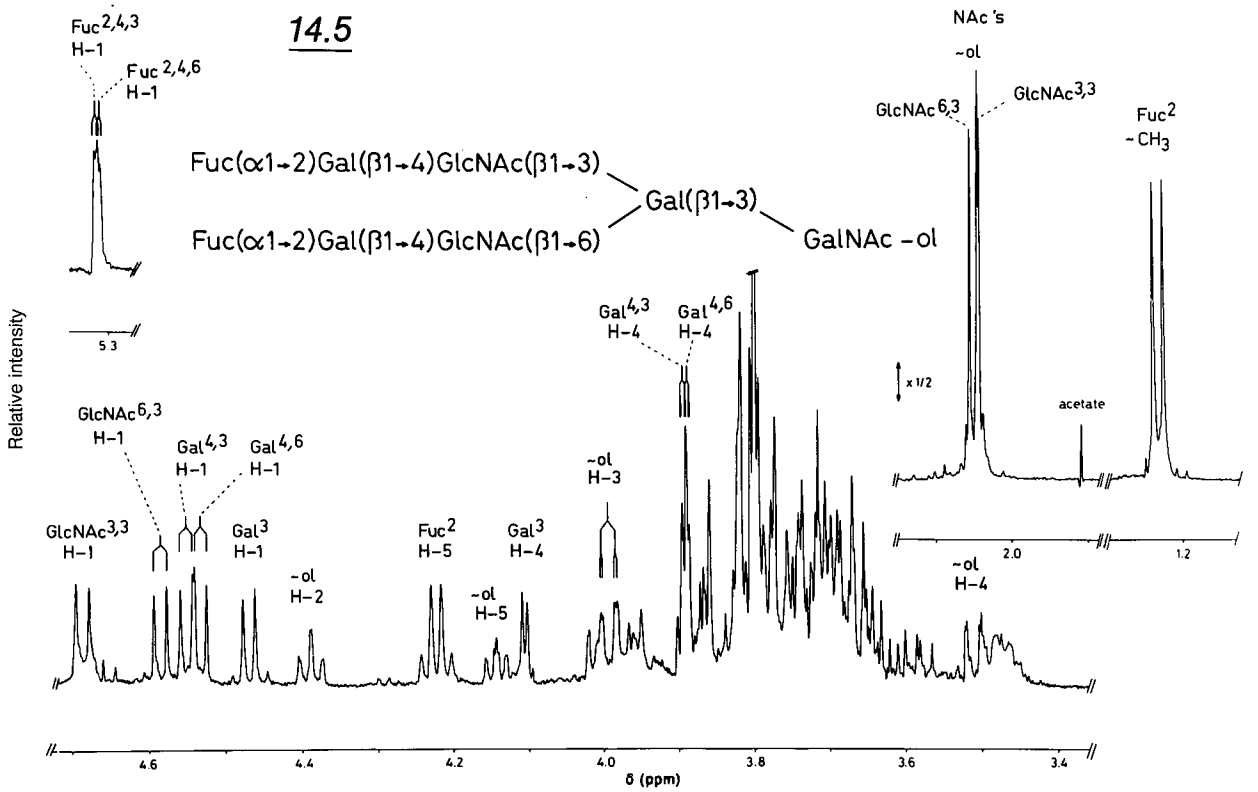


Fig. 3. *Structural-reporter-group regions of the resolution-enhanced 500-MHz ^1H-NMR spectrum of octasaccharide-alditol 14.5 from porcine blood-group H substance in ^2H_2O at 300 K.* For designation of the residues, see Table 1. The relative-intensity scale of the *N*-acetyl and Fuc methyl proton regions (insertion) deviates from that of the other parts of the spectrum, as indicated

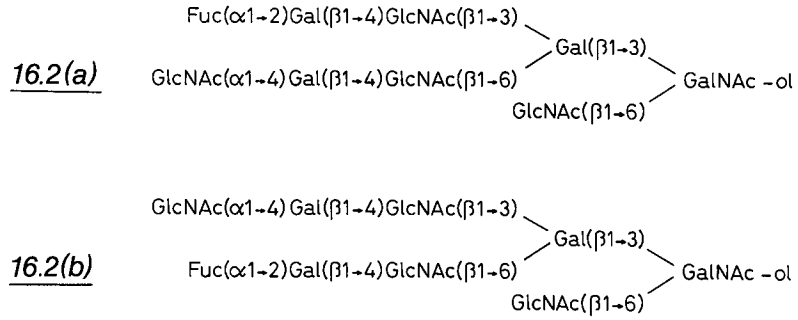
OS3. In addition, the H-1 signal of GlcNAc^6 is shifted slightly downfield ($\Delta\delta \approx 0.02$ ppm) due to introduction of (α1→4)-linked GlcNAc into this branch.

Each of the terminal, α-linked GlcNAc residues in fraction 11.1 gives rise to its own set of H-1, H-5, and H-4 signals at the aforementioned characteristic positions in the spectrum (see Fig. 2). The assignments of the signals within each pair to GlcNAc in a particular branch (GlcNAc^{4,3} or GlcNAc^{4,4,6}, respectively) are based upon comparison with the spectral data of compound 4 (see Table 2), lacking part of the lower branch as compared to fraction 11.1. The relative positions of the α-GlcNAc reporter-group signals are in line with the observation that the reporter groups of the upper-branch Gal^3 residue resonate at slightly higher field than the corresponding ones of the lower-branch Gal^4 residue (compare compound 3, Table 2). The application of homonuclear *J*-resolved two-dimensional (2D-*J*) ^{1}H-NMR spectroscopy to fraction 11.1 enabled the unraveling of the total spectrum, thereby confirming the aforementioned assignments (unpublished results).

Four well-resolved singlets occur in the *N*-acetyl region of the spectrum of fraction 11.1 (see Fig. 2). The signals at $\delta = 2.104$ and 2.054 ppm are ascribed to GlcNAc (α1→4)-linked to Gal^3, and to GalNAc-ol, respectively. These assignments are based upon comparison with compound 4. The typical shift alteration observed for the *N*-acetyl signal of GlcNAc^4 in the upper branch as compared to fraction OS3 is caused by the attachment of GlcNAc in (β1→6) linkage to GalNAc-ol (see compound 4, Table 2). The remaining singlets at $\delta = 2.069$ and 2.067 ppm are assigned to the lower-branch GlcNAc^4 and GlcNAc^6, respectively, assuming no significant effect of introduction of GlcNAc^4 on the chemical shift of the *N*-acetyl signal of GlcNAc^6 (compare with compound 3).

The 500-MHz ^{1}H-NMR spectrum of fraction 14.5, depicted in Fig. 3, reveals the characteristics of an oligosaccharide possessing Gal^3 as branching point. The Gal H-4 atom serves as the reporter group for the 3,6-disubstitution of Gal^3. This can be deduced from comparison of the data of fraction 14.5 (Table 1) with those of compounds 1 and 5, and also with compounds 6–8 (Table 2). The aforementioned H-4 atom has previously been established to be the reporter group for (β1→3) substitution of Gal by GlcNAc [6,9] (compare compound 5 with compound 1; for Gal^3 H-4, $\Delta\delta \approx 0.23$ ppm) or by Gal [19]. Assuming independency of the influences of attachment of the (β1→3)-linked and the (β1→6)-linked *N*-acetyllactosamine units on the chemical shifts of the reporter groups of Gal^3, it may be concluded from comparison of fraction 14.5 with compound 5 that the introduction of the (β1→6)-linked branch causes an additional shift alteration for the Gal^3 H-4 atom ($\Delta\delta \approx -0.02$ ppm). Moreover, extension of Gal^3 in compound 5 with the (β1→6)-linked *N*-acetyllactosamine unit causes considerable shift effects on some GalNAc-ol protons, e.g. on H-3 ($\Delta\delta \approx -0.06$ ppm) and on H-5 ($\Delta\delta \approx -0.04$ ppm).

The H-1 signal of GlcNAc^3 is observed at $\delta = 4.685$ ppm ($J_{1,2} = 8.5$ Hz) (compare with compound 5), that of GlcNAc^6 at $\delta = 4.584$ ppm ($J_{1,2} = 8.3$ Hz). The remaining two doublets ($J_{1,2} = 7.8$ Hz) in the β-anomeric region of the spectrum belong to the Gal^4 residues. The signal at $\delta = 4.550$ ppm is assigned to Gal^{4,3}, because of its shift increment in comparison to Gal^{4,3} in compound 5 ($\Delta\delta \approx 0.07$ ppm), which is typical of substitution of a Gal^4 by an (α1→2)-linked Fuc residue [6,7,9,10,19]. Consequently, the signal at $\delta = 4.532$ ppm is



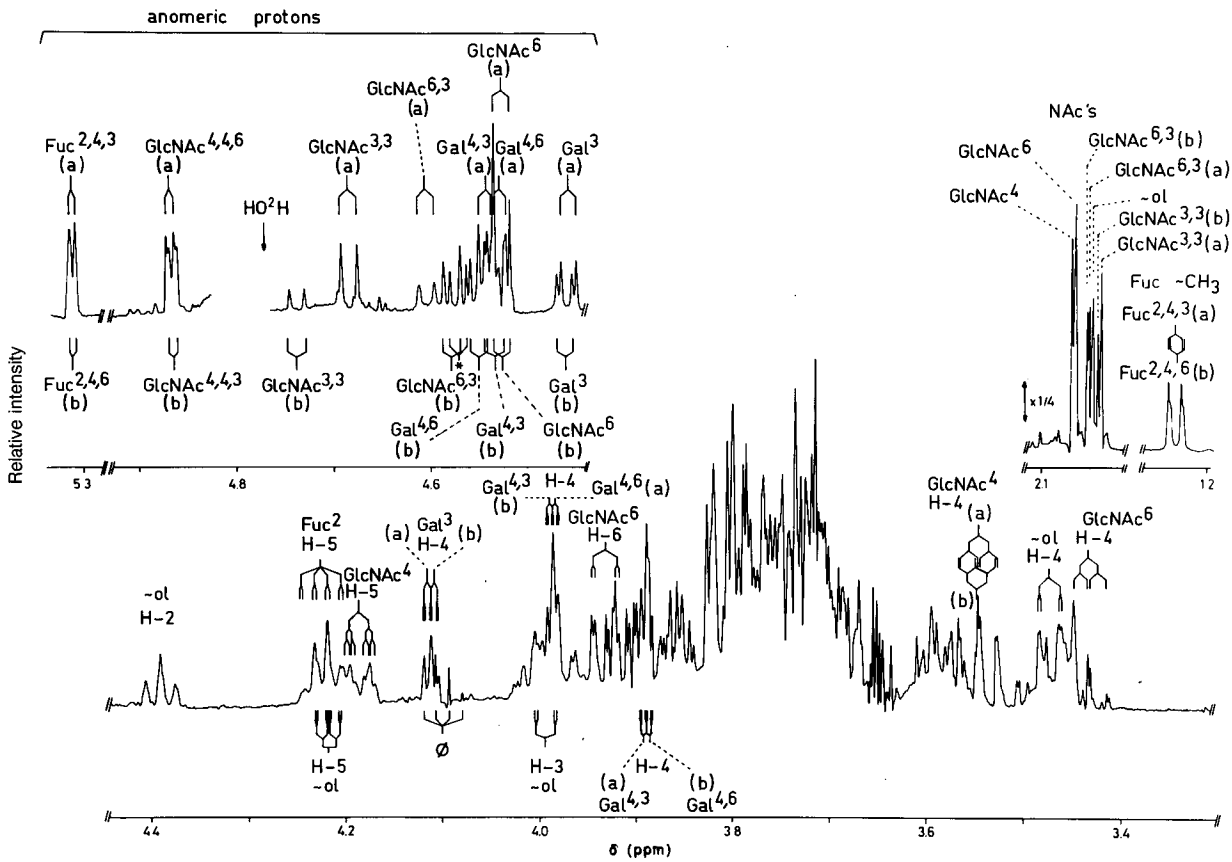


Fig. 4. *Structural-reporter-group regions of the resolution-enhanced 500-MHz ^1H-NMR spectrum of the mixture of nonasaccharide-alditols 16.2(a) and 16.2(b) (ratio 4:3), obtained from porcine blood-group H substance, in ^2H_2O at 300 K.* The HO^2H resonance has been omitted. For designation of the residues, see Table 1. Signals of corresponding protons in the two components coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl and Fuc methyl proton regions (insertion) deviates from that of the other parts of the spectrum, as indicated. The quartet at $\delta \approx 4.11$ ppm marked by ϕ, originates from a frequently occurring non-protein, non-carbohydrate contaminant of unknown structure. The sample is contaminated with small amounts of oligosaccharide-alditols of unidentified structure; this gives rise *inter alia* to the signal marked by an asterisk

attributed to $Gal^{4,6}$. The assignment of the three singlets in the *N*-acetyl region of the spectrum of fraction 14.5 is consistent with that of the *N*-acetyl signals in the spectra of compounds 6–8 (see Table 2) and of partial structures thereof (unpublished results).

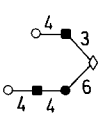
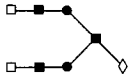
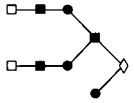
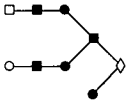
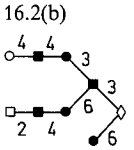
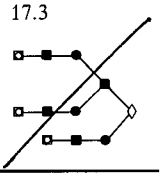
Apart from the effects on the chemical shifts of the reporter groups of the neighbouring Gal^4 residues, the set of chemical shifts for the Fuc structural-reporter groups themselves (δ H-1 $\approx$ 5.31 ppm; δ H-5 $\approx$ 4.22 ppm; δ CH_3 $\approx$ 1.23 ppm) is indicative of the ($\alpha 1 \rightarrow 2$) linkage between Fuc and Gal [6, 7, 9, 10]. The H-1 doublets of the Fuc residues in fraction 14.5 are separated (see Fig. 3); their assignment is based on the trend observed for the positions of anomeric proton signals of upper-branch residues, having more downfield positions than the corresponding signals of lower-branch residues (see Table 1 and Fig. 3).

Compound 16.4 can be conceived as an extension of compound 14.5 with GlcNAc in ($\beta 1 \rightarrow 6$) linkage to GalNAc-ol. This terminal $GlcNAc^6$ residue is characterized by its H-1 and *N*-acetyl signals. The chemical shift of the $GlcNAc^6$ H-1 atom (δ = 4.526 ppm) in compound 16.4 differs from that observed for compounds 2 and 4 (see Table 2). This may be ascribed to the presence of the bulky, di-antennary part of the structure at Gal^3. However, the chemical shift of the *N*-acetyl signal of $GlcNAc^6$ (δ = 2.066 ppm) is insensitive to the branching of Gal^3 with fucosylated *N*-acetyllactosamine units.

The attachment of GlcNAc to GalNAc-ol in ($\beta 1 \rightarrow 6$) linkage is also evidenced by the chemical shift of GalNAc-ol

Table 1. *^1H chemical shifts of structural-reporter-group protons of constituent monosaccharides for some oligosaccharide-alditols obtained from porcine blood-group H substance*

Chemical shifts are given at 300 K, relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O (but were actually measured relative to internal acetone: $\delta = 2.225$ ppm). For complete structures of the compounds, see Fig. 1. In the table heading the structures are represented by short-hand symbolic notation: ◇ = GalNAc-ol; ■— = Gal; ●— = β-linked GlcNAc; ○— = α-linked GlcNAc; □— = Fuc; ⊡— = α-linked GlcNAc or Fuc. The (first) superscript at the name of a sugar residue indicates the type of glycosidic linkage in which it is involved; any more superscripts indicate the type(s) of the next linkage(s) in the sequence containing the residue concerned. Types of linkage that figure in this series of compounds are given in schematic structures 11.1 and 16.2(b). n.d. means that the various shifts for this proton could not be determined precisely due to the complexity of the mixture; (f) indicates clearly observable shift that corresponds with Fuc being present in the terminal position of this branch

Residue	Reporter group	Chemical shift in oligosaccharide-alditols						
		OS3	11.1	14.5	16.4	16.2(a)	16.2(b)	17.3
								
		ppm						
GalNAc-ol	H-2	4.404	4.405	4.386	4.392	4.392	4.392	4.390
	H-3	4.081	4.083	3.991	3.993	3.994	3.994	n.d.
	H-4	3.541	3.574	3.509	3.472	3.474	3.474	n.d.
	H-5	4.193	4.264	4.142	4.219	4.218	4.218	≈4.22
	H-6	3.76	3.952	3.8	3.9	3.9	3.9	n.d.
	NAc	2.058	2.054	2.045	2.046	2.047	2.047	2.048
Gal^3	H-1	4.525	4.521	4.468	4.457	4.457	4.461	≈4.46
	H-4	3.971	3.980	4.104	4.103	4.115	4.108	≈4.10
$GlcNAc^6$	H-1	–	4.579	–	4.526	4.526	4.526	4.579(f)
	H-6	–	4.019	–	3.931	3.932	3.932	n.d.
	NAc	–	2.067	–	2.066	2.065	2.065	n.d.
$GlcNAc^{3,3}$	H-1	–	–	4.685	4.681	4.684	4.722	4.687(f)
	NAc	–	–	2.043	2.041	2.039	2.042	n.d.
$GlcNAc^{6,3}$	H-1	–	–	4.584	4.577	4.604	4.578	n.d.
	NAc	–	–	2.055	2.051	2.051	2.053	n.d.
$Gal^{4,3}$	H-1	–	–	4.550	4.551	4.542	4.533	n.d.
	H-4	–	–	3.890	3.892	3.896	3.990	n.d.
$Gal^{4,6}$	H-1	–	4.526	4.532	4.529	4.529	4.549	4.529(f)
	H-4	–	3.987	3.888	3.888	3.983	3.892	n.d.
$GlcNAc^{4,(4,)3}$	H-1	4.869	4.870	–	–	–	4.866	4.868
	H-4	3.542	3.542	–	–	–	3.545	n.d.
	H-5	4.183	4.182	–	–	–	4.185	≈4.19
	NAc	2.089	2.104	–	–	–	2.069	2.069
$GlcNAc^{4,(4,)6}$	H-1	–	4.873	–	–	4.869	–	4.871
	H-4	–	3.544	–	–	3.548	–	n.d.
	H-5	–	4.167	–	–	4.185	–	≈4.19
	NAc	–	2.069	–	–	2.069	–	2.069
$Fuc^{2,4,3}$	H-1	–	–	5.313	5.311	5.312	–	5.311
	H-5	–	–	4.221	4.225	4.226	–	≈4.22
	CH_3	–	–	1.230	1.232	1.233	–	1.234
$Fuc^{2,4,6}$	H-1	–	–	5.310	5.311	–	5.312	5.311
	H-5	–	–	4.221	4.225	–	4.226	≈4.22
	CH_3	–	–	1.230	1.232	–	1.231	1.231

H-5 ($\delta \approx 4.22$ ppm). The deviation of this value from that observed so far in structures possessing the GlcNAc(β1→6)-GalNAc-ol element ($\delta \approx 4.28$ ppm, compare compounds 2–4 in Table 2) [8, 9] is accounted for by the presence of the *N*-acetyllactosamine unit in (β1→6) linkage to Gal^3 (see compound 14.5). Mutually, the chemical shift of H-1 of Gal^3 is slightly affected by introduction of $GlcNAc^6$ ($\Delta\delta \approx -0.01$ ppm, as compared to compound 14.5).

The spectrum of fraction 16.2 (see Fig. 4) shows doubling of some structural-reporter group signals, in the intensity ratio of 4:3, thereby giving rise to nearly twice as many anomeric doublets as would be expected on the basis of the structure previously proposed for this fraction (see Fig. 1). Also the number of *N*-acetyl singlets in the spectrum is larger than that of the amino sugars in the aforementioned structure. These features indicate heterogeneity of the sample. Comparison of the spectra of fractions 16.4 and 16.2 shows that the structural-reporter group signals of GalNAc-ol and $GlcNAc^6$ have the same chemical shifts. Moreover, they are not doubled in the spectrum of fraction 16.2, therefore being of the highest relative intensity in their spectral regions.

Apparently, the doubling is restricted to reporter-group signals of peripheral residues, which leads to the assumption that microheterogeneity occurs in the sample with respect to the location of α-GlcNAc and Fuc in the two branches. A clue to this feature is given by the presence of two H-1 signals for (α1→4)-linked GlcNAc, namely at $\delta = 4.869$ and 4.866 ppm, as well as of two Fuc CH_3-group doublets at $\delta = 1.233$

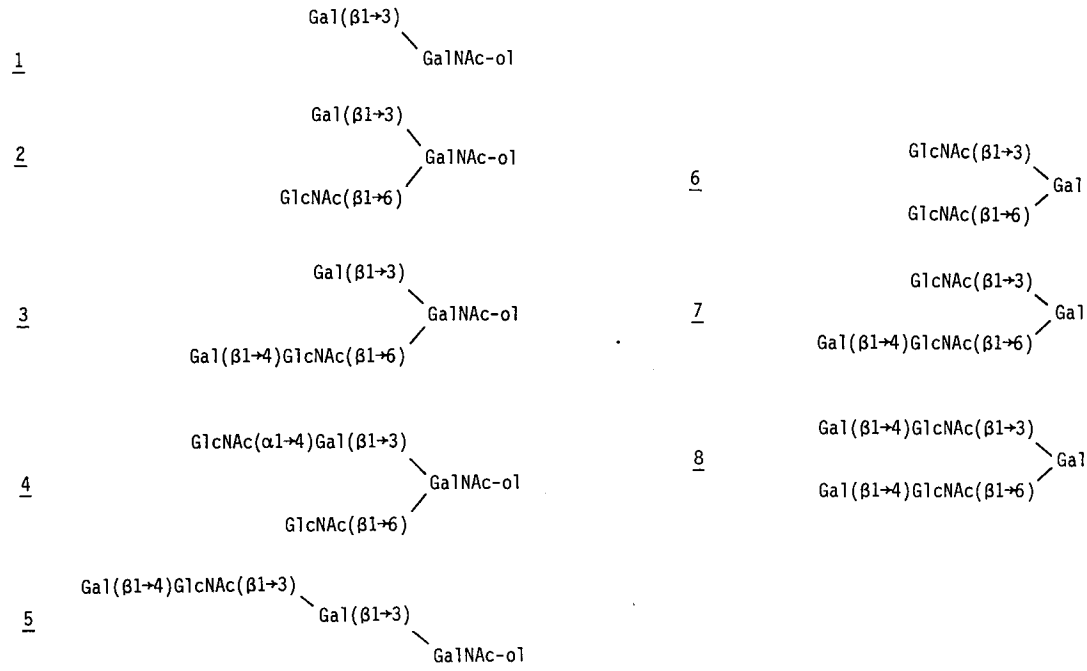


Fig. 5. *Structures of mucin-type oligosaccharide-alditols (1–5) and of related oligosaccharides (6–8), used as reference compounds*

Table 2. *^1H chemical shifts of structural-reporter-group protons of constituent monosaccharides for oligosaccharide-alditols (1–5) and oligosaccharides (6–8), used as reference compounds*

Chemical shifts are given at 300 K, relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O (but were actually measured relative to internal acetone: $\delta = 2.225$ ppm). For complete structures of the compounds, see Fig. 5. In the table heading, the structures are represented by short-hand symbolic notation: ◇ = GalNAc-ol; ■— = Gal; ●— = β-linked GlcNAc; ○— = α-linked GlcNAc. The superscript at the name of a sugar residue indicates the type of glycosidic linkage in which it is involved. Types of linkage that figure in this series of compounds are given in schematic structures 3 and 8. For comparison with 1–5, possessing Gal[3] in β-linkage to GalNAc-ol, only the shifts for the β anomers of 6–8 are given

Residue	Reporter group	Chemical shift in oligosaccharides							
		1	2	3	4	5	6β	7β	8β
		ppm							
GalNAc-ol	H-2	4.395	4.395	4.394	4.405	4.400	–	–	–
	H-3	4.065	4.061	4.060	4.081	4.051	–	–	–
	H-4	3.507	3.468	3.465	3.451	3.497	–	–	–
	H-5	4.196	4.281	4.282	4.261	4.185	–	–	–
	H-6	3.69	3.931	3.931	3.90–3.95	3.7	–	–	–
	NAc	2.050	2.066	2.067	2.054	2.047	–	–	–
Gal[3](Galβ)	H-1	4.478	4.468	4.465	4.518	4.464	4.531	4.528	4.530
	H-4	3.901	3.901	3.900	3.976	4.126	4.111	4.105	4.112
GlcNAc[6]	H-1	–	4.538	4.560	4.546	–	–	–	–
	NAc	–	2.066	2.064	2.067	–	–	–	–
GlcNAc[3,β]	H-1	–	–	–	–	4.688	4.686	4.683	4.706
	NAc	–	–	–	–	2.042	2.037	2.034	2.032
GlcNAc[6,β]	H-1	–	–	–	–	–	4.559	4.579	4.581
	NAc	–	–	–	–	–	2.043	2.041	2.040
Gal[4]	H-1	–	–	4.470	–	4.481	–	4.466	4.477[a] 4.468
	H-4	–	–	3.925	–	3.927	–	3.920	3.923[b]
GlcNAc[4]	H-1	–	–	–	4.868	–	–	–	–
	H-4	–	–	–	3.476	–	–	–	–
	H-5	–	–	–	4.174	–	–	–	–
	NAc	–	–	–	2.103	–	–	–	–

[a] Belonging to Gal[4,3] and Gal[4,6], respectively.
[b] Signal from two protons.

and 1.231 ppm; the intensity ratio within each of the pairs is that of 4:3.

In the spectral region of the β-anomeric doublets, the H-1 signals of the branching Gal[3] residue are readily recognizable at $\delta = 4.457$ and 4.461 ppm (ratio 4:3). The remaining eight doublets can be divided into two groups on the basis of their relative intensity. Within such a group, they can be attributed to a Gal or a GlcNAc residue on the basis of their $J_{1,2}$ value (for Gal, $J_{1,2} < 8$ Hz; for GlcNAc, $J_{1,2} > 8$ Hz), as has been demonstrated for fraction 14.5. Comparison with the data for the reporter groups of the fucosylated upper and lower branches in fraction 16.4 finally leads to the interpretation given in Table 1 and illustrated in Fig. 4. A similar procedure can be applied for the unraveling of the *N*-acetyl region, taking into consideration that the *N*-acetyl signals of the (α1→4)-linked GlcNAc residues in the two components coincide, at $\delta = 2.069$ ppm, despite their difference in location. The remaining singlets are assigned on the basis of their relative intensities and on comparison with the spectrum of fraction 16.4. With regard to the remaining structural-reporter groups, the doubling of the Gal[3] H-4 signal should be mentioned. Its position is typical of the branching pattern of Gal[3].

Assuming additivity of the shift effects brought about on the reporter groups of neighbouring residues by attaching Fuc in (α1→2) and GlcNAc in (α1→4) linkages, the influences of the latter type of termination can be deduced from comparison of the spectral data of compounds 16.2 with those of 16.4, 7 and 8 (see Tables 1 and 2). These effects are in agreement with those described for the steps from compounds 1 to OS3, 3 to 11.1 and 2 to 4. It can be concluded that sample 16.2 consists of two components denoted 16.2(a) and 16.2(b), occurring in the ratio of 4:3. The first component has a Fuc residue in (α1→2) linkage to Gal[4,3], i.e. in the upper branch, and, consequently, a terminal GlcNAc in (α1→4) linkage to Gal[4,6], in the lower branch. The structure of 16.2(b) is identical to that originally proposed [16].

The structure of fraction 17.3 has been reported to be as depicted in Fig. 1, having two (α1→2)-linked Fuc residues, and one (α1→4)-linked GlcNAc in well-defined terminal positions [16]. The 500-MHz ^{1}H-NMR spectrum of this fraction, however, reveals that in fact the sample consists of a mixture of various components, differing in the distribution of these terminal residues over the three branches. The substitution pattern of GalNAc-ol can be inferred to be as depicted in Fig. 1 from the chemical shifts of its easily recognized H-2 and H-5 signals. The branching of Gal[3] can also be characterized on the basis of its H-1 and H-4 chemical shifts (see Table 1), despite the fact that the H-1 of Gal[3] resonates in three or four signals, instead of a single one. With respect to the β-anomeric doublets, it should be noted that the H-1 signals of the Gal and GlcNAc residues of fucosylated branches dominate this spectral region (see Table 1); the overall ratio Fuc:α-GlcNAc, being 2:1, is reflected in every branch. Regarding the terminating residues, two H-1 doublets are observed for GlcNAc in (α1→4) linkage, at $\delta = 4.868$ and 4.871 ppm; also two doublets are observed for Fuc CH_3 protons, at $\delta = 1.234$ ppm and 1.231 ppm. The intensity ratio of Fuc and α-GlcNAc H-1 signals is in accord with the aforementioned overall ratio. In summary, sample 17.3 consists of various components, having in common the Gal-(β1→3)GalNAc-ol unit that bears three *N*-acetyllactosamine branches, as is indicated in Fig. 1. The components differ in the way the branches are terminated; they bear either Fuc or GlcNAc in α-linkage. The total amount of fucosylated branches in the mixture is twice as high as that of the glucosaminylated branches.

CONCLUSIONS

The 500-MHz ^{1}H-NMR spectroscopic analysis of six oligosaccharide-alditol fractions obtained from porcine blood-group H substance shows that their typical structural features can readily be elucidated by ^{1}H-NMR. In this respect, the type of substitution of GalNAc-ol, the possible branching of Gal[3], the number of *N*-acetyllactosamine branches, and the nature and location of the terminal sequences, i.e. GlcNAc-(α1→4)Gal and Fuc(α1→2)Gal, in different branches can be deduced.

This study demonstrates that high-resolution ^{1}H-NMR spectroscopy is suitable for detecting microheterogeneity in mucin-type, *O*-glycosidic carbohydrate structures. The extent to which the structures of individual constituents can be unraveled depends on the complexity of the mixture. For example, for fraction 16.2, microheterogeneity could be defined in terms of primary structures. However, fraction 17.3 is so complex that all of the constituent oligosaccharide-alditols could not be recognized. It is obvious that, for a true understanding of the physiological meaning of microheterogeneity, the precise assessment of this phenomenon in terms of primary structures is obligatory.

Thanks are due to Dr W. E. Hull (Bruker Analytische Messtechnik GmbH, Rheinstetten, FRG) for recording the 500-MHz ^{1}H-NMR spectra, and to Prof. Dr M. F. Kramer and Dr H. L. Smits (Laboratory of Histology and Cell Biology, Utrecht, The Netherlands) and Drs D. H. van den Eijnden and W. M. Blanken (Department of Medical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands) for providing us with reference compounds 4 and 6–8, respectively. This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), and by the Netherlands Foundation for Cancer Research (KWF, grant UUKC-OC 79-13).

REFERENCES

1. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Schmid, K., Montreuil, J., Fournet, B. & Hull, W. E. (1980) *FEBS Lett. 114*, 11–16.
2. Van Halbeek, H., Dorland, L., Veldink, G. A., Vliegenthart, J. F. G., Michalski, J.-C., Montreuil, J., Strecker, G. & Hull, W. E. (1980) *FEBS Lett. 121*, 65–70.
3. Van Halbeek, H., Dorland, L., Veldink, G. A., Vliegenthart, J. F. G., Strecker, G., Michalski, J.-C., Montreuil, J. & Hull, W. E. (1980) *FEBS Lett. 121*, 71–77.
4. Dorland, L., Van Halbeek, H., Vliegenthart, J. F. G., Lis, H. & Sharon, N. (1981) *J. Biol. Chem. 256*, 7708–7711.
5. Michalski, J.-C., Strecker, G., Van Halbeek, H., Dorland, L. & Vliegenthart, J. F. G. (1982) *Carbohyd. Res. 100*, 351–363.
6. Vliegenthart, J. F. G., Dorland, L. & Van Halbeek, H. (1982) *Adv. Carb. Chem. Biochem. 41*, in the press.
7. Vliegenthart, J. F. G., Van Halbeek, H. & Dorland, L. (1981) *Pure Appl. Chem. 53*, 45–77.
8. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Fiat, A.-M. & Jollès, P. (1981) *FEBS Lett. 133*, 45–50.
9. 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.
10. Van Halbeek, H., Dorland, L., Haverkamp, J., Veldink, G. A., Vliegenthart, J. F. G., Fournet, B., Ricart, G., Montreuil, J., Gathmann, W. D. & Aminoff, D. (1981) *Eur. J. Biochem. 118*, 487–495.

11. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Montreuil, J., Fournet, B. & Schmid, K. (1981) *J. Biol. Chem. 256*, 5588–5590.
12. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Jouanneau, J. & Bourrillon, R. (1981) *Biochem. Biophys. Res. Commun. 99*, 886–892.
13. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Spik, G., Chéron, A. & Montreuil, J. (1981) *Biochim. Biophys. Acta, 675*, 293–296.
14. Derevitskaya, V. A., Arbatsky, N. P. & Kochetkov, N. K. (1975) *Izv. Akad. Nauk. S.S.R. Ser. Khim.* 1411–1415.
15. Kochetkov, N. K., Derevitskaya, V. A. & Arbatsky, N. P. (1976) *Eur. J. Biochem. 67*, 129–136.
16. Derevitskaya, V. A., Arbatsky, N. P. & Kochetkov, N. K. (1978) *Eur. J. Biochem. 86*, 423–437.
17. Derevitskaya, V. A. (1981) *Pure Appl. Chem. 53*, 89–106.
18. Shashkov, A. S., Arbatsky, N. P., Derevitskaya, V. A. & Kochetkov, N. K. (1979) *Carbohyd. Res. 72*, 218–221.
19. Vliegenthart, J. F. G., Van Halbeek, H. & Dorland, L. (1980) in *IUPAC 27th Int. Congr. Pure Appl. Chem.* (Varmavuori, A., ed.) pp. 253–262, Pergamon Press, Oxford.

H. van Halbeek, L. Dorland, and J. F. G. Vliegenthart,
Organisch-Chemisch Laboratorium, Afdeling Bio-Organische Chemie, Rijksuniversiteit (te) Utrecht,
Croesestraat 79, NL-3522 AD Utrecht, The Netherlands

N. K. Kochetkov, N. P. Arbatsky, and V. A. Derevitskaya, Institut Organicheskoj Khimii imeni N.D. Zelinskogo, Akademiya Nauk S.S.S.R.,
Leninsky prospect 47, Moskva B-334, U.S.S.R. 117913