THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 262, No. 14, Issue of May 15, pp. 6650-6657, 1987
Printed in U.S.A.

Structure of the Monosialyl Oligosaccharides Derived from Salivary Gland Mucin Glycoproteins of the Chinese Swiftlet (Genus *Collocalia*)

CHARACTERIZATION OF NOVEL TYPES OF EXTENDED CORE STRUCTURE, Galβ(1→3)[GlcNAcβ(1→6)] GalNAcα(1→3)GalNAc(-ol), AND OF CHAIN TERMINATION, $[Gal\alpha(1\rightarrow4)]_{0-1}[Gal\beta(1\rightarrow4)]_2GlcNAc\beta(1\rightarrow.)$*

(Received for publication, April 1, 1986)

Jean-Michel Wieruszeski, Jean-Claude Michalski, Jean Montreuil, and Gérard Strecker

From the Laboratoire de Chimie Biologique et Unite Associée au Centre National de la Recherche Scientifique 217, Université des Sciences et Techniques de Lille I, F-59655 Villeneuve d'Ascq Cédex, France

Jasnà Peter-Katalinic and Heinz Egge

From the Institut für Physiologische Chemie, Universität Bonn, Nussallee 11, D-5300 Bonn 1, Federal Republic of Germany

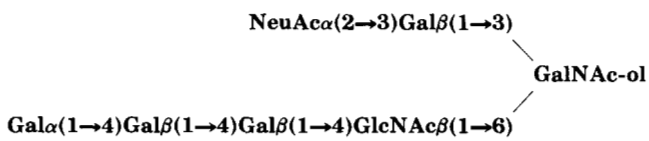
Herman van Halbeek, Johanna H. G. M. Mutsaers, and Johannes F. G. Vliegenthart

From the Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, NL-3522 AD Utrecht, The Netherlands

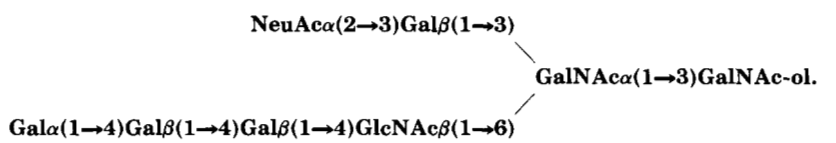
The mucus glycoproteins, the so-called nest-cementing substance, from the salivary gland of Chinese swiftlets (genus *Collocalia*) are mainly constituted of sialic acid-rich *O*-glycosylproteins. Alkaline reductive treatment of the crude material led to the release of some neutral and numerous monosialyl and disialyl oligosaccharides. These were fractionated by gel filtration, anion-exchange chromatography, and high-performance liquid chromatography.

The structures of the monosialyl oligosaccharides were established by combination of sugar and methylation analysis, fast atom bombardment-mass spectrometry, and electron impact-mass spectrometry after permethylation and ^1H NMR spectroscopy (at 500 MHz). Typically, some of the monosialyl oligosaccharides appeared to possess the core structure Galβ(1→3)[GlcNAcβ(1→6)]GalNAcα(1→3)GalNAc-ol. Moreover, the (1→6)-linked branch consisted of an unusual di- or trigalactosyl sequence, $[Gal\alpha(1\rightarrow4)_{0-1}Gal\beta(1\rightarrow4)Gal\beta(1\rightarrow4)GlcNAc\beta(1\rightarrow6)$.

Thus, the most complex representatives of the monosialyl fraction from *Collocalia* mucin were found to be:



and



The other compounds identified are partial structures thereof.

The nest-cementing substance of Chinese swiftlets (genus *Collocalia*) is a mixture of algae and mucins produced by the salivary glands of the swallows. It contains both *N*- and *O*-glycosylproteins and represents a natural source of a carbohydrate-rich material (1). Some of the neutral, *O*-linked oligosaccharide chains have been characterized (2). We have studied the sialyl oligosaccharide-alditols released from the crude edible bird's nest material by reductive β-elimination and subseqently separated by high-performance liquid chromatography (HPLC).[1] Here we report on the primary structure of five novel oligosaccharide-alditols, determined by fast atom bombardment (FAB)- and electron impact (EI)-mass

* This investigation was supported by the Centre National de la Recherche Scientifique, by the Deutsche Forschungsgemeinschaff, by the Netherlands Foundation for Chemical Research (SON/Zuiver Wetenschappelijk Onderzoek), and by Grant UUKC 83-13 from the Netherlands Cancer Foundation (KWF). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations used are: HPLC (hplc in Miniprint), high-performance liquid chromatography; glc, gas-liquid chromatography; EI-MS, electron impact-mass spectrometry; FABS-MS, fast atom bombardment-mass spectrometry; NOE, nuclear Overhauser enhancement; DSS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate.

spectrometry (MS), methylation analysis, and 500-MHz ^{1}H NMR spectrometry. A preliminary account of this investigation was presented previously.[2]

EXPERIMENTAL PROCEDURES[3]

RESULTS

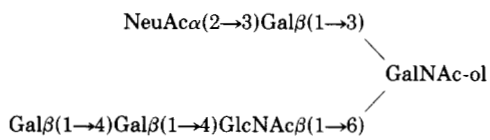
Fractionation of Reduced Oligosaccharides—Oligosaccharides released by alkaline reductive treatment were fractionated as described under "Experimental Procedures" (see Miniprint). Four fractions (A–D) were obtained after Bio-Gel P-4 chromatography (Fig. 1). The main fraction C was subdivided by ion-exchange chromatography into neutral, monosialyl, and disialyl oligosaccharides. The mixture of apparently monosialyl compounds (fraction C-20) was subfractionated by semipreparative HPLC (Fig. 2). Five monosialyl oligosaccharides (denoted as M5 to M9) were obtained in pure state. Components denoted as M1 to M4 were not carbohydrate material as shown by further investigations.

Structure Determination of the Monosialyl Oligosaccharide-alditols—The molar carbohydrate composition of the HPLC-separated monosialyl oligosaccharide-alditols M5 to M9 is given in Table I. The results of the methylation analyses of M5 to M9 are summarized in Table II. A typical example of a gas chromatogram representing the mixture of partially methylated methyl glycosides derived from such a monosialyl oligosaccharide-alditol is shown for M5 in Fig. 3. It is interesting to note that the methanolysis of methylated oligosaccharide-alditols leads to the formation of 1,4,5-Me$_3$-3,6-anhydro-GalNAc(Me)-ol (Fig. 4). In order to reveal their molecular masses, the permethylated compounds M5 to M9 were subjected to FAB-MS in the positive-ion mode. In addition, EI-MS was applied for obtained sequence information. The high-mass ranges of the FAB spectra, showing the pseudomolecular ions, and the EI spectra, including structures and *m/z* values of the predominant EI fragment ions, are presented in Fig. 5. The native samples M5 to M9 were investigated by 500-MHz ^{1}H NMR spectroscopy in aqueous solution. Their relevant NMR parameters are compiled in Table III.

Compound M5—The FAB-mass spectrum of permethylated M5 (Fig. 5*a*) revealed a pseudomolecular ion of *m/z* 1548 (M + Na$^+$), giving rise to M = 1525. This value in combination with the carbohydrate composition (Table I) proves M5 to be a hexasaccharide consisting of Gal, GlcNAc, NeuAc, and GalNAc-ol in the ratio of 3:1:1:1.

The results of methylation analysis (Table II) and the complete absence of a fragment at *m/z* 276 in the EI-MS normally representing a terminal monosubstituted hexosaminitol (Fig. 5*a*) point to a 3,6-disubstitution of GalNAc-ol. The fragment *m/z* 580 of low density in the EI-MS is correlated to the NeuAc-Gal sequence. A preferred cleavage occurs between C-4 and C-5 of GalNAc-ol giving rise to *m/z* 783. Elimination of NeuAc-Gal from the molecular ion produces an ion at *m/z* 929 which is accompanied by *m/z* 975 and *m/z* 897 (elimination of CH_3OH). The ion pairs *m/z* 219, 187 and *m/z* 668, 636 in the EI spectrum reveal the composition of the second branch to be Gal-Gal-GlcNAc; the sequence of the latter is established from the presence of fragments *m/z* 1086 (M$^+$ minus Gal-Gal), *m/z* 1290 (M$^+$ minus Gal) and the absence of an ion pair at *m/z* 464, 432 (terminal Gal-GlcNAc, compare M6, see below). The intense ion *m/z* 182 is a further indication for the 1–4 linkage at the GlcNAc residue.

In the 500-MHz ^{1}H NMR spectrum of M5 (Fig. 6*a*), the GalNAc-ol H-2 and H-5 resonances are observed at δ 4.385 and 4.267, respectively; this points to the Galβ(1→3)[GlcNAcβ(1→6)]GalNAc-ol core type (3–6). The resonance position[4] of Gal^3 H-1 (δ 4.530, $J_{1,2}$ = 7.9 Hz), in combination with that of its H-3 signal, and the set of chemical shifts for H-3ax[5] and H-3eq of NeuAc (Table III) indicate that the Gal^3 residue is substituted with NeuAc in an α(2→3) linkage (7–9). The H-1 and H-6 signals of GlcNAc[6], being found at δ 4.555 ($J_{1,2}$ = 8.3 Hz) and 3.996, respectively, are indicative of the occurrence of this GlcNAc in an *N*-acetyllactosamine unit (Galβ(1→4)GlcNAcβ(1→6) (6). The third Gal residue present in M5 (Table I) is β(1→4)-linked to Gal^{4,6}. The β-configuration of the linkage is evident from the value of $J_{1,2}$ (7.8 Hz) of the anomeric signal at δ 4.595. In a NOE experiment, presaturation of H-1 at δ 4.595 gave rise to a clearly observable interglycosidic NOE effect on H-4 of Gal^{4,6} at δ 4.188 (compare Fig. 8*c*), suggesting a 1→4 linkage. This was proven by the results of the methylation analysis; 2 substituted Gal residues are present. A 3-substituted Gal is present in the NeuAcα(2→3)Gal sequence, so the 4-substituted residue must bear the terminal Gal. Therefore, the structure of M5 was established as follows.



Apparently, attachment of Gal in a β(1→4) linkage to another Gal residue affords considerable downfield shift effects for H-1 and H-4 of the latter residue, as can be inferred from comparison of the data for M5 with those for reference compound R (7) (Table III). The rather downfield positions of these reporter-group signals make this structural element recognizable; they are sufficiently different from the chemical shifts that are characteristic of the Galβ(1→3)Galβ(1→4) element (10), to render possible their unambiguous identification.

Compound M7—Combining the results of FAB-MS (M + Na$^+$ = 1752) (Fig. 5*b*) with the data of sugar analysis of M7 (Table I), the compound appeared to be a heptasaccharide containing Gal, GlcNAc, NeuAc, and GalNAc-ol in the ratio of 4:1:1:1. Comparison of the MS spectra (Fig. 5*b*) and ^{1}H NMR spectrum (Fig. 6*b*) of M7 with those of M5 (Figs. 5*a* and 6*a* and Table III) shows that M7 can be considered as an extension of M5 with another Gal residue. The ion pair at *m/z* 872, 840 indicates that this additional Gal residue is attached to the Gal-Gal-GlcNAc moiety present in M5. Other sequence ions like *m/z* 1133, 1101 and *m/z* 1179 are accordingly shifted to values 204 mass units higher. The ions representing the NeuAc-Gal-GalNAc-ol moiety, *m/z* 376, 344, 580, and 783, remain present, unaltered as compared to M5.

Comparison of the NMR spectra and structures of M5 and M7 indicates that the signal at δ 4.366 can be associated with the presence of the Galα(1→4)Galβ(1→4) sequence. The shape of the signal at δ 4.366 is unmistakenly the shape of a

[2] Wieruszeski, J.-M., Michalski, J.-C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., van Halbeek, H., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1984) *Abstracts of the XIIth International Carbohydrate Symposium*, Utrecht, The Netherlands, p. 435, Vonk, Zeist.

[3] Portions of this paper (including "Experimental Procedures," Tables I and II, Figs. 1–8, and Refs. S1–S7) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 86 M-1050, cite the authors, and include a check or money order for $9.20 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

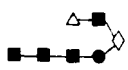
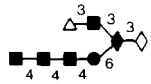
[4] For explanation of superscript notation, see footnote to Table III.

[5] H-3ax, H of carbon 3 in axial position; H-3eq, H of carbon 3 in equatorial position.

TABLE III

^1H chemical shifts of structural reporter-group protons of the constituent monosaccharides for the monosialyl oligosaccharide-alditols M5 to M9 derived from Collocalia mucin and for a related compound (R) derived from cow colostrum κ-casein (7)

n.d., not determined.

Residue[c]	Reporter group	Chemical shift[a] in[b] R (7)	M5	M7	M6	M8	M9
							
GalNAc-ol	H-2	4.389	4.385	4.387	4.380	4.376	4.377
	H-3	4.069	4.065	4.067	3.886	3.886	3.887
	H-4	3.435	3.435	3.435	n.d.	3.6	n.d.
	H-5	4.271	4.267	4.268	n.d.	n.d.	n.d.
	H-6	3.927	3.894	3.9	n.d.	n.d.	n.d.
	NAc	2.065	2.064	2.066	2.043	2.043	2.045
GalNAc^3	H-1	-	-	-	5.109	5.110	5.111
	H-2	-	-	-	4.396	4.395	4.396
	H-3	-	-	-	4.043	4.039	4.041
	H-4	-	-	-	4.262	4.262	4.262
	H-5	-	-	-	4.155	4.153	4.153
	H-6	-	-	-	4.081	4.082	4.082
	NAc	-	-	-	2.075	2.075	2.076
Gal^3	H-1	4.532	4.530	4.530	4.540	4.538	4.539
	H-3	4.114	4.113	4.113	4.074	4.074	4.070
	H-4	3.922	3.928	3.930	3.931	3.931	3.931
GlcNAc^6	H-1	4.558	4.555	4.558	4.607	4.605	4.607
	H-6	3.993	3.996	3.997	3.998	3.997	4.003
	NAc[d]	2.065	2.064	2.066	2.037	2.037	2.038
Gal^{4,6}	H-1	4.467	4.498	4.503	4.467	4.499	4.503
	H-4	3.931	4.188	4.204	3.923	4.189	4.204
Gal^{4,4,6}	H-1	-	4.595	4.709	-	4.597	4.710
	H-4	-	3.903	4.028	-	3.902	4.027
Gal^{4,4,4,6}	H-1	-	-	4.953	-	-	4.953
	H-4	-	-	4.035	-	-	4.034
	H-5	-	-	4.366	-	-	4.365
NeuAc^3	H-3ax	1.801	1.799	1.800	1.783	1.783	1.784
	H-3eq	2.773	2.770	2.773	2.760	2.759	2.759
	NAc	2.032	2.033	2.033	2.028	2.028	2.029

[a] Chemical shifts are given for D_2O solutions at pD 7 and at 27 °C, in parts/million downfield from DSS (using internal acetone at δ 2.225).

[b] For complete structures of the compounds, see text. In the table heading, the structures are represented by short-hand symbolic notation as follows: ◇–, GalNAc-ol; ◆–, GalNAc; ●–, GlcNAc; ■–, Gal, and Δ–, NeuAc.

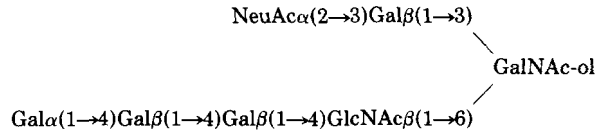
[c] A superscript at the name of a sugar residue indicates to which position of the adjacent monosaccharide it is glycosidically linked. More than one superscript is used to discriminate between identically bound Gal residues; they indicate the types of linkages of the neighboring residues in the sequence.

[d] The *N*-acetyl signal of GlcNAc^6 at δ 2.037 is assigned on the basis of comparison with another GlcNAc residue β(1→6)-linked to an α-linked sugar, namely the GlcNAc$^{7'}$ in a tetraantennary structure. This NAc signal is found at δ 2.040 (11). NMR parameters of GalNAc α(1→3)GalNAc-ol are described by Hounsell *et al.* (12). The two-*N*-acetyl signals of this oligosaccharide present at δ 2.049 and 2.060 are assigned to GalNAc-ol and GalNAc, respectively. On the basis of analogy with this structure, we assigned the signal at δ 2.043 to GalNAc-ol and the signal at 2.075 to GalNAc^3. It should be noted that these assignments are tentative.

H-5 signal.

Examination of the spectra of structures possessing α-linked Gal or GalNAc residues shows that the H-5 of such residues is found in the structural reporter-group region downfield from the bulk of skeleton protons (13, 15, 16, 21). Therefore, we assigned the signal at δ 4.366 to H-5 of Gal.

The structure of M7 is proposed as follows.



Owing to extension with the $\alpha(1\rightarrow4)$-linked Gal residue, the H-1 and H-4 of $Gal^{4,4,6}$ undergo considerably large downfield shift effects (compare M5 and M7) in Table III: for H-1, $\Delta\delta$ = 0.114 ppm, and for H-4, $\Delta\delta$ = 0.125 ppm). The parameters typical of the Galα(1→4)Galα(1→.) sequence are sufficiently different from those of Galα(1→3)Galα(1→.) (13, 17) to permit their identification on the basis of chemical shifts of structural-reporter groups only.

Compound M8—The molecular weight of permethylated M8, as determined by FAB-MS (M + Na^+ = 1793) is 1770 (Fig. 5*d*). This result, together with the sugar composition (Table I), allowed us to predict that M8 is a heptasaccharide consisting of Gal, GlcNAc, NeuAc, GalNAc, and GalNAc-ol in the molar ratio of 3:1:1:1:1. The occurrence of a 3,6-disubstituted GalNAc in M8 is indicated by the finding of 4-mono-MeGalNAc(Me) in the methylation analysis (Table II), whereas the presence of 1,4,5,6-Me_4-GalNAc(Me)-ol points to a mono-3-substituted, terminal GalNAc-ol. These features are reflected in the EI-MS by a series of fragmentations. The terminal GalNAc-ol is represented by *m/z* 276, the NeuAc by *m/z* 376, 344, and the Gal-Gal-GlcNAc unit by *m/z* 219, 187 and *m/z* 668, 636, 182. The hexasaccharide ion up to and including the core GalNAc is represented by *m/z* 1478 both in the EI- and FAB-MS. This type of fragmentation is similar to that observed for the permethylated reduced chitobiosyl residue of *N*-linked oligosaccharides (18).

As in M5 and M7, the NeuAc-Gal residue linked to position 3 of the GalNAc is eliminated under electron impact with the positive charge remaining with the GalNAc residue leading to *m/z* 1174. In addition a fragment 46 mass units higher at *m/z* 1220 is present that corresponds to the Jo fragment ion first described by Sharp and Albersheim (19) in the EI spectra of per-*O*-alkylated (1→3)-linked oligosaccharide-alditols.

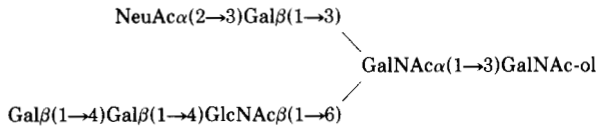
Analogous Jo fragment ions are observed in M5 at *m/z* 975 (Fig. 5*a*), in M7 at *m/z* 1179 (Fig. 5*b*), and in M9 at *m/z* 1424 (Fig. 5*c*). Under these conditions of fragmentation, hardly any fragment ion *m/z* 580 corresponding to the sequence NeuAc-Gal is formed.

The ^1H NMR spectrum of M8 is depicted in Figs. 7*c* and 8*a*. Comparison of the NMR data of M8 with those of M5 (Table III) shows the common presence of the sequences Galβ(1→4)Galβ(1→4)GlcNAcβ(1→.) and NeuAcα(2→3)Galβ(1→.) in both compounds. Therefore, the H-1 resonance at δ 5.110 ($J_{1,2}$ = 4.1 Hz) has to be attributed to the GalNAc residue present in M8. The $J_{1,2}$ value indicates that this GalNAc is α-linked. Various other structural reporter-group signals can be ascribed to this GalNAc residue, namely the signal at δ 4.395 to its H-2, at δ 4.039 to H-3, at δ 4.262 to H-4, at δ 4.153 to H-5, and at δ 4.082 to its H-6; the assignments can be achieved unambiguously by consecutive double-resonance (spin decoupling and NOE) experiments starting from the H-1 doublet at δ 5.110.

The core GalNAc serves as a C-3, C-6 branching point: the disaccharide unit NeuAcα(2→3)Gal(1→.) is attached to C-3, while the sequence Galβ(1→4)Galβ(1→4)GlcNAcβ(1→.) is linked to C-6. The sites of substitution have been established by the observation of interresidue NOE couplings over the glycosidic bond between H-1 of Gal^3 and H-3 of GalNAc, on one hand (Fig. 8*b*), and H-1 of $GlcNAc^6$ and H-6 of GalNAc on the other hand (Fig. 8*c*) (compare Refs. 13, 20, and 21). The H-1 of $GlcNAc^6$ (δ 4.605) could not be irradiated selectively because its resonance nearly coincides with that of H-1 of $Gal^{4,4,6}$ (δ 4.597). Therefore, the NOE difference spectrum (Fig. 8*c*) showed also a pronounced interresidue NOE effect on H-4 of $Gal^{4,6}$ at δ 4.189. Again, as in M5, this suggests the linkage between the two adjacent Gal residues to be (1→4).

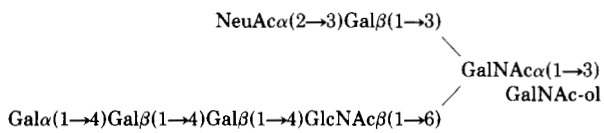
GalNAc in its turn is α-linked to C-3 of GalNAc-ol. This became evident from methylation analysis (Table II) and, independently, from ^1H NMR. As illustrated in Fig. 8*d*, an NOE effect on the H-3 signal of GalNAc-ol at δ 3.886 could be detected due to presaturation of H-1 of GalNAc, and vice versa. (In a separate experiment, the signal of H-3 of GalNAc-ol had been located at δ 3.886 by selective irradiation of the H-2 signal of GalNAc-ol at δ 4.376).

In conclusion, the oligosaccharide-alditol M8 was completely sequenced by employment of 500-MHz ^1H NMR; its structure is proposed as follows.



Compound M9—Sugar analysis of M9 (Table I) revealed its carbohydrate composition to be Gal, GlcNAc, NeuAc, GalNAc, and GalNAc-ol in the ratio of 4:1:1:1:1. In combination with the *m/z* value of the pseudomolecular ion M + Na^+ 1997 (Fig. 5*e*), M9 could be deduced to be an octasaccharide. As in M8, methylation analysis furnished a mono-3-substituted GalNAc-ol and a 3,6-disubstituted GalNAc; besides, two mono-4-substituted Gal residues were detected (Table II). The location of this additional Gal at the Gal-Gal-GlcNAc chain is proven by the fragments *m/z* 872, 840 present in the EI-MS (Fig. 5*e*). The terminal GalNAc-ol gives rise to an intense fragment *m/z* 276 and the branched heptasaccharide moiety gives an intense *m/z* 1682 in the FAB-MS (not shown). The ions produced after elimination of the NeuAc-Gal chain are *m/z* 1378 and 1424, both being 204 mass units shifted with respect to those of M8.

Comparison of the ^1H NMR data of M9 (Fig. 7*d* and Table III) with those of M7 and M8 showed that M9 can be considered as an extension of M8 with an additional Gal residue in $\alpha(1\rightarrow4)$ linkage to $Gal^{4,4,6}$. The typical parameters of this trigalactosyl moiety which were established for M7 and the effects of attachment of $\alpha(1\rightarrow4)$-linked Gal (compare the step from M5 to M7, Table III) are found again for M9 (and the step from M8 to M9). Therefore, the structure of M9 has been established as follows.

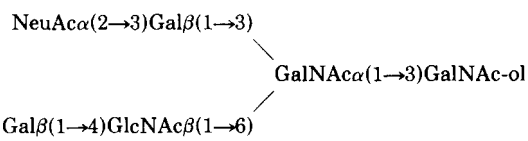


Compound M6—FAB-MS analysis revealed the presence of two compounds in HPLC fraction 6 with M = 1525 and 1566, respectively. From the pseudomolecular ion *m/z* 1548 (Fig. 5*c*-1), it can be deduced that this component is at least isomeric to M5 (M = 1525). A closer inspection of the FAB- and EI-MS shows that all key ions present in the spectra of M5 are also present in those of HPLC fraction 6. The other component (M6) was then concluded to be a hexasaccharide consisting of Gal, GlcNAc, NeuAc, GalNAc, and GalNAc-ol in the ratio of 2:1:1:1:1 (compare Table I). The EI spectrum of M6, after subtraction of that of M5 from that of HPLC fraction 6, proportional to *m/z* 668, yields the spectrum shown in Fig. 5*c*-2 which can be conceived as that of pure M6. Characteristic features are *m/z* 276 (terminal mono-3-substituted GalNAc-ol), *m/z* 464, 432 (terminal Gal(1→4)GlcNAc), and *m/z* 1274 that represents the branched pentasaccharide moiety. Several other fragment ions like *m/z* 899, 970 characterize M6 as the lower homologue of M8 and M9, missing the nonreducing Gal-Gal sequence. The ion *m/z* 899 deserves

special attention because of its high intensity. It can be explained formally by the elimination of the NeuAc residue from the fragment ion *m/z* 1274 (1274-NeuAc + H). Linked scan measurement indicates that this ion can serve as mother ion for *m/z* 867, 464, 432, and 356. This points to a stepwise degradation of the fragment ion *m/z* 1274 finally leading to the disaccharide ion *m/z* 464. The fact that analogous ions are found in M8 shifted by 204 mass units to *m/z* 1103 and 1350 and in M9 at *m/z* 1307 and 1554 indicates that these fragment ions are derived from the terminal penta-, hexa-, and heptasaccharides as pointed out in the schemes of fragmentation of Fig. 5, *c–e*.

From the 500-MHz ^{1}H NMR spectrum of this HPLC fraction 6 (Fig. 7*a*), it could be deduced that the minor component (20%) is identical with M5. Subtraction of the ^{1}H NMR spectrum of compound M5 (Fig. 6*a*) from that of fraction 6 (Fig. 7*a*) yielded the spectrum of pure M6 (Fig. 7*b*). Once the structural reporter-group signals of the core type Galβ(1→3)[GlcNAcβ(1→6)]GalNAcα(1→3)GalNAc-ol were assigned in the NMR spectrum of M8 (Fig. 7*c*), they could be recognized in the spectrum of "NMR-pure" compound M6. The Gal^3 residue is extended with NeuAc in α(2→3) linkage. GlcNAc^6 is substituted with a Gal residue in β(1→4) linkage which can be deduced from comparison of the NMR data of M6 with those for reference substance R (7) and M8 (see Table III).

Therefore, the structure of the major component M6 of HPLC fraction 6 has been revealed as follows.



This structure is in agreement with the methylation analysis data (Table II) obtained on the 1:4 mixture of M5 and M6.

DISCUSSION

Previous structural studies have shown that *Collocalia* mucin contains at least two glycoproteins (CMI and CMII), separable by ion-exchange chromatography on DEAE-cellulose (1). Recently, the structures of neutral di-, tetra-, and hexasaccharides released from a purified sialylglycoprotein fraction were established as follows: Galβ(1→3)GalNAc-ol, Galβ(1→4)GlcNAcβ(1→6)[Galβ(1→3)]GalNAc-ol, and Galβ(1→4)GlcNAcβ(1→3)Galβ(1→4)GlcNAcβ(1→3)Galβ(1→3)GalNAc-ol (2). However, the structural investigations were carried out on a purified sialylglycoprotein which corresponded to less than 10% of the mucin, while our studies have been undertaken on the crude material.

We established the structures of a novel type of *O*-glycosidically linked carbohydrates, with the typical feature of the core GalNAcα(1→3)GalNAc-ol, of which GalNAc serves as branching point. A similar type of core structure has been characterized in human rectal adenocarcinoma glycoprotein: NeuAcα(2→6)[GalNAcα(1→3)]GalNAc-ol (22). A second interesting, novel structural element is represented by the di- and trigalactosyl sequences [Galα(1→4)]$_{0\text{-}1}$ Galβ(1→4)Galβ(1→4)-linked to GlcNAc. The terminal Galα(1→4)Galβ(1→4) sequence is known to occur in blood group P$_1$ antigenic determinants (23), and particularly in another material of avian origin, turtledove ovomucoid (24).

Acknowledgments—We wish to thank Catherine Alonso and Gisèle Tinel for excellent technical assistance.

REFERENCES

1. Keilich, G., Kopp, E. C., and Brossmer, R. (1977) *Hoppe Seyler's Z. Physiol. Chem.* **358,** 257
2. Hanisch, F. G., and Uhlenbruck, G. (1984) *Hoppe-Seyler's Z. Physiol. Chem.* **365,** 119–128
3. van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Hull, W. E., Lamblin, G., Lhermitte, M., Boersma, A., and Roussel, P. (1982) *Eur. J. Biochem.* **127,** 7–20
4. Lamblin, G., Boersma, A., Lhermitte, M., Roussel, P., Mutsaers, J. H. G. M., van Halbeek, H., and Vliegenthart, J. F. G. (1984) *Eur. J. Biochem.* **143,** 227–236
5. Lamblin, G., Boersma, A., Klein, A., Roussel, P., van Halbeek, H., and Vliegenthart, J. F. G. (1984) *J. Biol. Chem.* **259,** 9051–9058
6. van Halbeek, H. (1984) *Biochem. Soc. Trans.* **12,** 601–605
7. van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Fiat, A. M., and Jollès, P. (1981) *FEBS Lett.* **133,** 45–50
8. Korrel, S. A. M., Clemetson, K. J., van Halbeek, H., Kamerling, J. P., Sixma, J. J., and Vliegenthart, J. F. G. (1984) *Eur. J. Biochem.* **140,** 571–576
9. Akiyama, K., Simons, E. R., Bernasconi, P., Schmid, K., van Halbeek, H., Vliegenthart, J. F. G., Haupt, H., and Schwick, H. G. (1984) *J. Biol. Chem.* **259,** 7151–7154
10. van Halbeek, H., Dorland, L., Veldink, G. A., Vliegenthart, J. F. G., Garegg, P. J., Norberg, T., and Lindberg, B. (1982) *Eur. J. Biochem.* **127,** 1–6
11. Vliegenthart, J. F. G., Dorland, L., and van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41,** 209–374
12. Hounsell, E. F., Lawson, A. M., Feeney, J., Gooi, H. C., Pickering, N. J., Stoll, M. S., Lui, S. C., and Feizi, T. (1985) *Eur. J. Biochem.* **148,** 367–377
13. van Halbeek, H., Vliegenthart, J. F. G., Winterwerp, H., Blanken, W. M., and Van den Eijnden, D. H. (1983) *Biochem. Biophys. Res. Commun.* **110,** 124–131
14. François-Gérard, C., Pierce-Crétel, A., André, A., Dorland, L., van Halbeek, H., Vliegenthart, J. F. G., Montreuil, J., and Spik, G. (1983) in *Proceedings of the 7th International Symposium on Glycoconjugates* (Chester, M. A., Heinegard, D., Lundblad, A., and Svensson, S., eds) pp. 169–170, Rahms i Lund, Sweden
15. Nasir-ud-Dim, Jeanloz, R. W., Lamblin, G., Roussel, P., van Halbeek, H., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1986) *J. Biol. Chem.* **261,** 1992–1997
16. Dua, V. K., Rao, B. N. N., Wu, S.-S., Dube, V. E., and Bush, C. A. (1986) *J. Biol. Chem.* **261,** 1599–1608
17. Dorland, L., van Halbeek, H., and Vliegenthart, J. F. G. (1984) *Biochem. Biophys. Res. Commun.* **122,** 859–866
18. Egge, H., Peter-Katalinic, J., Paz-Parente, J., Strecker, G., Montreuil, J., and Fournet, B. (1983) *FEBS Lett.* **156,** 357–362
19. Sharp, J. K., and Albersheim, P. (1984) *Carbohydr. Res.* **128,** 193–202
20. Prestegard, J. H., Koerner, T. A. W., Demou, P. C., and Yu, R. K. (1982) *J. Am. Chem. Soc.* **104,** 4993–4995
21. Dabrowski, J., Hanfland, P., and Egge, H. (1982) *Methods Enzymol.* **83,** 69–86
22. Kurosaka, A., Nakajima, H., Funakoshi, I., Matsuyama, M., Nagayo, T., and Yamashina, I. (1983) *J. Biol. Chem.* **258,** 11594–11598
23. Naiki, M., and Marcus, M. (1975) *Biochemistry* **14,** 4837–4841
24. François-Gérard, C., Brocteur, J., André, A., Gerday, C., Pierce-Crétel, A., Montreuil, J., and Spik, G. (1980) *Blood Transf. Immunohaematol.* **5,** 579–588

Additional references are found on p. 6655.

Supplement to : STRUCTURE OF THE MONOSIALYL OLIGOSACCHARIDES DERIVED FROM SALIVARY-GLAND MUCIN GLYCOPROTEINS OF THE CHINESE SWIFTLET (GENUS COLLOCALIA)

Jean-Michel Wieruszeski, Jean-Claude Michalski, Jean Montreuil, Gérard Strecker, Jasna Peter-Katalinic, Heinz Egge, Herman van Halbeek, Johanna H.G.M. Mutsaers and Johannes F.G. Vliegenthart.

EXPERIMENTAL PROCEDURES

Fractionation of reduced oligosaccharides - Edible bird's nest material was obtained from Yuen Wo Bird's Nest Co., Hong Kong. The material (6 g) was powdered with a Waring blendor and submitted to alkaline reductive degradation in 50 mM NaOH containing 1.0 M $NaBH_4$ (250 ml) at 37°C for 48 h. The reaction was stopped by addition of Dowex 50 x 8 (25-50 mesh ; H^+) in the presence of 0.2 ml octanol, at 4°C. The solution was filtered, adjusted to pH 5.5 with 0.1 M NaOH and then concentrated with methanol. The material was then fractionated on a Bio-Gel P-4 column (3 x 50 cm). The main fraction (fraction C) was applied on a column (2 x 15 cm) of Dowex 1 x 2 (200- 400 mesh ; acetate). After washing with 200 ml water, the sialyloligosaccharides were eluted with a discontinuous gradient (20, 50, 100 and 200 mM) of pyridine-acetic acid buffer (pH 5.6).

Analytical procedures - Reduced oligosaccharides were isolated by hplc on primary amine bonded silica (5 µm Amino AS-5A column ; 0.4 x 25 cm). A solution of CH_3CN-15 mM KH_2PO_4 buffer, pH 5.2 (75 : 25) was applied for 25 min, followed by a linear gradient from 75 : 25 to 55 : 45 for 65 min. The flow rate was 1 ml/min and the oligosaccharides were detected by UV spectroscopy at 200 nm.

The molar ratios of hexoses, N-acetylhexosamines, N-acetylhexosaminitols and N-acetylneuraminic acid were determined by glc of the trifluoroacetyl derivatives (S1), obtained after methanolysis of oligosaccharide-alditols (methanol-0.5 M HCl, 24 h, 80°C).

The methylation analysis was performed according to Finne et al. (S2). The partially methylated methylglycosides released by methanolysis were acetylated (pyridine acetic anhydride ; 1:10 ; 0.2 ml) and the products were analysed by glc-MS (S3) with a capillary column (0.33 mm x 25 m) coated with fused CP-SIL 5 CB (temperature programme, 100 → 240°C, at 5°/min).

Mass spectrometry was performed on a ZAB HF reversed geometry mass spectrometer (VG Analytical, Manchester, U.K.). For FAB-MS, the sample was bombarded with xenon atoms having a kinetic energy equivalent to 9 keV. Permethylated oligosaccharides (3-5 µg) were added in methanolic solution (1 µl) to a thioglycerol matrix (1-mercapto-2,3-propanediol, Ega Chemie, Steinheim, F.R.G.). Positive ion spectra were recorded in a mass controlled scan of 300-500 sec duration. The spectra were evaluated by counting the spectral lines. The values presented are therefore nominal masses. They are lower by about one mass unit per 2000 units as compared to the calculated exact physical masses.

EI-MS was performed on the same instruments. Samples of 5-10 µg of permethylated oligosaccharide were introduced in shallow quartz tubes close to the electron beam and heated indirectly by the heater of the ion source to 250-290°C until relevant signals were obtained. The ionization energy was 25 eV and the acceleration voltage 8kV. Linked scan measurements were performed under electron impact ionization in order to establish fragmentation pathways. For the detection of mother ions, B^2/E measurements were used and for the detection of daughter ions B/E and MIKES (mass-analysed ion kinetic energy spectrometry) (S4).

Prior to ^{1}H-NMR spectroscopic analysis, the oligosaccharide-alditol fractions were repeatedly treated with D_2O at pD 7 and room temperature. After each exchange treatment the materials were lyophilized. Finally, the samples were redissolved in 400 µl D_2O (99.96 atom % D, Aldrich, Milwaukee, WI). ^{1}H-NMR spectroscopy was performed on a Bruker WM-500 spectrometer (SON hf-NMR facility, Department of Biophysical Chemistry, Nijmegen University, The Netherlands), operating at 500 MHz in the Fourier transform mode and equipped with a Bruker Aspect 2000 computer. Further experimental details have been reported (S5). For recording NOE-difference spectra, 20 µl acetone-D6 was added to the samples ; the deuterium resonance of the acetone was used as field-frequency lock signal. NOE-difference spectra were obtained according to (S6) in combination with a DANTE pulse sequence for selective suppression of the HOD-line (S7). Resolution-enhancement of the spectra was achieved by Lorentzian-to-Gaussian transformation. The indicated probe temperature was 27°C, and was kept constant within 0.1°. Chemical shifts (δ) are expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS), but were actually measured by reference to internal acetone (δ 2.225 in D_2O at 27°C), with an accurancy of 0.002 ppm.

REFERENCES

S1. Zanetta, J.P., Breckenridge, W.C. and Vincendon, G. (1972) J. Chromatogr. 69, 291-304.
S2. Finne, J., Krusius, T. and Rauvala, H. (1980) Carbohydr. Res. 80, 336-339.
S3. Fournet, B., Strecker, G., Leroy, Y. and Montreuil, J. (1981) Anal. Biochem. 116, 489-502.
S4. Bogzorgzadeh, M.H., Morgan, R.P. and Beynon, J.H. (1978) Analyst 103, 613-622.
S5. Vliegenthart, J.F.G., Dorland, L. and van Halbeek, H. (1983) Adv.Carbohydr. Chem. Biochem. 41, 209-374.
S6. Wagner, G. and Wuthrich, K. (1979) J. Magn. Reson. 33, 675-680.
S7. Haasnoot, C.A.G. (1983) J. Magn. Reson. 52, 153-158.

Table I

Molar carbohydrate composition of the monosialyl oligosaccharide-alditols M5 to M9 derived from Collocalia mucin.

Sample	Molar ratio				
	GalNAc-ol[a]	Gal	GalNAc	GlcNAc	NeuAc
M5	1.0	3.0	-	1.0	0.8
hplc fraction 6[b]	1.0	2.5	0.8	0.9	0.8
M7	1.0	4.2	-	1.1	0.8
M8	1.0	3.3	0.9	1.1	0.9
M9	1.0	4.2	0.9	1.1	0.9

[a] including anhydro derivatives
[b] consisting of M6 and M5 in the ratio of 4:1

Table II

Molar ratios of partially methylated monosaccharides present in the methanolysates of the permethylated oligosaccharide-alditols M5 to M9 derived from Collocalia mucin.

Methyl ether	Molar ratio[a] in oligosaccharide-alditol				
	M5	fraction 6[c]	M7	M8	M9
2,3,4,6-Me4-Gal[b]	0.9	0.8	0.8	0.9	0.8
2,3,6-Me3-Gal	1.1	0.3	2.1	1.3	2.1
2,4,6-Me3-Gal	1.1	1.1	1.2	1.1	1.1
3,6-Me2-GlcNAc(Me)	1.0	1.0	1.0	1.0	1.0
1,4,5,6-Me4-GalNAc(Me)-ol	-	0.8	-	0.7	0.9
1,4,5-Me3-GalNAc(Me)-ol					
1,4,5-Me3-3,6-anhydro-GalNAc(Me)-ol	0.9	0.2	1.1	-	-
4-Me-GalNAc(Me)	-	0.7	-	0.7	0.8
4,7,8,9-Me4-NeuAc(Me)	1.1	1.0	1.1	0.9	1.1

[a] calculated relative to 3,6-Me_2-GlcNAc = 1.0
[b] 2,3,4,6-Me_4-Gal, 2,3,4,6-tetra-O-methyl methylgalactoside, etc.
[c] consisting of M6 and M5 in the ratio of 4:1

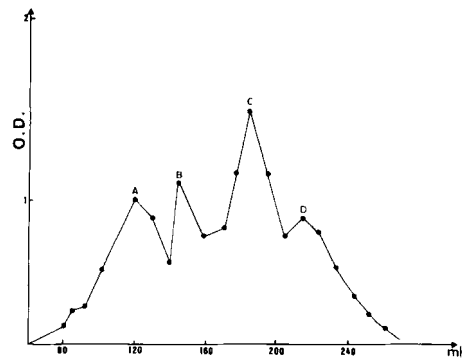


Fig. 1. Bio-Gel P-4 chromatography of oligosaccharide-alditols released from Collocalia mucin by alkaline reductive treatment. Fractions A-D were pooled as indicated by the bars ; the yields were as follows : 1, 572 mg ; B, 250 mg ; C, 914 mg ; D, 380 mg. OD optical density.

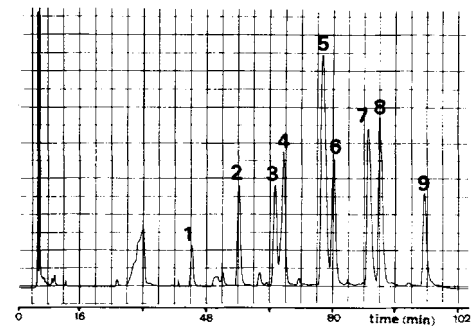


Fig. 2. Semi-preparative hplc of the monosialyl oligosaccharide-alditols (fraction C2D) on 5 µm Amino AS-5A. For chromatographic conditions, see Experimental Procedures.

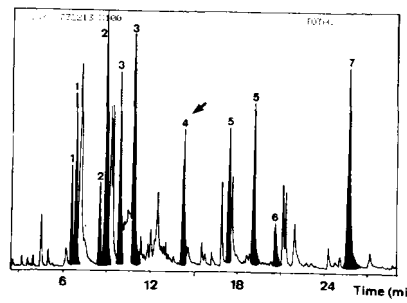


Fig. 3. Gas chromatogram of the partially methylated methylglycosides derived from M5. (Note the occurrence of peak4 ; its EI-MS pattern is given in Fig. 4). 1 : 2,3,4,6-Me_4-Gal ; 2 : 2,3,6-Me_3-Gal ; 3 : 2,4,6-Me_3-Gal ; 4 : 1,4,5-Me_3-3,6-anhydro-GalNAc(Me)-ol ; 5 : 3,6-Me_2GlcNAc(Me) ; 6 : 1,4,5-Me_3-GalNAc(Me)-ol ; 7 : 4,7,8,9-Me_4-NeuAc(Me).

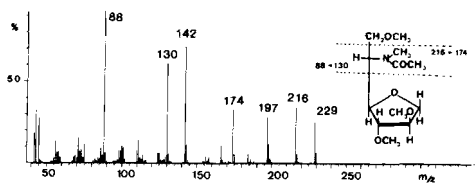


Fig. 4. EI mass spectrum of 1,4,5-Me_3-3,6-anhydro-GalNAc(Me)-ol.

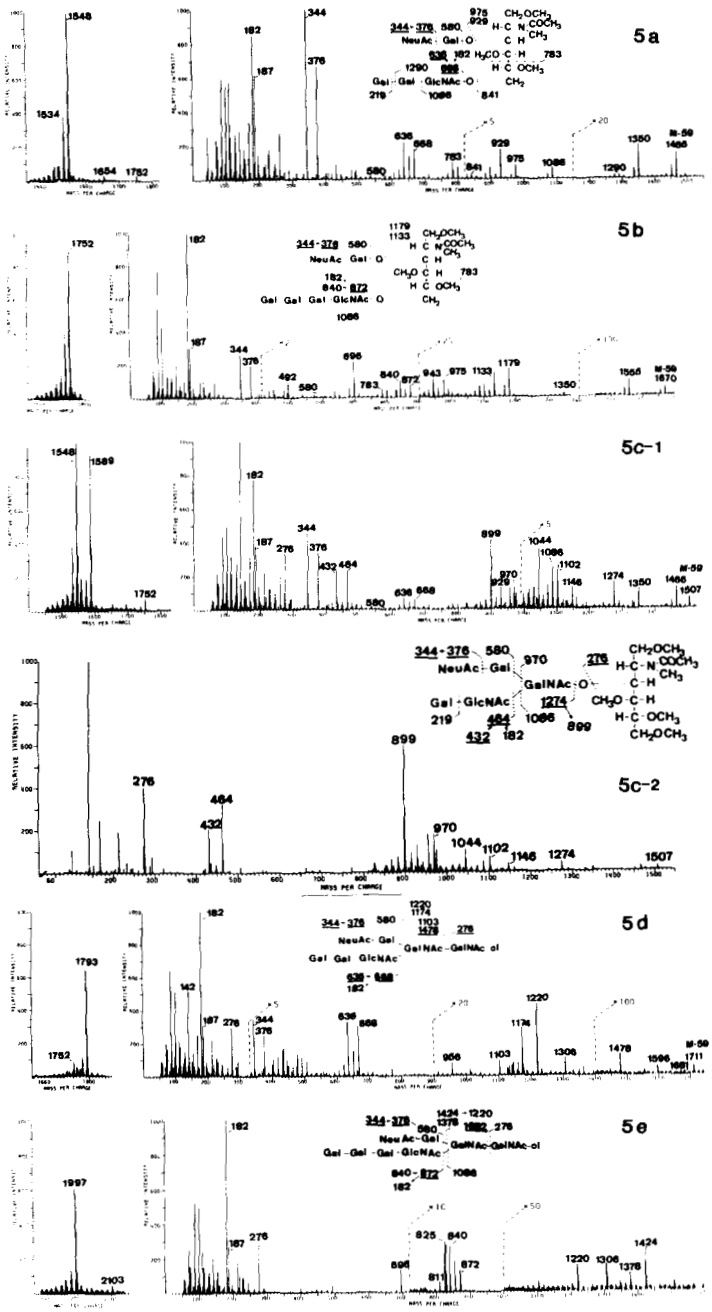


Fig. 5. FAB- AND EI mass spectra and fragmentation schemes of permethylated oligosaccharide-alditols. (a) M5 ; (b) M7 ; (c_1) hplc fraction 6 ; (c_2) M6 ; (d) M8 ; (e) M9.

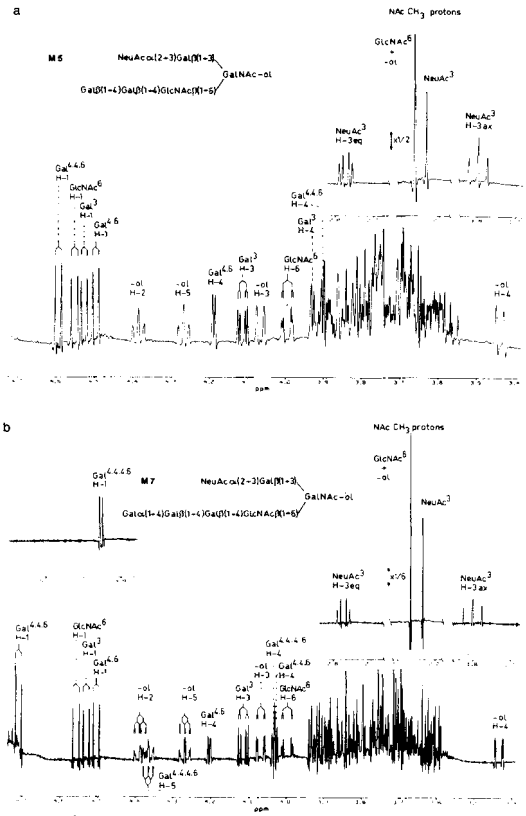


Fig. 6. 500-MHz ^{1}H-NMR spectra (D_2O ; pD 7 ; 27°C) of monosialyl oligosaccharide-alditols M5 (a) and M7 (b). The relative-intensity scale of the N- acetyl proton regions deviates from that of the other parts of the respective spectra, as indicated.

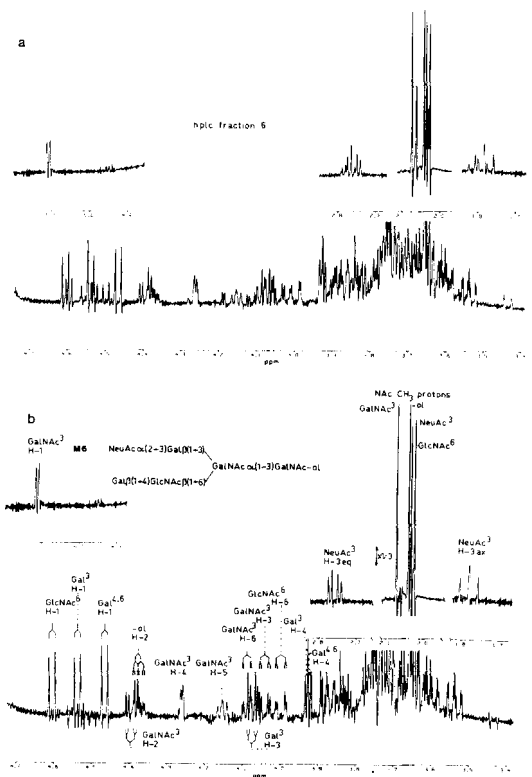


Fig. 7. 500-MHz ^{1}H-NMR spectra (D_2O ; pD 7 ; 27°C) of (a) the mixture of monosialyl oligosacchariche-alditols M5 and M6 (hplc fraction 6) ; (b) compound M6 (resulting from substraction of spectrum 6a from 7a) ; (c) compound M8 and (d) compound M9. The relative-intensity scale of the N-acetyl proton regions deviates from that of the other parts of the respective spectra, as indicated.

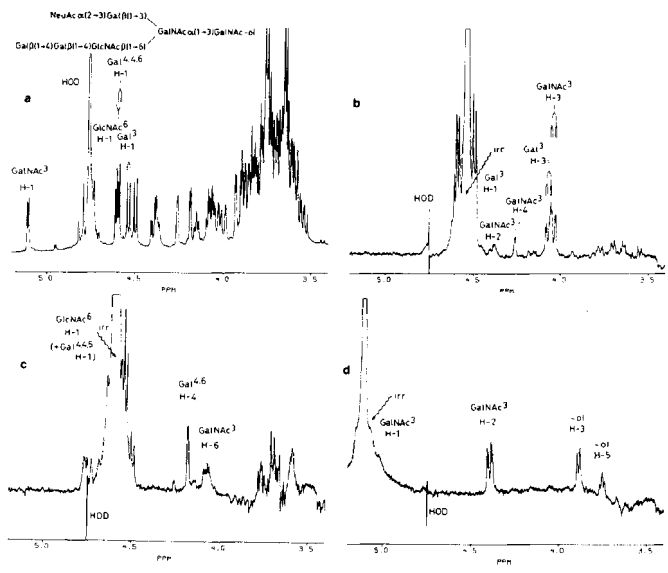


Fig. 8. (a) Relevant part of the 500-MHz ^{1}H-NMR spectrum of monosialyl oligosaccharide -alditol M8.
(b) NOE-difference spectrum, with on-resonance irradiation of H-1 of Gal^3.
(c) NOE-difference spectrum with simultaneous on-resonance irradiation of H-1 of $GlcNAc^6$ and $Gal^{4,4,6}$.
(d) NOE-difference spectrum, with on-resonance irradiation of H-1 of $GalNAc^3$.