Structural characterization of neutral oligosaccharides with blood-group A and H activity isolated from bovine submaxillary mucin

Angela V. SAVAGE,* Sinead M. T. D'ARCY and Carmel M. DONOGHUE

Department of Chemistry, University College, Galway, Ireland

In this study we investigated the structures of 11 neutral oligosaccharides released from bovine submaxillary mucin by alkaline borohydride treatment and isolated by h.p.l.c. One hexa-, one penta-, three tetra-, four tri- and two di-saccharides containing core types 1, 2, 3 or 4 were obtained. We report their structures, determined by a combination of one- and two-dimensional ^1H n.m.r. spectroscopy at 270 MHz and methylation analysis involving g.l.c.–m.s., along with their approximate molar ratios. Only three of these oligosaccharides have previously been reported in this source. Of the new oligosaccharides, one contains the blood-group-A antigenic determinant, two contain the blood-group-H type 2 determinant, while another contains the blood-group-H type 3 determinant. The oligosaccharide $\text{GlcNAc}\beta(1\rightarrow6)[\text{GlcNAc}\beta(1\rightarrow3)]\text{GalNAc}$ ol, although previously found as a core structure, has been isolated here as a novel trisaccharide.

INTRODUCTION

Previous investigations of the oligosaccharides released by alkaline borohydride treatment of bovine submaxillary mucin have revealed the presence of two neutral disaccharides, a trisaccharide and a tetrasaccharide [1,2] along with a series of acidic oligosaccharides [3–6]. In this study we have extended this work by investigating a number of neutral oligosaccharides present in minor amounts. The structures were determined by analysis of the one-dimensional ^1H n.m.r. spectra at 270 MHz in conjunction with two-dimensional correlation experiments (COSY) and methylation analysis. A preliminary account of these structures has appeared [7].

This study should contribute to a better understanding of the function of the carbohydrate chains on salivary mucins. Furthermore, the structures found reflect the glycosylation potential of the bovine submaxillary gland and give indications of the specificity of the glycosyltransferases involved in their synthesis.

MATERIALS AND METHODS

Isolation of the mucin, β -elimination and purification of the oligosaccharides and methods for carbohydrate composition, methylation analysis and n.m.r. experiments were essentially as previously described [3,5].

RESULTS

Isolation of oligosaccharides

Approx. 90% of the total mucin-bound sialic acid was precipitated in the 65–75% (v/v) ethanol fraction. From 4 g of this mucin 2.5 g of oligosaccharides was released by β -elimination. The neutral fraction from ion-exchange chromatography (approx. 1 g), when applied to a Bio-Gel P-4 column, gave four peaks (Fig. 1), N1, N2, N3 and N4, which were further fractionated by h.p.l.c. Peak N1 gave six fractions (N1a–N1f, Fig. 2), peak N2 gave seven fractions (N2a–N2g, Fig. 3), peak N3 gave four fractions (N3a–N3d, Fig. 4), while peak N4 was homogeneous on h.p.l.c. These fractions were isolated by semipreparative h.p.l.c. and desalted using Bio-Gel P-2

chromatography, before n.m.r. investigations and methylation analysis.

Composition of oligosaccharides

Using g.l.c. the oligosaccharides were shown to have the monosaccharide composition shown in Table 1.

Fig. 1. Gel-filtration fractionation of the neutral products released from bovine submaxillary mucin by β -elimination

Oligosaccharides were applied to a column (1.4 cm $\times$ 200 cm) of Bio-Gel P-4 (200–400 mesh), equilibrated and eluted with 0.05 M-ammonium acetate, pH 5.2, at a flow rate of 20 ml/h. Fractions of 2 ml were collected and assayed for carbohydrate [26]. The bars show the fractions that were combined. The arrows show the elution position of (1) starch and (2) glucose.

Abbreviations used: GalNAc, reduced *N*-acetylgalactosamine; COSY, correlation spectroscopy.

* To whom correspondence and reprint requests should be addressed.

Fig. 2. H.p.l.c. separation of oligosaccharides constituting peak N1 from Bio-Gel P-4 fractionation of the neutral oligosaccharide alditols released from bovine submaxillary mucin by β -elimination

Chromatography was performed on a column (25 cm $\times$ 4.6 mm i.d.) of Spherisorb S5 amino (Phase Separations Ltd, Queensferry, Clwyd, U.K.). The column was eluted with a mixture of acetonitrile (h.p.l.c. grade) and deionized distilled water containing 15 mM-potassium phosphate, pH 5.2, in a ratio of 4:1 followed by a linear gradient, reducing the acetonitrile by 0.10 %/min (-----). The flow rate was 2 ml/min. The arrows indicate the elution positions of (1) $\text{Fuca}(1 \rightarrow 2)\text{Gal}\beta(1 \rightarrow 3)\text{GalNAcol}$ and (2) $\text{GalNAc}\alpha(1 \rightarrow 3)[\text{Fuca}(1 \rightarrow 2)]\text{Gal}\beta(1 \rightarrow 3)\text{GalNAcol}$.

Fig. 3. H.p.l.c. separation of oligosaccharides constituting peak N2 from Bio-Gel P-4 fractionation of the neutral oligosaccharide alditols released from bovine submaxillary mucin by β -elimination

Chromatography was performed on a column (25 cm $\times$ 4.6 mm i.d.) of Spherisorb S5 amino (Phase Separations Ltd). The column was eluted with a mixture of acetonitrile (h.p.l.c. grade) and deionized distilled water containing 15 mM-potassium phosphate, pH 5.2, in a ratio of 4:1 for 30 min, followed by a linear gradient, reducing the acetonitrile by 0.20 %/min (-----). The flow rate was 2 ml/min. The arrows indicate the elution positions of (1) $\text{Fuca}(1 \rightarrow 2)\text{Gal}\beta(1 \rightarrow 3)\text{GalNAcol}$ and (2) $\text{GalNAc}\alpha(1 \rightarrow 3)[\text{Fuca}(1 \rightarrow 2)]\text{Gal}\beta(1 \rightarrow 3)\text{GalNAcol}$. * denotes a non-protein, non-carbohydrate, contaminant.

Methylation analysis

The partially methylated monosaccharides obtained on methanolysis and acetylation of the permethylated oligosaccharide alditols are indicated in Table 2.

^1H n.m.r. spectroscopy

In order to identify independently the constituent sugars, to determine the anomeric configurations of the linkages, and to obtain structural information on the linkage positions and the sequence of the sugar residues in the structures, each oligosaccharide was investigated by ^1H n.m.r. spectroscopy [8].

Peak N4, which was eluted from the Bio-Gel P-4 column with a volume of approximately two hexose units, was shown to

Fig. 4. H.p.l.c. separation of oligosaccharides constituting peak N3 from Bio-Gel P-4 fractionation of the neutral oligosaccharide alditols released from bovine submaxillary mucin by β -elimination

Chromatography was performed on a column (25 cm $\times$ 4.6 mm i.d.) of Spherisorb S5 amino (Phase Separations Ltd). The column was eluted with a mixture of acetonitrile (h.p.l.c. grade) and deionized distilled water containing 15 mM-potassium phosphate, pH 5.2, in a ratio of 84:16 (-----). The flow rate was 2 ml/min. The arrows indicate the elution positions of (1) GalNAcol and (2) $\text{Fuca}(1 \rightarrow 2)\text{Gal}\beta(1 \rightarrow 3)\text{GalNAcol}$.

Table 1. Carbohydrate composition of oligosaccharides obtained from bovine submaxillary mucin

Oligosaccharide	Monosaccharide residues present (mol/mol)				
	GalNAcol	GlcNAc	Gal	Fuc	GalNAc
N3a	1.00				
N3b	1.00	0.94			
N3c	1.00		1.06		
N3d	1.00		1.08	0.95	
N2b	1.00	0.90	1.12		
N2c	1.00	1.90			
N2d	1.00	0.85	0.90	0.08	0.09
N2f	1.00	1.85	0.93		
N2g	1.00	0.93	2.04		
N1e	1.00	1.80	0.92	0.90	
N1f	1.00	0.85	1.97	1.85	

contain only one component, N4a, by h.p.l.c. ^1H n.m.r. analysis (Table 3) identified the compound as the neutral monosaccharide alditol, GalNAcol [9]. ^1H n.m.r. analysis also identified h.p.l.c. fraction N3a as GalNAcol .

In the ^1H n.m.r. spectrum of fraction N3b the characteristic signals of GalNAcol H-2 at 4.290 p.p.m. and H-5 at 4.144 p.p.m. (Table 3) are typical of the core-type 3 structure, $\text{GlcNAc}\beta(1 \rightarrow 3)\text{GalNAcol}$ [10]. The chemical shifts of GalNAcol H-6 and H-6' were found to occur in the region 3.60–3.65 p.p.m. by examination of the COSY spectrum, confirming the absence of a substituent at GalNAcol C-6 [11]. The two *N*-acetyl methyl resonances of equal intensity at 2.085 and 2.037 p.p.m. have been assigned to GlcNAc^3 and GalNAcol respectively. The single anomeric signal at 4.606 p.p.m. ($J_{1,2} = 8.24$ Hz) is consistent with a terminal β -linked GlcNAc^3 residue [12,13]. The structure of N3b is identified as:

In the ^1H n.m.r. spectrum of N3c the GalNAcol H-2 signal is found at 4.397 p.p.m. (Table 3) indicating the occurrence of a core-type 1 structure [10]. The absence of a substituent at C-6 of GalNAcol is evident from the chemical shifts of H-5

Table 2. Molar ratios of monosaccharide methyl ethers present in the methanolysate of the permethylated reduced oligosaccharides

Me-3,4,6-Me₃Gal = methyl 3,4,6-*O*-methyl galactoside, etc. 1,4,5-Me₃-GalNAcNMe-ol was calculated as the sum of 1,4,5-Me₃-3,6-anhydro-GalNAcNMe-ol and 1,4,5-Me₃-GalNAcNMe-ol.

Per- <i>O</i> -acetylated derivative of	Oligosaccharide alditol (mol/mol)									
	N3b	N3c	N3d	N2b	N2c	N2d	N2f	N2g	N1e	N1f
Me-2,3,4-Me ₃ Fuc			0.85			0.08			0.92	1.80
Me-2,3,4,6-Me ₄ Gal		1.00		1.00		1.00	1.00	2.00		
Me-3,4,6-Me ₃ Gal			1.00						1.00	2.00
Me-4,6-Me ₂ Gal						0.10				
Me-3,6-Me ₂ GlcNAc				0.97			0.83	0.91	1.10	0.87
Me-3,4,6-Me ₃ GlcNAc	1.00				2.00	1.15	0.94		0.84	
Me-3,4,6-Me ₃ GalNAc						0.09				
1,4,5-Me ₃ -GalNAcNMe-ol					1.10	0.95	0.95	0.97	0.95	0.90
1,4,5,6-Me ₄ -GalNAcNMe-ol	0.96	1.13	0.95	0.94		0.09				

Table 3. ¹H n.m.r. chemical shift data for the constituent monosaccharides of the oligosaccharides N3a, N3b, N3c, N3d, N2b, and N2c derived from bovine submaxillary mucin

Residue	Reporter group	Chemical shift (p.p.m.) of:					
		N3a	N3b	N3c	N3d	N2b	N2c
GalNAcol	H-2	4.254	4.290	4.397	4.400	4.288	4.286
	H-3	3.848	3.996	4.066	4.092	4.003	3.988
	H-4	3.385	3.541	3.506	3.521	3.553	
	H-5	3.931	4.144	4.198	4.168	4.143	4.237
	H-6		3.60–3.65	3.60–3.65	3.6–3.7	3.6–3.65	~ 3.9
	H-6'		3.60–3.65	3.60–3.65	3.6–3.7	3.6–3.65	~ 3.7
	H-1		3.60–3.63		~ 3.8	3.55–3.60	3.60–3.63
	H-1'	3.743	3.60–3.63		~ 3.8	3.55–3.60	3.60–3.63
	NAc	2.056	2.037	2.051	2.047	2.038	2.045
GlcNAc ³	H-1		4.606			4.632	4.599
	H-2		3.7–3.8			~ 3.75	3.7–3.8
	H-3					3.7–3.75	
	H-4		3.439			3.532	
	H-5		3.45–3.50				
	H-6		3.952			4.021	
	H-6'		3.7–3.8				
	NAc		2.085			2.083	2.082
Gal ³	H-1			4.478	4.584		
	H-2			3.5–3.6	~ 3.7		
	H-4			3.902	3.928		
Fuc ^{2,3}	H-1				5.256		
	H-2				~ 3.8		
	H-5				4.281		
	H-6				1.244		
GlcNAc ⁶	H-1						4.541
	H-2						3.65–3.75
	NAc						2.064
Gal ⁴	H-1					4.455	
	H-2					3.50–3.55	
	H-3					3.60–3.65	
	H-4					3.925	

(4.198 p.p.m.), and H-6 and H-6' (3.60–3.65 p.p.m.) of GalNAcol [11]. The single β -anomeric signal at 4.478 p.p.m. ($J_{1,2} = 7.69$ Hz) is in agreement with that observed for the structure Gal β (1 $\rightarrow$ 3)GalNAcol [12,13], as is the occurrence of a single *N*-acetyl methyl resonance at 2.051 p.p.m., which is attributable to GalNAcol [10]. The assignment of structural reporter groups was confirmed by examination of the COSY spectrum. The structure of oligosaccharide N3c was determined to be:

In the ¹H n.m.r. spectrum of fraction N3d, the chemical shift of GalNAcol H-2 at 4.400 p.p.m. (Table 3) indicates the core-type 1 structure [10]. The absence of a substituent at GalNAcol C-6 is evident from the chemical shifts of H-5 (4.168 p.p.m.), H-6 (3.6–3.7 p.p.m.) and H-6' (3.6–3.7 p.p.m.) of GalNAcol [11]. The fucose structural reporter group signals, H-1 (5.256 p.p.m., $J_{1,2} = 3.66$ Hz), H-5 (4.281 p.p.m.) and H-6 (1.244 p.p.m.), reveal the presence of a Fuc residue linked α (1 $\rightarrow$ 2). The anomeric signal of the Gal³ residue at 4.584 p.p.m. ($J_{1,2} = 7.33$ Hz) shows the characteristic downfield shift observed on substitution by

Fuc α (1 $\rightarrow$ 2). Structural reporter groups, which were confirmed by examination of the COSY spectrum, are identical with those found previously for the trisaccharide alditol consisting of a fucosylated core-type 1 structure [10,12,14]. The structure of oligosaccharide N3d was determined to be:

Analysis of the ^1H n.m.r. spectrum of N2a, which was identical with that of N3b, identified it as the disaccharide alditol GlcNAc β (1 $\rightarrow$ 3)GalNAcol.

The ^1H n.m.r. spectrum of N2b (Table 3) showed the presence of a trisaccharide alditol consisting of two β -linked sugar residues in addition to GalNAcol in a core-type 3 structure [10,11]. The chemical shift (Table 3) of the H-1 doublet (4.632 p.p.m., $J_{1,2} = 7.33$ Hz), which has been assigned to the core GlcNAc 3 , as well as the *N*-acetyl methyl signal at 2.083 p.p.m., are typical of a GlcNAc residue substituted by Gal β (1 $\rightarrow$ 4) as opposed to Gal β (1 $\rightarrow$ 3) [10]. The H-1 doublet at 4.455 p.p.m. ($J_{1,2} = 7.51$ Hz) has therefore been attributed to Gal 4 . This chemical shift is typical of a Gal residue in a terminal position [10], as are the chemical shifts of H-2 and H-3, determined from the COSY and 1-relayed-COSY spectra. Assignment of H-2 of GlcNAc 3 from the COSY spectrum and hence H-3 of GlcNAc 3 (3.70–3.75 p.p.m.) from the 1-relayed-COSY spectrum confirms the presence of a β (1 $\rightarrow$ 4) linkage between Gal and GlcNAc rather than a β (1 $\rightarrow$ 3) linkage [15]. Assignment of the multiplet at 4.021 p.p.m. to H-6 of GlcNAc 3 is also in agreement with a β (1 $\rightarrow$ 4) linkage [13,16]. On the basis of these data, the structure of N2b is identified as:

Analysis of the ^1H n.m.r. spectrum and chemical shift data (Table 3) of h.p.l.c. fraction N2c reveals the presence of a trisaccharide alditol containing two β -anomeric linkages. The resonance positions of GalNAcol H-2 (4.286 p.p.m.) and H-5 (4.237 p.p.m.) and the three *N*-acetyl methyl signals of equal intensity at 2.045, 2.064 and 2.082 p.p.m., which have been assigned to GalNAcol, GlcNAc 6 and GlcNAc 3 respectively, are characteristic of the core-type 4 structure, GlcNAc β (1 $\rightarrow$ 3)-[GlcNAc β (1 $\rightarrow$ 6)]GalNAcol [10,13,17]. Assignment of H-6 of GalNAcol (δ approx. 3.9 p.p.m.) via its connectivity to H-5 seen in the COSY spectrum is in agreement with substitution of GalNAcol at C-6 by β GlcNAc [10,11]. The two β -anomeric signals at 4.599 p.p.m. ($J_{1,2} = 8.42$ Hz) and 4.541 p.p.m. ($J_{1,2} = 8.24$ Hz) have been assigned to unsubstituted GlcNAc 3 and GlcNAc 6 respectively, in accordance with the trends observed for (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 6)-linked β GlcNAc residues. The structure of oligosaccharide N2c is therefore determined as:

Examination of the ^1H n.m.r. spectrum of fraction N2d (Table 4) showed it to be composed of one major and one minor component. The major component, designated N2d(i), is a trisaccharide alditol consisting of the core-type 2 element, Gal β (1 $\rightarrow$ 3)[GlcNAc β (1 $\rightarrow$ 6)]GalNAcol. This is evident from the chemical shifts (Table 4) of GalNAcol H-2 (4.395 p.p.m.) and H-5 (4.279 p.p.m.) along with the *N*-acetyl methyl signals for GalNAcol and GlcNAc 6 which coincide at 2.066 p.p.m., which is consistent with the core-type 2 structure [10,13,18]. The H-1 doublets at 4.466 and 4.539 p.p.m., assigned to Gal 3 and GlcNAc 6

respectively, indicate that they are unsubstituted. The structure of the major component N2d(i) has been determined to be:

The presence of a minor component, designated N2d(ii), in fraction N2d is inferred from the low-intensity signals in the α -anomeric, *N*-acetyl and fucose CH_3 regions of the spectrum (Table 4). The α -anomeric signal at 5.389 p.p.m. is typical of that found for Fuc α (1 $\rightarrow$ 2) in the blood-group A determinant [19]. On examination of the COSY spectrum of fraction N2d, the α -anomeric doublet at 5.191 p.p.m. ($J_{1,2} = 3.3$ Hz) is seen to be coupled to a proton (its H-2 proton) in the spectral region 4.2–4.3 p.p.m. This is consistent with the presence of an α GalNAc residue [5]. The occurrence of an anomeric signal in the spectral region 4.7–4.8 p.p.m. (hidden under the HO 2 H signal) is suggested from the COSY spectrum by the appearance of a single connectivity between a peak in this region and a peak (presumably H-2) in the region 3.9–4.0 p.p.m. The chemical shifts of these anomeric protons and those of their coupled H-2 protons agree well with those reported by Dua *et al.* for the structure Fuc α (1 $\rightarrow$ 2)[GalNAc α (1 $\rightarrow$ 3)]Gal β (1 $\rightarrow$ 3)GalNAcol, possessing the blood-group A determinant [19]. The structure of the minor component N2d(ii) is therefore proposed as:

^1H n.m.r. analysis of h.p.l.c. fraction N2e showed it to be a non-carbohydrate contaminant.

The ^1H n.m.r. spectrum of fraction N2f showed the presence of a tetrasaccharide alditol containing three β -anomeric linkages. The chemical shifts of the GalNAcol structural reporter groups H-2 and H-5, found at $\delta = 4.281$ p.p.m. and $\delta = 4.239$ p.p.m. respectively (Table 4), are indicative of the GlcNAc β (1 $\rightarrow$ 3)[GlcNAc β (1 $\rightarrow$ 6)]GalNAcol core-type 4 structure [10,13,17]. The H-1 doublets at $\delta = 4.598$ p.p.m. ($J_{1,2} = 8.8$ Hz) and $\delta = 4.566$ p.p.m. ($J_{1,2} = 8.6$ Hz) have been attributed to GlcNAc β (1 $\rightarrow$ 3) and GlcNAc β (1 $\rightarrow$ 6) respectively. The H-1 value for GlcNAc β (1 $\rightarrow$ 3), together with the chemical shift of its *N*-acetyl methyl protons ($\delta = 2.081$ p.p.m.), indicate its terminal position within the tetrasaccharide alditol [13,17]. Furthermore the assignment of H-2 of GlcNAc 3 from the COSY spectrum and hence H-3 of GlcNAc 3 (δ approx. 3.57 p.p.m.) from the 1-relayed-COSY is in agreement with an unsubstituted GlcNAc 3 residue. On the other hand, the chemical shift of H-1 of the GlcNAc β (1 $\rightarrow$ 6) residue and that of its *N*-acetyl methyl protons ($\delta = 2.061$ p.p.m.) indicate that it is substituted by Gal in the β (1 $\rightarrow$ 4) linkage [13]. Such substitution by a terminal Gal residue is also evident from the chemical shift of H-1 of Gal ($\delta = 4.474$ p.p.m., $J_{1,2} = 7.88$ Hz), which distinguishes between the structural elements Gal β (1 $\rightarrow$ 4)GlcNAc β (1 $\rightarrow$ 3) and Gal β (1 $\rightarrow$ 4)GlcNAc β (1 $\rightarrow$ 6) [10,13]. Also, the chemical shift of H-3 of GlcNAc 6 is known to occur in the region 3.7–3.8 p.p.m. (Table 4), which is consistent with Gal β (1 $\rightarrow$ 4) as opposed to

Table 4. ¹H n.m.r. chemical shift data for the constituent monosaccharides of the oligosaccharides N2d(i), N2d(ii), N2f, N2g, N1e and N1f derived from bovine submaxillary mucin

Residue	Reporter group	Chemical shift (p.p.m.) in:					
		N2d(i)	N2d(ii)	N2f	N2g	N1e	N1f
GalNAcol	H-2	4.395		4.281	4.393	4.282	4.407
	H-3	4.062		3.988	4.060		4.085
	H-4	3.469			3.460		3.497
	H-5	4.279		4.239	4.282	4.243	4.290
	H-6	3.9–4.0		~ 3.9	~ 3.90		3.85–3.9
	H-6'	3.6–3.7		~ 3.7	3.6–3.7		3.65–3.7
	H-1			3.60–3.65	3.7–3.8		~ 3.75
	H-1'			3.60–3.65	3.7–3.8		~ 3.75
GlcNAc ⁶	NAc	2.066	2.049	2.045	2.068	2.045	2.060
	H-1	4.539		4.566	4.559	4.541	4.550
	H-2	3.7–3.8		3.7–3.8	3.75–3.80		3.6–3.8
	H-3			3.7–3.8	3.70–3.75		
	H-4	3.428					
	H-6	3.9–4.0		3.998			4.000
Gal ³	NAc	2.066		2.061	2.068	2.066	2.060
	H-1	4.466	4.7–4.8		4.464		4.573
	H-2	3.5–3.6	3.9–4.0		3.5–3.6		3.6–3.8
	H-3				3.6–3.7		
	H-4	3.902			3.900		3.899
GalNAc ³	H-1		5.191			4.597	
	H-2		4.2–4.3				
	NAc		2.049			2.081	
Fuc ^{2,3}	H-1		5.389				
	H-2		3.80–3.85				
	H-5		4.2–4.3				
	H-6		1.236				
GlcNAc ³	H-1			4.598			
	H-2			3.7–3.8			
	H-3			~ 3.57			
	H-4			3.427			
	H-5			~ 3.45			
	H-6			3.951			
	NAc			2.081			
Gal ⁴	H-1			4.474	4.470	4.541	4.540
	H-2			~ 3.55	3.5–3.6		3.6–3.8
	H-3			~ 3.65	3.6–3.7		
	H-4			3.927	3.926		3.923
Fuc ^{2,4}	H-1					5.310	5.310
	H-2						3.75–3.8
	H-5					4.230	4.229
	H-6					1.233	1.232
Fuc ^{2,3}	H-1						5.220
	H-2						3.75–3.8
	H-5						4.254
	H-6						1.243

Gal β (1 $\rightarrow$ 3) [15]. On the basis of these data, the structure of N2f is identified as:

Analysis of the ^1H n.m.r. spectrum of fraction N2g (Table 4) indicates the presence of a tetrasaccharide alditol containing three β -anomeric linkages. The occurrence of the core-type 2 structure is evident from the chemical shifts of the GalNAc

structural reporter group signals H-2 (4.393 p.p.m.), H-5 (4.282 p.p.m.) and H-6 (approx. 3.9 p.p.m., assigned from the COSY spectrum) along with the *N*-acetyl methyl signals for GalNAc_{ol} and GlcNAc⁶ which coincide at 2.068 p.p.m. [10,13,18]. The H-1 doublet at 4.559 p.p.m. ($J_{1,2} = 7.69$ Hz), which has been attributed to the GlcNAc⁶ residue, indicates that it is substituted by Gal in the $\beta(1 \rightarrow 4)$ linkage [10]. The observation that the chemical shift of H-3 of GlcNAc⁶ occurs in the region 3.70–3.75 p.p.m. is in agreement with substitution by GalB(1 $\rightarrow$ 4) rather than GalB(1 $\rightarrow$ 3) [15].

The chemical shift of GlcNAc⁶ H-3 was assigned by examination of the 1-relayed-COSY spectrum, GlcNAc⁶ H-2 having already been assigned from the COSY spectrum. The β -anomeric signals at 4.464 p.p.m. and 4.470 p.p.m. have been assigned to

Table 5. Structure and approximate molar ratios of the neutral oligosaccharides isolated from bovine submaxillary mucin

Code	Structure	Molar ratio (mol %)
N3b (N2a)	GlcNAc β (1 $\rightarrow$ 3)GalNAcol	45.5
N3c	Gal β (1 $\rightarrow$ 3)GalNAcol	13.0
N3d	Fuc α (1 $\rightarrow$ 2)Gal β (1 $\rightarrow$ 3)GalNAcol	8.5
N2b	Gal β (1 $\rightarrow$ 4)GlcNAc β (1 $\rightarrow$ 3)GalNAcol	5.0
N2c	GlcNAc β 1 6 GalNAcol 3	4.0
N2d(i)	GlcNAc β 1 6 GalNAcol 3 Gal β 1	4.5
N2d(ii)	GalNAc α 1 3 Gal β (1 $\rightarrow$ 3)GalNAcol 2 Fuc α 1	0.5
N2f (N1c)	Gal β (1 $\rightarrow$ 4)GlcNAc β 1 6 GalNAcol 3 GlcNAc β 1	4.5
N2g (N1d)	Gal β (1 $\rightarrow$ 4)GlcNAc β 1 6 GalNAcol 3 Gal β 1	8.0
N1e	Fuc α (1 $\rightarrow$ 2)Gal β (1 $\rightarrow$ 4)GlcNAc β 1 6 GalNAcol 3 GlcNAc β 1	2.5
N1f	Fuc α (1 $\rightarrow$ 2)Gal β (1 $\rightarrow$ 4)GlcNAc β 1 6 GalNAcol 3 Fuc α (1 $\rightarrow$ 2)Gal β 1	4.0

the linear trisaccharide N2b was eluted before the branched trisaccharide N2d(i). The co-elution of the trisaccharide N2d(i) with the tetrasaccharide N2d(ii) is attributable to the presence of both Fuc α and GalNAc α in the latter structure [5].

Quantification of the oligosaccharides using h.p.l.c. with a u.v. detector operating at 195 nm, and with regard to the number of chromophores in each oligosaccharide [20], indicated that the oligosaccharides were present in the approximate molar regions shown in Table 5. Integration of the anomeric signals in the n.m.r. spectrum of N2d indicated that N2d(i) and N2d(ii) were present in the approximate molar ratio 4.5:0.5.

Previous investigations of the neutral oligosaccharides of bovine submaxillary mucin have led to the characterization of two disaccharides, a trisaccharide and a tetrasaccharide, in addition to GalNAcol [1,2]. These correspond to structures N2d(i), N3b, N3c and N4a of the present study. The tetrasaccharide Fuc α (1 $\rightarrow$ 2)Gal β (1 $\rightarrow$ 3)[GlcNAc β (1 $\rightarrow$ 6)]GalNAcol, previously found by Tsuji & Osawa [2], has not been identified here. However, 12 neutral structures, ranging in size from mono- to hexa-saccharide, eight of which have not previously been reported in bovine submaxillary mucin, have been characterized. Eleven of these structures have previously been isolated from other mucin sources, in particular, cystic fibrosis bronchial mucins and ovarian cyst mucins [10,12,13,19]. Although oligosaccharide N2c has been identified as a core structure in many

mucins (core-type 4) [10,17,22], the occurrence of this oligosaccharide as a distinct entity is novel. ^1H n.m.r. data for this structure are presented here for the first time. Oligosaccharides N2b, N3b and N3c represent the asialo analogues of the acidic oligosaccharides A1/2a/A1/2c [4,6], A4c/A4f [5] and A4d [3] respectively. Interestingly, the asialo analogues of the structures A1/2b, Fuc α (1 $\rightarrow$ 2)Gal β (1 $\rightarrow$ 4)GlcNAc β (1 $\rightarrow$ 3)GalNAcol [3,6] and A4b/A4e [3], GalNAc α (1 $\rightarrow$ 3)GalNAcol, have not been found among the neutral oligosaccharides. They may, however, be present in minor amounts. The occurrence of GalNAc α (1 $\rightarrow$ 3)-GalNAcol as a minor component in peak N3 from the Bio-Gel P-4 column was suggested by the presence of an α -anomeric signal at 5.103 p.p.m. in the ^1H n.m.r. spectrum of this mixture, before h.p.l.c. purification. This structure, however, was not identified by ^1H n.m.r. analysis of any of the major h.p.l.c.-purified components of peak N3.

All of the acidic and neutral structures reported here are biosynthetically allowed, assuming that the biosynthesis of oligosaccharides in bovine submaxillary mucin proceeds in the same manner as has been reported for other mucins [23,24]. It has previously been suggested that the activity of the enzyme *N*-acetylglucosaminide β -(1 $\rightarrow$ 4)-galactosyltransferase, which catalyses the synthesis of the Gal β (1 $\rightarrow$ 4)GlcNAc sequence, is relatively low in bovine submaxillary glands [2]. This suggestion was made because no carbohydrate chain bearing this sequence

had previously been characterized. The finding in this study of eight structures possessing the Gal β (1 $\rightarrow$ 4)GlcNAc β (1 $\rightarrow$ 3/6)-sequence shows that this enzyme is indeed active in bovine submaxillary glands.

Earlier investigations of the immunological properties of bovine submaxillary mucin showed some blood-group-A and -H activity, but no blood-group-B activity [25]. The present study is in agreement with such findings. Oligosaccharides N1e and N1f contain the blood-group-H type 2 determinant, while oligosaccharide N3d contains the blood-group-H type 3 determinant, and oligosaccharide N2d(ii) represents the blood-group-A determinant.

The fact that bovine submaxillary mucin has now been shown to contain all the five known mucin core types, demonstrating the heterogeneous nature of its carbohydrate chains, supports its use as a model mucin to relate biological activity to glycoconjugate structure.

We thank Professor Dirk van den Eijnden, Vrije Universiteit, Amsterdam, The Netherlands, for performing the methylation analyses and for helpful discussion.

REFERENCES

- Bertolini, M. & Pigman, W. (1970) *Carbohydr. Res.* **14**, 53–63
- Tsuji, T. & Osawa, T. (1986) *Carbohydr. Res.* **151**, 391–402
- D'Arcy, S. M., Donoghue, C. M., Koeleman, C., Van den Eijnden, D. H. & Savage, A. V. (1989) *Biochem. J.* **260**, 389–393
- D'Arcy, S. M., Donoghue, C. M. & Savage, A. V. (1989) *Proc. R. Ir. Acad. Sect. B* **89**, 125–134
- Savage, A. V., Donoghue, C. M., D'Arcy, S. M., Koeleman, C. & Van den Eijnden, D. H. (1990) *Eur. J. Biochem.* **192**, 427–432
- Savage, A. V., Donoghue, J. J., Koeleman, C. & Van den Eijnden, D. H. (1990) *Eur. J. Biochem.* **193**, 837–843
- Savage, A. V., Donoghue, C. M., D'Arcy, S. M., Koeleman, C. & Van den Eijnden, D. H. (1988) *Proc. Int. Carbohydr. Symp.* 14th, Stockholm, Abstr. C76
- Vliegthart, J. F. G., Dorland, L. & Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41**, 209–374
- Van Halbeek, H., Dorland, L., Vliegthart, J. F. G., Fiat, A.-M. & Jollès, P. (1980) *Biochim. Biophys. Acta* **623**, 295–300
- Van Halbeek, H., Dorland, L., Vliegthart, J. F. G., Hull, W. E., Lamblin, G., Lhermitte, M., Boersma, A. & Roussel, P. (1982) *Eur. J. Biochem.* **127**, 7–20
- Van Halbeek, H., Dorland, L., Vliegthart, J. F. G., Fiat, A.-M. & Jollès, P. (1981) *FEBS Lett.* **133**, 45–50
- Dua, V. K., Dube, V. E. & Bush, C. A. (1984) *Biochim. Biophys. Acta* **802**, 29–40
- Lamblin, G., Boersma, A., Lhermitte, M., Roussel, P., Mutsaers, J. H. G. M., Van Halbeek, H. & Vliegthart, J. F. G. (1984) *Eur. J. Biochem.* **143**, 227–236
- Van Halbeek, H., Dorland, L., Vliegthart, J. F. G., Kochetkov, N. K., Arbatsky, N. P. & Derevitskaya, V. A. (1982) *Eur. J. Biochem.* **127**, 21–29
- Gleeson, P. A., Feeney, J., Mills, G. & Hughes, R. C. (1984) *Eur. J. Biochem.* **144**, 143–150
- Breg, J., Van Halbeek, H., Vliegthart, J. F. G., Lamblin, G., Houvenaghel, M.-C. & Roussel, P. (1987) *Eur. J. Biochem.* **168**, 57–68
- Breg, J., Van Halbeek, H., Vliegthart, J. F. G., Klein, A., Lamblin, G. & Roussel, P. (1988) *Eur. J. Biochem.* **171**, 643–654
- Klein, A., Lamblin, G., Lhermitte, M., Roussel, P., Breg, J., Van Halbeek, H. & Vliegthart, J. F. G. (1988) *Eur. J. Biochem.* **171**, 631–642
- Dua, V. K., Rao, B. N. N., Wu, S.-S., Dube, V. E. & Bush, C. A. (1986) *J. Biol. Chem.* **261**, 1599–1608
- Savage, A. V., Koppen, P. L., Schiphorst, W. E. C. M., Trippelvit, L. A. W., Van Halbeek, H., Vliegthart, J. F. G. & Van den Eijnden, D. H. (1986) *Eur. J. Biochem.* **160**, 123–129
- Blanken, W. M., Bergh, M. L. E., Koppen, P. L. & Van den Eijnden, D. H. (1985) *Anal. Biochem.* **145**, 322–330
- Hounsell, E. F., Fukuda, M., Powell, M. E., Feizi, T. & Hakomori, S. (1980) *Biochem. Biophys. Res. Commun.* **92**, 1143–1150
- Brockhausen, I., Matta, K. L., Orr, J., Schachter, H., Koenderman, A. H. L. & Van den Eijnden, D. H. (1986) *Eur. J. Biochem.* **157**, 463–474
- Beyer, T. A., Rearick, J. I., Paulson, J. C., Priels, J.-P., Sadler, J. E. & Hill, R. L. (1979) *J. Biol. Chem.* **254**, 12531–12541
- Gottschalk, A., Bhargava, A. S. & Murty, V. L. N. (1972) in *Glycoproteins – Their Composition, Structure and Function* (Gottschalk, A., ed.), vol. 5, part B, pp. 817–821, Elsevier, Amsterdam
- Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, K. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356

Received 12 February 1991/26 March 1991; accepted 29 March 1991