Identification of a novel oligosaccharide backbone structure with a galactose residue monosubstituted at C-6 in human foetal gastrointestinal mucins

Elizabeth F. HOUNSELL,*§ Alexander M. LAWSON,† James FEENEY,‡ Geoffrey C. CASHMORE,† David P. KANE,* Mark STOLL* and Ten FEIZI*  

*Section of Glycoconjugate Research and †Section of Clinical Mass Spectrometry, Clinical Research Centre, Watford Road, Harrow HA1 3UJ, U.K., and ‡M.R.C. Biomedical NMR Centre, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

An oligosaccharide purified from a major penta- to hexa-saccharide fraction of human meconium glycoproteins has been shown by m.s. and n.m.r. analysis to have a novel backbone structure containing an internal galactose residue monosubstituted at C-6 by N-acetylgalcosamine:

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

This oligosaccharide may represent a biosynthetic product of a previously unrecognized N-acetylgalcosaminyltransferase catalysing formation of a linear GlcNAcβ1-6Gal sequence.

INTRODUCTION

Mucin-type glycoproteins are a rich source of oligosaccharides of oncodevelopmental type, i.e. antigens of foetal origin that accumulate in human tumours (Feizi et al., 1984). The relative abundance of mucin oligosaccharides has allowed for detailed structural and antigenic studies that have led to the characterization of the blood-group antigens as peripheral glycosylations that mask antigens (e.g. li) expressed in backbone regions (Kabat, 1973; Watkins, 1980a,b; Hounsell & Feizi, 1982; Feizi et al., 1984). In the absence of peripheral substitutions, epitopes on the backbone and core-region oligosaccharides are revealed (e.g. the li antigens and the receptor for peanut lectin). The antigenic changes in both normal development and oncogenesis can occur by either incomplete or extended oligosaccharide chain biosynthesis (Feizi, 1985; Hakomori, 1986).

The present study forms part of a programme to characterize oligosaccharides of oncodevelopmental type from meconium glycoproteins that are representative of human foetal gastrointestinal mucins. Earlier studies of the mono- to tetra-saccharides of a blood-group-O meconium pool revealed the presence of two core-region sequences not found so far in other human mucins, GlcNAcβ1-6GalNAc-ol and GalNAcβ1-3GalNAc-ol (Hounsell et al., 1985; Feeney et al., 1986). The present paper reports the identification of a novel backbone sequence having an internal galactose monosubstituted at C-6 by N-acetylgalcosamine.

MATERIALS AND METHODS

Oligosaccharide isolation

Glycoproteins were prepared from a blood-group-O-active meconium pool as described previously (Hounsell et al., 1985). The glycoproteins were depleted of li antigenic activity by immunoaffinity chromatography and defucosylated by mild acid hydrolysis (0.05 M-H₂SO₄ at 100 °C for 1 h) in order to reduce the oligosaccharide heterogeneity. Oligosaccharides were isolated by mild base/borohydride treatment (0.05 M-NaOH in 1 M-NaBH₄ at 50 °C for 16 h) and separated on a Bio-Gel P4 column (Hounsell et al., 1985). The oligosaccharides N1 and K2(4) described in the present paper were purified by h.p.l.c. from fractions N and K, which were eluted from the Bio-Gel P4 column in the regions of 7–8 and 10 glucose units respectively. The purification and structural assignment of oligosaccharide N1, Galβ1-4GlcNAcβ1-6(Galβ1-3)GalNAc-ol, has been reported previously (Hounsell et al., 1985). The p.l.c. of fraction K was performed sequentially on ODS-Hypersil and APS-Hypersil columns (4 mm × 250 mm; Shandon Southern Products, Runcorn, Cheshire, U.K.) with a Spectra-Physics SP9700 solvent-delivery system and SP8400 variable-wavelength detector at 208 nm (Spectra-Physics, Santa Clara, CA, U.S.A.). Fraction K2 was eluted at 5.4 min (solvent front 1.9 min) in 100% water on the ODS column, and fraction K2(4) was eluted at 11.7 min on the APS column with a gradient of 65% (v/v) acetonitrile in water to 60% (v/v) acetonitrile in water in 10 min (solvent front 2.0 min).

M.s.

Oligosaccharide K2(4) was analysed by l.s.i.-m.s. in negative-ion mode with a matrix of glycerol/thioglycerol (1:1, v/v). The oligosaccharide (5 nmol) was methylated by a procedure adopted from Ciucanu & Kerek (1984) as described previously (Scudder et al., 1987), and a sample was analysed by positive-ion l.s.i.-m.s. with the same matrix. For methylation analysis (Hakomori, 1964; Lind-
Fig. 1. (a) L.s.i.-m.s. of native oligosaccharide K2(4) in negative-ion mode and (b) L.s.i.-m.s. of permethylated oligosaccharide K2(4) in positive-ion mode

(a) L.s.i.-m.s. of the native oligosaccharide in negative-ion mode showing the major molecular ion m/z 1114 corresponding to Hex$_2$HexNAc$_2$HexNAc-ol and minor molecular ions m/z 1098, 952 and 749 corresponding to Fuc$_1$Hex$_2$HexNAc$_2$HexNAc-ol, Hex$_2$HexNAc$_2$HexNAc-ol and Hex$_2$HexNAc$_2$HexNAc-ol respectively. Asterisks (*) indicate ions from minor components. (b) L.s.i.-m.s. of the permethylated oligosaccharide in positive-ion mode. The molecular ions m/z 1410.7, 1380.7, 1206.6 and 961 correspond to the permethylated oligosaccharides giving rise to m/z 1114, 1098, 952 and 749 respectively in (a) representing one major and three minor components. Asterisks (*) indicate ions from minor components.

berg, 1972) the remaining permethylated oligosaccharide was further modified to give partially methylated [H]-alditol acetates as described by Scudder et al. (1987) and analysed by combined g.c. and e.i.-m.s. with a 25 m fused-silica capillary column (0.22 mm internal diam.) of CP-Sil5-CB with injection temperature 200 °C and temperature programme 50 °C for 0.5 min (splitless on column injection) then 25 °C/min to 150 °C, 5 °C/min to 220 °C and 25 °C/min to 280 °C. E.i. mass spectra at 70 eV were recorded every 3 s with a Jeol DX-303 mass spectrometer. L.s.i.-m.s. was carried out on a VG Analytical ZAB2-E instrument.

**N.m.r spectroscopy**

The 500 MHz $^1$H-n.m.r. spectra of oligosaccharides K2(4) and N1 were obtained at 22 °C in $^2$H$_2$O with a Bruker AM500 spectrometer operating in the Fourier-transform mode and equipped with an Aspect 3000 computer. Resolution enhancement was carried out by a Gaussian multiplication technique (Ferrige & Lindon, 1979). Proton assignments were made from the one-dimensional spectra, spin decoupling and phase-sensitive COSY experiments carried out with a ‘double-quantum filter’ and phase cycling (Piantini et al., 1982; States et al., 1982; Marion & Wuthrich, 1983). Chemical shifts were measured in p.p.m. from acetone, which resonates at 2.225 p.p.m. from sodium 4,4-dimethyl-4-silapentane-1-sulphonate at 22 °C. Coupling constants were obtained from a first-order analysis of the spectra.

**RESULTS**

Oligosaccharide K2(4) made up approx. 10% of the penta- and hexa-saccharide fraction K and hence represents approx. 1% of the oligosaccharides from the meconium pool studied (Hounsell et al., 1985).

L.s.i.-m.s of oligosaccharide K2(4) showed a major component with an $M_r$ of 1115 corresponding to Hex$_2$-HexNAc$_2$HexNAc-ol giving an [M − H]$^+$ ion in negative-ion mode of the native oligosaccharide of m/z 1114.
Novel oligosaccharide of human foetal gastrointestinal mucin

(Fig. 2) and an $[M + H]^+$ ion of the permethylated oligosaccharide in positive-ion mode of $m/z$ 1410.7 (Fig. 1b). L.s.i.-m.s. of the permethylated oligosaccharide also identified chain-terminating Hex1-4HexNAc sequences giving the ion pair $m/z$ 464 and 432 formed by fragmentation at the anomeric oxygen atom of $N$-acetylhexosamine residues and loss of methanol (suggestive of a 1–4 linkage). The absence of an ion at $m/z$ 913 shows that the longer linear sequence Hex–HexNAc–Hex–HexNAc– is not present in this oligosaccharide. The relatively low abundance of the fragmentation ions of $m/z$ 260, 228 and 709 compared with $m/z$ 464 and 432 show that HexNAc, HexNAc–Hex–HexNAc or Hex–HexNAc–HexNAc are not major component sequences.

The g.c.–m.s. analysis of the partially methylated $\text{[}^2\text{H}\text{]}$alditol acetates of this oligosaccharide (Fig. 2) gave approximate 2:1:1:2 proportions of terminal non-reducing galactose, 6-linked galactose, 3,6-linked $N$-acetylglactosaminol and 4-linked $N$-acetylglucosamine. These results, together with the l.s.i.-m.s. data, are indicative of the oligosaccharide:

\[
\text{Gal}1\text{-4GlcNAc} - 6\text{/3GlcNAc-ol}
\]

\[
\text{Gal}1\text{-4GlcNAc} - 6\text{Gal1-3} \text{GlcNAc-ol}
\]

The 500 MHz $^1$H-n.m.r. spectrum of oligosaccharide K2(4) is shown in Fig. 3(a) with the $^1$H-$^1$H homonuclear COSY spectrum (Fig. 3b), from which a near-complete assignment of the signals could be made. The chemical shifts and coupling constants are given in Table I together with the chemical shifts for the major tetrasaccharide of meconium glycopeptides designated N1, Gal/$\beta$1–4GlcNAc/$\beta$1–6(Gal/$\beta$1–3)GalNAc-ol (Hounsell et al., 1985), for which a near-complete proton assignment has now been achieved by using a COSY experiment (spectrum not shown). For oligosaccharide K2(4) the proportions of the $\beta$ anomic signals at 4.45–4.6 p.p.m. were shown in the non-deconvoluted spectrum to be 1:1:3, the first two signals having the coupling constant indicative of an $N$-acetylglucosamine residue (8.3 Hz) and the third signal that for galactose (7.8 Hz). The

<table>
<thead>
<tr>
<th>Chemical shift (p.p.m.)</th>
<th>Coupling constant (Hz)</th>
<th>Chemical shift (p.p.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-1</td>
<td>3.76</td>
<td>11.4</td>
</tr>
<tr>
<td>H-1'</td>
<td>3.76</td>
<td>7.5</td>
</tr>
<tr>
<td>H-2</td>
<td>4.389</td>
<td>4.400</td>
</tr>
<tr>
<td>H-3</td>
<td>4.013</td>
<td>4.063</td>
</tr>
<tr>
<td>H-4</td>
<td>3.476</td>
<td>3.457</td>
</tr>
<tr>
<td>H-5</td>
<td>4.235</td>
<td>4.289</td>
</tr>
<tr>
<td>H-6</td>
<td>3.946</td>
<td>3.986</td>
</tr>
<tr>
<td>H-6'</td>
<td>3.70</td>
<td>10.3</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>4.599, 4.556</td>
<td>8.3</td>
</tr>
<tr>
<td>H-1</td>
<td>3.76</td>
<td>11.4</td>
</tr>
<tr>
<td>H-2</td>
<td>3.73</td>
<td>8.2</td>
</tr>
<tr>
<td>H-3</td>
<td>3.711</td>
<td>1</td>
</tr>
<tr>
<td>H-5</td>
<td>3.602, 3.593</td>
<td>2.1</td>
</tr>
<tr>
<td>H-6</td>
<td>4.002, 3.998</td>
<td>5.0</td>
</tr>
<tr>
<td>H-6'</td>
<td>3.858, 3.842</td>
<td>12.3</td>
</tr>
<tr>
<td>Gal/$\beta$1-4</td>
<td>4.470</td>
<td>7.8</td>
</tr>
<tr>
<td>H-1</td>
<td>4.470</td>
<td>7.8</td>
</tr>
<tr>
<td>H-2</td>
<td>3.537</td>
<td>9.9</td>
</tr>
<tr>
<td>H-3</td>
<td>3.668</td>
<td>3.4</td>
</tr>
<tr>
<td>H-4</td>
<td>3.922</td>
<td>3.923</td>
</tr>
<tr>
<td>Gal/$\beta$1-3</td>
<td>4.470</td>
<td>7.8</td>
</tr>
<tr>
<td>H-1</td>
<td>4.470</td>
<td>7.8</td>
</tr>
<tr>
<td>H-2</td>
<td>3.547</td>
<td>9.9</td>
</tr>
<tr>
<td>H-3</td>
<td>3.672</td>
<td>3.4</td>
</tr>
<tr>
<td>H-4</td>
<td>3.882</td>
<td>3.897</td>
</tr>
<tr>
<td>NAc</td>
<td>2.066, 2.055</td>
<td>2.068, 2.065</td>
</tr>
</tbody>
</table>

Vol. 256
Fig. 3. (a) 500 MHz 1H-n.m.r. resolution-enhanced spectrum of oligosaccharide K2(4) (500 nmol) and (b) 1H–1H homonuclear phasesensitive COSY spectrum of oligosaccharide K2(4) (500 nmol) showing some of the cross-peak assignments.

Asterisks (*) indicate peaks due to non-carbohydrate contaminants.

critical shifts for the anomeric protons of galactose and for the H-2–H-4 protons of two of the three galactose residues (Fig. 3 and Table 1) are identical with those for galactose in the sequence Galβ1–4GlcNAcβ1–6 of the tetrasaccharide N1.

An indication that the core N-acetylgalactosaminol of oligosaccharide K2(4) is substituted at C-6 by N-acetylglicosamine and at C-3 by galactose is given by the multiplet at 4.389 p.p.m., which has a similar shift to H-2 of N-acetylgalactosaminol substituted at C-6 by N-acetylglicosamine and at C-3 by galactose or at C-3 by galactose alone (e.g. 4.400 and 4.393 p.p.m. respectively for the tetrasaccharide N1 and the disaccharide Galβ1–3GalNAc-ol of meconium; Hounsell et al., 1985), but is dissimilar to the chemical shift of H-2 of N-acetylgalactosaminol substituted at C-3 or C-6 only by N-acetylglicosamine (e.g. 4.294 and 4.247 p.p.m. respectively for the trisaccharides Galβ1–4GlcNAcβ1–3GalNAc-ol and Galβ1–4GlcNAcβ1–6GalNAc-ol of meconium; Hounsell et al., 1985; Feeney et al., 1986) and at C-6 and C-3 by N-acetylglicosamine [e.g. 4.286 p.p.m. for Galβ1–4GlcNAcβ1–6(Galβ1–4GlcNAcβ1–3)GalNAc-ol; Lamblin et al. (1984)]. The remainder of the signals for N-acetylgalactosaminol assigned from the two-dimensional COSY spectrum of oligosaccharide K2(4) are similar to those for the tetrasaccharide N1, the greatest divergence being for the H-3 and H-5 signals, which are 0.05 p.p.m. upfield of the corresponding signals in the tetrasaccharide N1. The chemical shifts for N-acetylgalactosaminol of oligo-
saccharide K2(4) are in closest agreement with those given by Van Halbeek et al. (1982) for a monosaccharide of pig blood-group-H gastric mucins with the structure:

\[ \text{GlcNAc}\beta_\text{l--6GalNAc-ol} \]
\[ \text{Fuc}\beta_\text{l--2Gal}\beta_\text{l--4GlcNAc}\beta_\text{l--6Gal}\beta_\text{l--3} \]

\[ \text{(N-acetylgalactosaminitol H-2 4.392 p.p.m., H-3 3.993 p.p.m., H-4 3.472 p.p.m. and H-5 4.219 p.p.m.)} \]

It is thus concluded that oligosaccharide K2(4) has the sequence:

\[ \text{Gal}\beta_\text{l--4GlcNAc}\beta_\text{l--6GalNAc-ol} \]
\[ \text{Gal}\beta_\text{l--4GlcNAc}\beta_\text{l--6Gal}\beta_\text{l--3} \]

and it is suggested that in solution the conformation of oligosaccharide K2(4) and the nonasaccharide shown above is such that the Gal\beta_\text{l--4GlcNAc} sequence joined by 1--6 linkage to galactose lies in close proximity to the N-acetylgalactosaminitol and hence perturbs the shielding of H-3 and H-5 of N-acetylgalactosaminitol, compared with the tetrasaccharide N1 for example where there is no substituent on the Gal\beta_\text{l--3GalNAc-ol} branch. Comparison of the chemical shifts for the N-acetylgalactosamine residues in oligosaccharides K2(4) and N1 (Table 1) and of their molecular models suggests that it is the C-1 and N-acetamido region of N-acetylgalactosamine linked \beta_\text{l--6} to galactose in oligosaccharide K2(4) that interacts with the face of N-acetylgalactosaminitol bearing the H-3 and H-5 protons.

**DISCUSSION**

The unique feature in oligosaccharide K2(4) of an unbranched GlcNAc\beta_\text{l--6Gal backbone sequence in the absence of a GlcNAc\beta_\text{l--3} substituent is analogous to the previously reported GlcNAc\beta_\text{l--6GalNAc-ol} core of meconium glycopeptides (Hounsell et al., 1985; Feeney et al., 1986), which again occurs in the absence of the normally found GlcNAc\beta_\text{l--3} substituent. The identification of these novel oligosaccharides throws open speculation as to whether the GlcNAc\beta_\text{l--6Gal/GalNAc sequence represents a new biosynthetic step and casts doubt on the generally accepted dogma that the 1,6-N-acetylgalactosaminyltransferases are branching enzymes acting on an already 3-substituted galactose or N-acetylgalactosamine residue.

Other explanations are possible and include digestion by gastrointestinal glycosidases, bacterial degradation and chemical breakdown during the isolation procedure. However, these processes are thought to be unlikely to account for the novel oligosaccharide sequences because of the relatively large amounts of these oligosaccharides found, the supposed lack of bacteria in the foetal gastrointestinal tract and the generally greater chemical stability of 1--3 linkages compared with 1--6 linkages.

We are grateful to Mrs. Sheila Kingsley for typing the manuscript.

**REFERENCES**


Hakomori, S. (1964) J. Biochem. (Tokyo) 55, 205--208


Received 1 June 1988; accepted 8 July 1988