Enzymatic Synthesis of Octadecameric Saccharides of Multiply Branched Blood Group I-Type, Carrying Four Distal α1,3-Galactose or β1,3-GlcNAc Residues†

Antti Seppo, Leena Penttilä, Ritva Niemelä, Hannu Maaheimo, and Ossi Renkonen*

Institute of Biotechnology and Department of Biochemistry, University of Helsinki, P.O. Box 45 SF-00014 University of Helsinki, Finland

Anita Keane

Finnigan MAT Ltd., Paradise, Hemel, Hempstead, Herts, HP2 4TK, U.K.

Received November 1, 1994; Revised Manuscript Received January 11, 1995[⊗]

ABSTRACT: Radiolabeled oligosaccharide constructs were prepared to evaluate carbohydrate determinants involved in gamete adhesion in mice. The octasaccharide primer GlcNAcβ1−3Galβ1−4GlcNAcβ1−3(GlcNAcβ1−3Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAc (**1**) was incubated with UDP-GlcNAc and β1,6-GlcNAc-transferase of hog gastric microsomes, producing the tetraantennary decasaccharide GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4GlcNAcβ1−3[GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**2**). The decasaccharide was then incubated with UDP-Gal and β1,4-galactosyltransferase from bovine milk, yielding the tetradecasaccharide Galβ1−4GlcNAcβ1−3(Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−3[Galβ1−4GlcNAcβ1−3(Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**3**). Incubation of the tetradecasaccharide **3** with UDP-Gal and α1,3-galactosyltransferase from bovine thymus gave the octadecameric glycan Galα1−3Galβ1−4GlcNAcβ1−3(Galα1−3Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−3[Galα1−3Galβ1−4GlcNAcβ1−3(Galα1−3Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**4**). Incubation of the tetradecasaccharide **3** with UDP-GlcNAc and β1,3-GlcNAc-transferase present in human serum gave the octadecameric saccharide GlcNAcβ1−3Galβ1−4GlcNAcβ1−3(GlcNAcβ1−3Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−3[GlcNAcβ1−3Galβ1−4GlcNAcβ1−3(GlcNAcβ1−3Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**5**). From 104 nmol of **1**, the octadecamer **4** was obtained in a yield of 53 nmol, enough for characterization by 1D proton NMR, matrix-assisted laser desorption ionization mass spectrometry, and degradative experiments. The constructs were studied as inhibitors of mouse gamete binding in experiments that are reported in detail in the accompanying paper [Litscher, E. S., *et al.* (1995) *Biochemistry 34*, 4662−4669].

Poly-*N*-acetyllactosamine glycans are presently of considerable interest because of their antigenicity (Finne *et al.*, 1989), the developmental regulation of their expression (Muramatsu, 1988), and their involvement in signaling (Velupillai & Harn, 1994) as well as cell adhesion (Hughes, 1992; Feizi, 1993). Important aspects of their biological roles are emerging from the study of saccharide-mediated gamete interactions in mouse (Litscher & Wassarman, 1993). These include primary binding of sperm to egg (Bleil & Wassarman, 1980), acrosomal reaction of sperm (Florman *et al.*, 1984), and secondary binding of acrosome-reacted sperm to egg (Bleil & Wassarman, 1986).

The primary binding of mouse gametes is mediated by large oligosaccharides bound to the egg extracellular coat glycoprotein ZP3[1] and its counter molecule, an as yet unidentified membrane component of sperm (Florman & Wassarman, 1985). The highest sperm-binding activity resided in a neutral fraction of O-glycans of 3.9 kDa, the equivalent of about 20−25 monosaccharide residues. Miller *et al.* (1992) have also found that the sperm-binding saccharides of ZP3 are 4 kDa. Hence, the sperm receptor oligosaccharides of mouse ZP3 may represent one or several poly-*N*-acetyllactosamine glycans similar to those of granulocytes (Fukuda *et al.*, 1986), epithelial mucins (Derevitskaya *et al.*, 1978; Wu *et al.*, 1984; Lamblin *et al.*, 1991), and keratan sulfate II (Tai *et al.*, 1993). This notion is supported also by the finding that murine embryonal carcinoma cells, known to be particularly rich in poly-*N*-acetyllactosamine glycans (Muramatsu *et al.*, 1978; Renkonen, 1983), produced sperm-binding ZP3 when transfected with mouse ZP3 gene (Kinloch *et al.*, 1991), whereas many other transfected cell lines did not.

The large size of the naturally occurring sperm receptor saccharide is in line with observations showing that multiantennary oligosaccharide ligands often possess particularly high affinities for the appropriate carbohydrate-binding proteins (Lee *et al.*, 1983; Sabesan *et al.*, 1992; DeFrees *et al.*, 1993). The increased affinity of branched oligosaccharides is probably caused by their ability to cross-link neighboring carbohydrate recognition domains of the lectins. Such cross-linking is directly visualized in some crystallographic studies of lectin−saccharide complexes (Weis *et*

† Grants from Jenny and Antti Wihuri foundation, from the University of Helsinki, and from the Academy of Finland are acknowledged.

* To whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1995.

[1] Abbreviations: ZP, zona pellucida; ZP3, zona pellucida glycoprotein 3; Gal, D-galactose; GlcNAc, *N*-acetyl-D-glucosamine; UDP, uridine-5′-diphospho; Lac, lactose; MT, maltotriose; MTet, maltotetraose; IMO, isomaltooctaose; TBS, Tris-buffered saline; WEFT, water elimination−Fourier transform; MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry.

0006-2960/95/0434-4655$09.00/0

al., 1992; Wright, 1992).

The poly-*N*-acetyllactosamine glycans of murine embryonal carcinoma cells contain the branched backbone sequence GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4GlcNAc (Muramatsu *et al.*, 1983; Seppo *et al.*, 1990), and both the 1→3- and the 1→6-linked branches appear to be β1,4-galactosylated (Renkonen *et al.*, 1989). The sperm-binding oligosaccharides of mouse ZP3 carry distal α-linked galactosyl residues that are essential for the biological activity (Bleil & Wassarman, 1988). The data of Bleil and Wassarman imply that distal Galα1−3Gal sequences are probably present, because Cummings and Mattox (1988) have found α1,3-galactosyltransferase activity in murine embryonal cancer cells of line F9, and we have noted the presence of Galα1−3Gal sequences, but the absence of other GalαGal sequences, in cells of line PC13 (A. Seppo *et al.*, unpublished observations). On the other hand, Miller *et al.* (1992) have provided evidence showing the importance of distal GlcNAc residues in the sperm-binding oligosaccharides of ZP3.

The present report describes enzyme-assisted synthesis of poly-*N*-acetyllactosamine glycans that contain many of the structural features believed to be characteristic of the sperm-binding saccharides on mouse eggs. However, the core portion of the O-linked glycans was omitted from the constructs, because the pyranosidic form of the proximal GalNAc was not essential for the binding activity (Florman & Wassarman, 1985).

MATERIALS AND METHODS

The primer octasaccharide GlcNAcβ1−3[^{3}H]Galβ1−4GlcNAcβ1−3(GlcNAcβ1−3[^{3}H]Galβ1−4GlcNAcβ1−6)-[^{14}C]Galβ1−4GlcNAc (**1**) was obtained as described in Vilkman *et al.* (1992). The structural reporter group signals from the ^{1}H-NMR spectrum of **1** (Figure 2A) were assigned as follows: The 2.3-Hz doublet at 5.207 ppm and the 8.4-Hz doublet at 4.742 ppm were assigned to H-1 of the reducing end GlcNAc of α and β anomers of the molecule, respectively. At the region of H-1:s of β1−3-linked GlcNAc:s (4.66−4.71 ppm), the 8.0-Hz doublet at 4.681 ppm with an intensity of 2 protons was assigned to the terminal GlcNAc:s (^{3,3}GlcNAc and ^{3,6}GlcNAc),[2] and the pattern of two 8.3-Hz doublets at 4.697/4.693 ppm (intensity ratio, 0.5:0.3) was assigned to ^{3}GlcNAc H-1:s of α and β anomers of the molecule, respectively. The pattern of two 8.0-Hz doublets at 4.621/4.613 ppm, with a total intensity of 1 proton, was assigned to the ^{6}GlcNAc H-1. At the region of Galβ1−4 H-1:s, three doublets with 1 proton intensity each were observed. The signal at 4.455 ppm (7.8 Hz) was assigned to the "branching" Gal H-1 as in GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4GlcNAc (Koenderman *et al.*, 1987). The two remaining 7.7-Hz doublets at 4.467 and 4.451 ppm were assigned to H-1:s of ^{3}Gal and ^{6}Gal, respectively. The distinction between the two branches was based on the order of resonances of monosubstituted galactoses in branched structures as in Grönberg *et al.* (1990) and Seppo *et al.* (1994). The doublet with 2 protons intensity at 4.152 ppm was assigned to H-4 of ^{3}Gal and ^{6}Gal, while the 1-proton doublet at 4.145 ppm was assigned to H-4 of the branching Gal.

Radiolabeled Galα1−3Galβ1−4GlcNAc and Galα1−3Galβ1−4GlcNAcβ1−3[Galα1−3Galβ1−4GlcNAcβ1−6]-Galβ1−4GlcNAc were synthesized as described (Seppo *et al.*, 1994).

β1,6-GlcNAc transferase (EC 2.4.1.148) reactions were carried out by incubating the acceptor saccharides with UDP-GlcNAc and hog gastric mucosal microsomes (Brockhausen *et al.*, 1983; Piller *et al.*, 1984; Seppo *et al.*, 1990). β1,3-GlcNAc transferase (EC 2.4.1.149) reactions were carried out by incubating the acceptor saccharides with UDP-GlcNAc and human serum as described (Seppo *et al.*, 1990). β1,4-Gal transferase reactions were carried out by incubating the acceptor saccharides with UDP-Gal or UDP-[^{3}H]Gal and *N*-acetyllactosamine synthase (EC 2.4.1.90) from bovine milk (Sigma) essentially as described by Brew *et al.* (1968). α1,3-Gal transferase reactions were carried out by incubating the radiolabeled acceptor saccharide with UDP-Gal and α1,3-Gal transferase purified from a Triton X-100 extract of bovine thymus according to Blanken and van den Eijnden (Blanken & van den Eijnden, 1985).

Digestions with β-galactosidase (EC 3.2.1.23) from *Diplococcus pneumoniae* (Boehringer) and from jack bean (Sigma) were carried out as described earlier (Renkonen *et al.*, 1989). Digestions with endo-β-galactosidase from *Bacteroides fragilis* were carried out as described in Renkonen *et al.* (1991).

^{1}H-NMR spectroscopy of the oligosaccharides was performed in ^2H_2O at 500 MHz in a Varian Unity 500 spectrometer. 1D spectra were recorded at 23 °C using a modification of the WEFT sequence as described in Hård *et al.* (1992) for water suppression. Chemical shifts were measured by reference to internal acetone (δ = 2.225 ppm) with an accuracy of 0.002 ppm.

Matrix-assisted laser desorption ionization mass spectroscopy (MALDI-MS) of the underivatized oligosaccharides was performed with a LASERMAT instrument (Finnigan MAT Ltd., U.K.). Operating conditions and procedures were modeled on the work of Karas *et al.* (Karas *et al.*, 1987; Karas & Hillenkamp, 1988): The sample was dissolved in 50 mM 2,5-dihydroxybenzoic acid (in acetonitrile/water, 70:30 by volume) and 1 μL of the mixture containing 10−30 pmol of oligosaccharide was applied to a standard stainless steel target. The droplet was allowed to dry in a microcrystalline form before insertion into the instrument. Oligomannose 9 ($Man_9GlcNAc_2$; M_w = 1884; source, porcine thyroglobulin) from Oxford Glycosystems, U.K., was used as an external calibrant for mass measurement accuracy purposes.

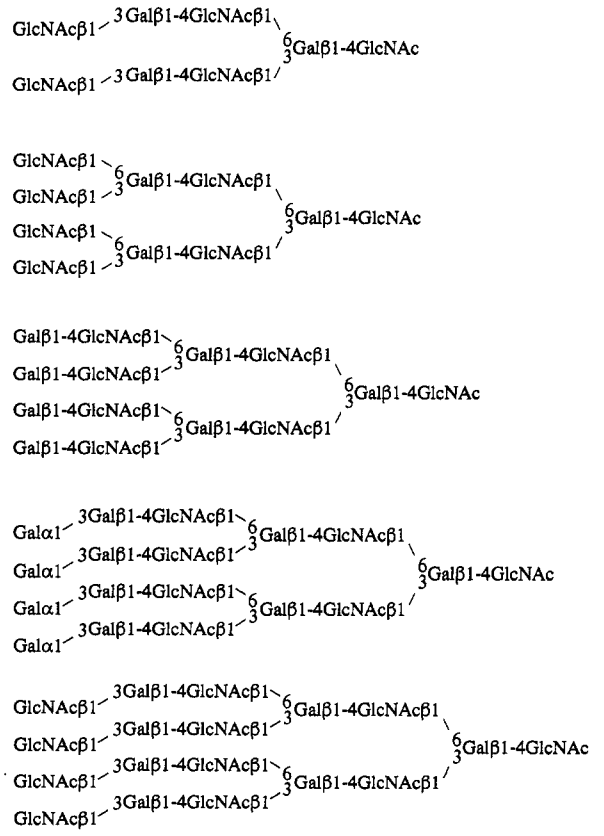
Partial acid hydrolysis was carried out using 0.1 M trifluoroacetic acid at 100 °C for 40 min (Renkonen *et al.*, 1989). Periodate oxidation of the radiolabeled galactobiose was carried out by a procedure of Hough (Hough, 1965) as described in Renkonen *et al.* (1989).

The oligosaccharides were desalted by filtration in water through AG1-X8 (AcO^-) and AG50W-X8 (H^+) (Bio-Rad, Richmond, CA). Gel filtration was performed as described in Rasilo and Renkonen (1982).

HPLC was carried out on Lichrosorb-NH_2 (Merck, Darmstadt, Germany) using acetonitrile−potassium phosphate buffer, the system I of Blanken *et al.* (Blanken *et al.*, 1985), as the solvent. Paper chromatography was carried out using the upper phase of a mixture of *n*-butanol/acetic acid/water

[2] Use of superscripts in labeling of individual monosaccharide residues in oligosaccharides is as in Kamerling and Vliegenthart (1992). Briefly, the superscript is derived from the shortest unambiguous sequence of glycoside linkage positions defining the route from the monosaccharide toward the reducing end.

Table 1: Structures of the Key Saccharides of the Present Experiments



(4:1:5 by volume) (solvent A), *n*-butanol/ethanol/water (10:1:2 by volume) (solvent E), or *n*-butanol/acetic acid/water (50:15:35 by volume) (solvent F) as solvents. Separation of oligosaccharides containing terminal α-linked galactose residues by affinity chromatography on immobilized *Griffonia simplicifolia* I isolectins was carried out according to Wang *et al.* (1988).

RESULTS

The structures of the key saccharides of the present experiments are shown in Table 1.

Synthesis and Characterization of the Decameric N-Acetyllactosaminoglycan **2**. The primer octasaccharide GlcNAcβ1−3[^{3}H]Galβ1−4GlcNAcβ1−3(GlcNAcβ1−3[^{3}H]-Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAc (**1**) was obtained by starting from GlcNAc and using a series of stepwise glycosyltransferase reactions as described in Vilkman *et al.* (1992). The 1H-NMR spectrum of **1** is shown in Figure 2A; for the assignment, see Materials and Methods. The [^{3}H]octasaccharide **1** was incubated with UDP-GlcNAc and pig gastric mucosal microsomes containing β1,6-GlcNAc transferase activity. The desalted reaction mixture was separated by HPLC in a LiChrosorb-NH_2 column (not shown) to a major product that proved to be the decasaccharide GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]Galβ1−4GlcNAcβ1−3[GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**2**) and to a minor product consisting of an intermediary mixture of nonasaccharides. The decasaccharide **2** was further purified by paper chromatography (R_{IMO} = 1.17, solvent F) (not shown).

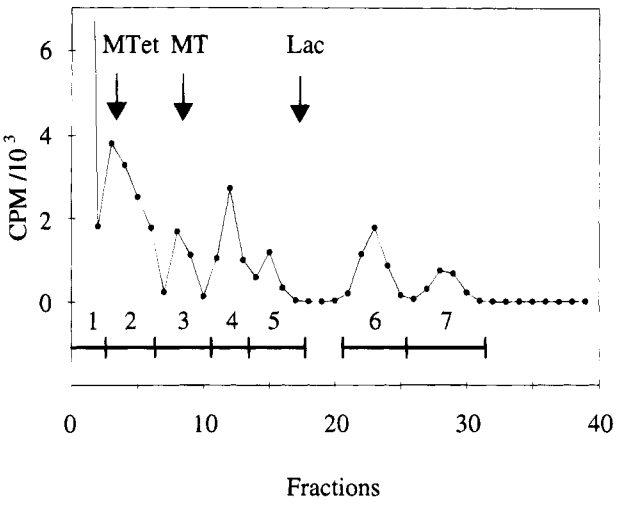


FIGURE 1: Paper chromatography of a partial acid hydrolysate of the [^{3}H]decasaccharide **2**: An 8-day run with solvent E. Peak 2 contained partially separated GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]-Galβ1−4GlcNAc and GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]Gal; peak 3 represented GlcNAc1−6[^{3}H]Galβ1−4GlcNAc; peak 4 was GlcNAcβ1−3[^{3}H]Galβ1−4GlcNAc; peak 5 contained the key disaccharide GlcNAc1−6[^{3}H]Gal; peak 6 represented GlcNAcβ1−3[^{3}H]Gal; peak 7 was [^{3}H]Galβ1−4GlcNAc. The markers run on the same paper were Lac = lactose, MT = maltotriose, and MTet = maltotetraose.

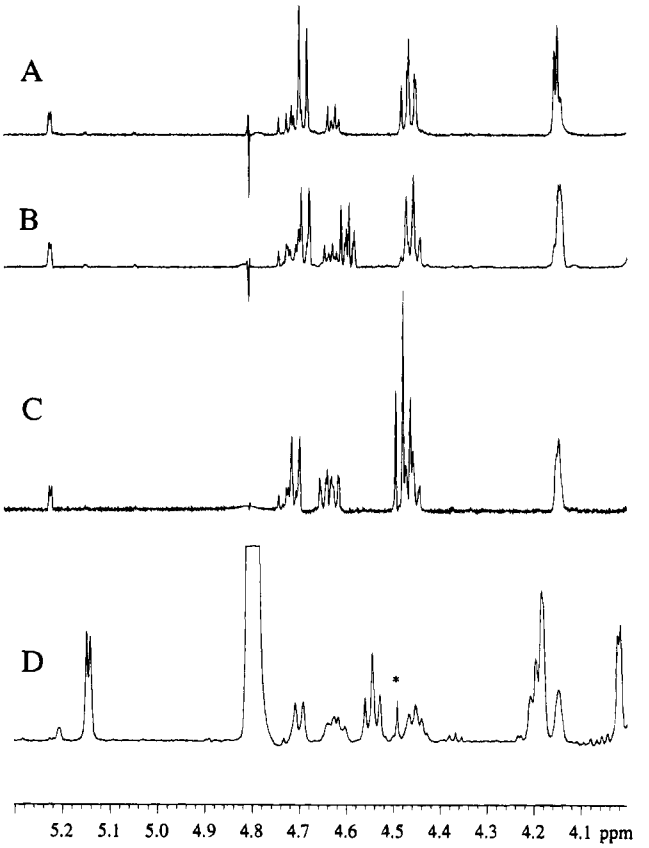


FIGURE 2: Expansions of the ^{1}H NMR spectra at 500 MHz. (A) The octamer **1**. (B) The decamer **2**. (C) The tetradecamer **3**. (D) The octadecamer **4**. The asterisk denotes a signal from an impurity which could be removed from the sample by gel filtration.

An aliquot of the decasaccharide **2** was subjected to partial acid hydrolysis, yielding a pattern of diagnostic cleavage products in paper chromatography (Figure 1; Table 2). These included GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]Galβ1−4GlcNAc, GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]Gal, GlcNAcβ1−6[^{3}H]-

Table 2: Diagnostic Cleavage Products Isolated from a Partial Acid Hydrolysate of **2**

cleavage product	mode of identification[a]
GlcNAcβ1 → 6 [^{3}H]Gal	Paper chromatography (solvents A & E) WGA-chromatography HPLC β-N-acetylhexosaminidase
GlcNAcβ1 → 6 [^{3}H]Galβ1-4GlcNAc	Paper chromatography (solvents A & E) β-N-acetylhexosaminidase
GlcNAcβ1 → 6 [^{3}H]Gal GlcNAcβ1 → 3	Paper chromatography (solvents A & E) WGA-chromatography β-N-acetylhexosaminidase
GlcNAcβ1 → 6 [^{3}H]Galβ1-4GlcNAc GlcNAcβ1 → 3	Paper chromatography (solvents A & E) WGA-chromatography β-N-acetylhexosaminidase

[a] See A. Seppo *et al.* (1990).

Galβ1−4GlcNAc, and GlcNAcβ1−6[^{3}H]Gal which establish the identity of the newly formed GlcNAcβ1−6[^{3}H]Gal linkages of the decamer **2**.

The branched sequence GlcNAcβ1−3(GlcNAcβ1−6)[^{3}H]-Galβ1−4GlcNAc is not cleaved by endo-β-galactosidase from *B. fragilis* (Scudder *et al.*, 1987). Thus, the decasaccharide **2** resisted the treatment with this enzyme under conditions where radiolabeled GlcNAcβ1−3Galβ1−4GlcNAc was completely cleaved (data not shown). The conversion of the cleavable octasaccharide **1** (Vilkman *et al.*, 1992) into the resistant **2** confirms that both branches of **1** had become β1,6-*N*-acetylglucosaminylated.

The anomeric proton region of the 1D ^{1}H NMR spectrum of the decasaccharide **2** is shown in Figure 2B. It confirms fully the postulated structure of **2**. In comparison to the spectrum of **1**, the addition of the two new GlcNAcs can be seen as GlcNAcβ1−6 H-1 signals of 1 proton intensity each, appearing at 4.593 ppm (8.5-Hz doublet) and 4.582/4.584 ppm (two 8.2-Hz doublets; α/β pair). These were assigned to ^{6,3}GlcNAc and ^{6,6}GlcNAc, respectively. The branches were distinguished by assuming that Gal substituents in 3-branches of oligo-*N*-acetyllactosaminoglycans resonate at lower or equal field when compared to similar substituents in 6-branches (Grönberg *et al.*, 1990; Seppo *et al.*, 1994). The H-1 signals of ^{6}Gal and ^{3}Gal appear shifted to higher field (4.457 and 4.442 ppm, respectively) than in the spectrum of **1**. The signals of H-4 of all Gals overlap to produce a broad signal at 4.141 ppm. The signals of GlcNAcβ1−3 H-1s as well as the reducing end GlcNAc H-1 signals remain practically the same as in **1**. The integrals suggest that, in addition to the reducing end GlcNAc, the decamer contains three β1,3-linked and three β1,6-linked GlcNAc residues together with three galactose units.

Synthesis and Characterization of the Tetradecameric N-Acetyllactosaminoglycan **3**. A ^{14}C-labeled decamer **2**, GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4[^{14}C]GlcNAcβ1−3[GlcNAcβ1−3(GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]-

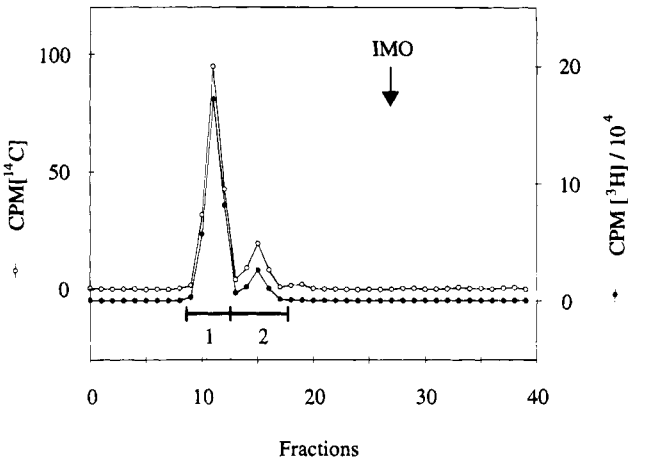


FIGURE 3: Paper chromatographic purification of the tetradecamer [^{3}H]Galβ1−4GlcNAcβ1−3([^{3}H]Galβ1−4GlcNAcβ1−6)Galβ1−4[^{14}C]GlcNAcβ1−3[[^{3}H]Galβ1−4GlcNAcβ1−3([^{3}H]Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**3**) resulting from enzymatic galactosylation of ^{14}C-labeled **2** with UDP-[^{3}H]Gal: A 5-day run with solvent F. The major peak represents saccharide **3**; the small peak 2 represents a mixture of intermediate 13-mers, carrying only three [^{3}H]Gal-units. IMO = isomaltooctaose marker. (○) ^{14}C radioactivity; (●) ^{3}H radioactivity.

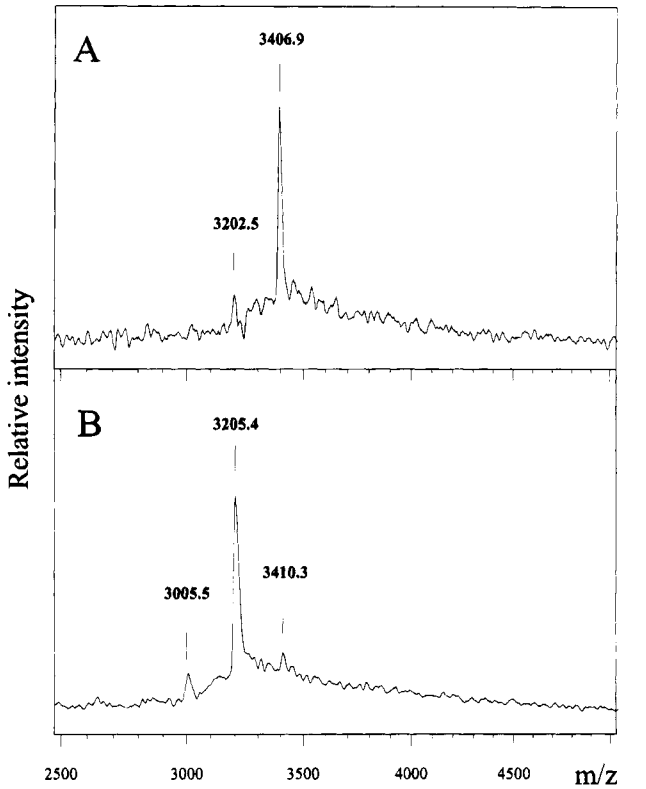


FIGURE 4: MALDI-MS of free oligosaccharides. (A) The octadecasaccharide **5**. (B) The fraction of 17-meric intermediates from synthesis of **5**.

Galβ1−4GlcNAc, was prepared as above. The product was β1,4-galactosylated with *N*-acetyllactosamine synthase and a 2-fold molar excess of UDP-[^{3}H]Gal. Paper chromatography of the reaction mixture revealed two products (Figure 3). The ^{14}C:^{3}H isotope ratios of the two products suggested that the major one contained four [^{3}H]Gal units, while the minor one contained only three. This implies that about 95% of available acceptor sites were galactosylated in the reaction and that the major product was the tetradecamer **3**.

Matrix-assisted laser desorption mass spectrometry of **3** revealed a single peak with *m/z* 2599 Da (not shown). The calculated mass for $Hex_7HexNAc_7$-Na^+ is 2598.4 Da.

All distal [^{3}H]Gal units were released from **3** upon treatment with β-galactosidase from *Diplococcus pneumoniae* (data not shown). This proved that the [^{3}H]Gal residues were β1,4-linked to the penultimate GlcNAc units (Paulson *et al.*, 1978).

The tetradecamer **3** resisted the action of endo-β-galactosidase from *B. fragilis* (not shown), implying that all internal galactoses remained substituted at position 6 (Scudder *et al.*, 1987; Renkonen *et al.*, 1991).

An expansion of the anomeric region of the ^{1}H NMR spectrum of the tetradecasaccharide **3** is shown in Figure 2C. It confirms fully the postulated structure of **3**. In comparison with the spectrum of **2**, the addition of four βGals can be seen as two 7.7-Hz doublets of an intensity of 2 protons each at 4.480 and 4.464 ppm. These were assigned to H-1 of ^{6,3}Gal + ^{6,6}Gal and H-1 of ^{3,3}Gal + ^{3,6}Gal, respectively, by comparison with the spectrum of lacto-*N*-neohexaose (Tarrago *et al.*, 1988). The signals of penultimate GlcNAcβ1−6 H-1s (4.614 ppm for ^{6,6}GlcNAc and 4.638 ppm for ^{6,3}GlcNAc) are shifted to lower field when compared to those of **2**, whereas the signal of ^{6}GlcNAc is unchanged at 4.626 ppm. Also ^{3,6}GlcNAc and ^{3,3}GlcNAc H-1s appear shifted to lower field at 4.697 ppm, in full agreement with Galβ1−4 substitution.

Synthesis and Characterization of the Octadecameric N-Acetyllactosaminoglycan **4**. The tritium-labeled tetradecamer **3** was incubated with UDP-galactose and α1,3-galactosyltransferase of bovine thymus (Blanken & van den Eijnden, 1985), and the resulting oligosaccharides were chromatographed on paper (not shown). The product appeared homogeneous and was well resolved from the acceptor. ($R_{\text{Tetradecamer 3}}$ = 0.46; solvent F). The data described below established that the product was the octadecamer Galα1−3[^{3}H]Galβ1−4GlcNAcβ1−3(Galα1−3[^{3}H]-Galβ1−4GlcNAcβ1−6)Galβ1−4[^{14}C]GlcNAcβ1−3[Galα1−3[^{3}H]Galβ1−4GlcNAcβ1−3(Galα1−3[^{3}H]Galβ1−4GlcNAcβ1−6)Galβ1−4GlcNAcβ1−6]Galβ1−4GlcNAc (**4**).

Matrix-assisted laser desorption mass spectrometry of **4** revealed a single peak (M + Na^+) with *m/z* 3245 Da (not shown). The calculated mass for $Hex_{11}HexNAc_7$-Na^+ is 3246.9 Da.

Treatment of the putative octadecamer **4** with α-galactosidase converted it completely into the tetradecamer **3** ($\mathbf{R}_{\text{Octadecamer 4}}$ = 2.26; solvent F), confirming the presence of distal α-linked galactose units. The linkage position on the subterminal [^{3}H]galactose unit of **4** was determined by subjecting the oligosaccharide to partial acid hydrolysis. This gave a series of cleavage products, including a [^{3}H]-galactobiose that was isolated by paper chromatography (not shown). Its linkage position was studied by periodate oxidation (Hough, 1965), followed by acid hydrolysis of the oxidation products. The procedure converted the [^{3}H]-galactose unit of the galactobiose into [^{3}H]lyxose (not shown), identifying the linkage position as 1 → 3.

The number of α1,3-linked galactose units in **4** was established by several experiments of different types. First, a partial α1,3-galactosylation experiment was carried out with the [^{3}H]galactose-labeled tetradecamer **3**. The paper chromatogram obtained from the resulting mixture revealed five components: the unreacted 14-mer, the completely reacted 18-mer, and, intermediary products representing putative 15-, 16-, and 17-mers (Figure 5). The reaction products were firmly identified by exhaustive digestions with jack bean β-galactosidase. This treatment removed nearly all [^{3}H]Gal residues from the 14-mer, showing that they were distally located. In contrast, from the 18-mer almost no [^{3}H]Gal residues were removed, proving that they were protected by α-galactose residues. Expected amounts of [^{3}H]galactose residues were removed by β-galactosidase also from the intermediary fractions (Figure 5).

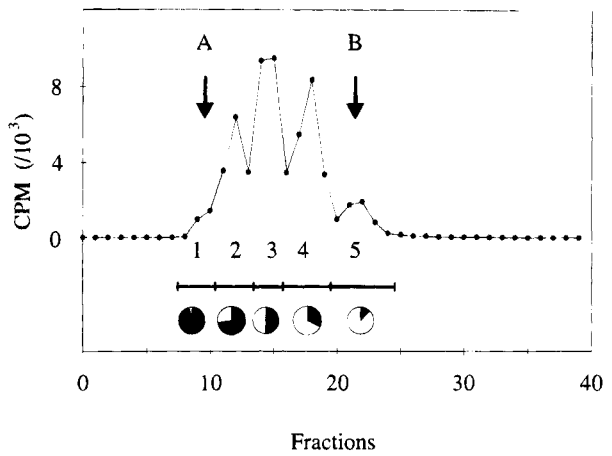


FIGURE 5: Paper chromatogram of oligosaccharides from partial α1,3-galactosylation of the [^{3}H]tetradecamer **3**: A 10-day run with solvent F. Arrows A and B denote positions of octadecamer **4** and tetradecamer **3**, respectively. The bars labeled 1−5 mark fractions pooled as tentative 18-, 17-, 16-, 15-, and 14-mers, respectively. Pie charts below each bar show the portion of β-galactosidase-released [^{3}H]Gal (white sector) vs β-galactosidase-resistant [^{3}H]-Gal (black sector).

Affinity chromatography on immobilized *Griffonia simplicifolia* isolectins I (Murphy & Goldstein, 1977; Wang *et al.*, 1988) gave peaks of positions distinct from all of the oligosaccharides containing zero, one, two, three, or four α1,3-galactose units (Table 3). Interestingly, small model compounds containing zero, one, or two α1,3-galactose units occupied almost the same positions in this system as their large counterparts. These data support the identification of the 14-mer **3** and the 18-mer **4**, as well as the intermediary 15-, 16-, and 17-mers.

An expansion of the anomeric region of the ^{1}H 1D NMR spectrum of the octadecasaccharide **4** is shown in Figure 2D. Comparison with the spectrum of the tetradecamer **3** shows the dramatic appearance of the H-1, H-4, and H-5 of the four αGal residues at 5.146 (3.6-Hz doublet), 4.021, and 4.194 ppm, respectively, and the shift of subterminal Gal H-1s to 4.553 and 4.537 ppm. These values agree well with those recorded for the anomeric protons of the αGal residue in the distal Galα1−3Gal sequence in Galα1−3Galβ1−4GlcNAc (van Halbeek *et al.*, 1983), in *N*-glycans (Debray *et al.*, 1991), and in medium-sized *N*-acetyllactosaminoglycans (Seppo *et al.*, 1994). The signals from H-4 of the β1,4-linked galactoses appear at 4.149 ppm, like in the decasaccharide **2** and the tetradecasaccharide **3**, integrating to 3 units. An additional H-4 signal of 4-units size appears at 4.185 ppm, close to the value reported for the Galα1−3Gal sequence in *N*-glycans (Debray *et al.*, 1991). The anomeric protons of the GlcNAc units in **4** resembled those of **3**. The integrals suggest that **4** consisted of four nonreducing αGal units, four 3-substituted and three 3,6-disubstituted βGal

Table 3: Affinity Chromatography[a] of the Partially α-Galactosylated Derivatives of **3**

saccharide	peak position (fraction)[b]
Pool (Fig. 5)	9
5 (no α-Gal units)	20
4 (one α-Gal unit)	37
3 (two α-Gal units)	50
2 (three α-Gal units)	62
1 (four α-Gal units)	
Model Compounds	10
GlcNAc	23
Galα1 ⁄ 3Galβ1-4GlcNAc	
	38
Galα1 ⁄ 3Galβ1-4GlcNAcβ1 \ 6 (3) Galβ1-4GlcNAc; Galα1 ⁄ 3Galβ1-4GlcNAcβ1 ⁄	

[a] The column (0.7 × 13.3 cm) of immobilized *Griffonia simplicifolia* isolectins I on ConA-agarose was prepared according to Wang *et al.* (1988). It contained 48.6 mg of *G. simplicifolia* isolectins I (Sigma). [b] Fractions of 0.63 mL were collected; tubes 1−20 were eluted in TBS (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 0.02% NaN_3), tubes 21−50 were eluted with 0.3 mM raffinose in TBS, and tubes 51−80 were eluted with 0.5 mM raffinose in TBS.

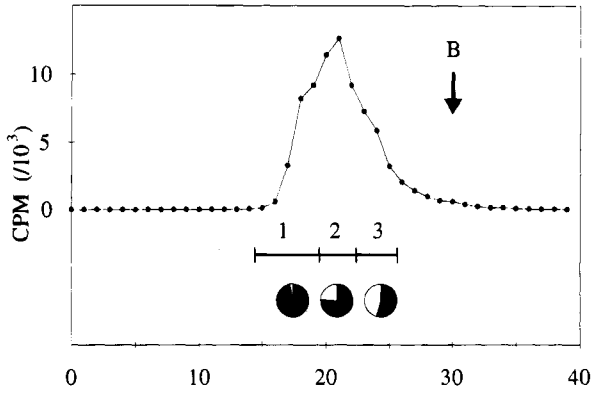


FIGURE 6: Paper chromatogram of oligosaccharides from partial β1,3-*N*-acetylglucosaminylation of the [^{3}H]tetradecamer **3**. Solvent F was used in a run of 9 days. The arrow B marks the position of the tetradecamer **3**. The bars labeled 1−3 mark fractions pooled as tentative 18-, 17-, and 16-mers, respectively. Pie charts below each bar show the portion of β-galactosidase-released [^{3}H]Gal (white sector) vs β-galactosidase-resistant [^{3}H]Gal (black sector).

units, three βGlcNAc residues in 1 → 3 linkage, and three in 1 → 6 linkage, together with one reducing end GlcNAc.

Synthesis and Characterization of the Octadecameric N-Acetyllactosaminoglycan **5**. The ^{3}H-labeled tetradecamer **3** was used also for preparation of the octadecameric saccharide **5**, carrying four distal β1,3-linked GlcNAc units. For this, **3** was incubated with UDP-GlcNAc and β1,3-*N*-acetylglucosaminyltransferase present in human serum (Piller & Cartron, 1983; Yates & Watkins, 1983; Zielenski & Koscielak, 1983), and the resulting oligosaccharides were chromatographed on paper (Figure 6). The data described below established that fractions 15−19 contained the octadecamer **5** (see Table 1 for the structure), while fractions 20−22 and 23−25 represented partially reacted heptadecameric and hexadecameric intermediates, respectively. Comparison of Figures 6 and 5 shows that the octadecamer **5** migrated faster than the octadecamer **4**; there was also a much better separation between the partially reacted intermediates in Figure 5 than in Figure 6.

Matrix-assisted laser desorption mass spectrometry of fractions 15−19 revealed a major peak ($M + Na^+$) with *m/z* 3406.9 Da; an additional small peak of 17-meric intermediates at 3206.6 Da was also observed (Figure 4A). The calculated mass for $Hex_7HexNAc_{11}$-Na^+ is 3411.2 Da. Fractions 20−22, in turn, revealed a major peak of 17-mers at 3205.4 Da (Figure 4B) with distinct small peaks of 18-mer and 16-mer at 3410.3 and 3005.5 Da, respectively.

Exhaustive β-galactosidase digestions confirmed the MALDI-MS data: As shown in Figure 6, fractions 15−19, representing the octadecamer **5**, did not release significant amounts of [^{3}H]galactose, while fractions 20−22, consisting of 17-mers, released one-quarter of their [^{3}H]galactose, and fractions 23−25, representing 16-meric intermediates, released half of their [^{3}H]galactose.

DISCUSSION

The present experiments constitute enzymatic *de novo* synthesis of a tetradecameric poly-*N*-acetyllactosamine backbone consisting of seven *N*-acetyllactosamine residues. The disaccharide units are interlinked in a highly branched tetraantennary array with 1 → 3′ and 1 → 6′ bonds, which are the principal ones also in natural oligo- and poly-*N*-acetyllactosamine sequences (structure **3** in Table 1). The tetradecamer **3** was subsequently capped with four α1,3-linked galactose residues to yield the octadecamer **4** or, in another experiment, with four β1,3-linked *N*-acetylglucosamine units to give the octadecamer **5**. The constructs **3**, **4**, and **5** represent the first pure samples of man-made poly-*N*-acetyllactosamine glycans consisting of 14−18 monosaccharide units. They were synthesized in order to assess the ability of pure saccharides of known structure as inhibitors of mouse gamete adhesion [for a recent review of oligosaccharide sperm receptors on mouse eggs, see Litscher and Wassarman (1993)].

The structures of the synthetic products of the present experiments were established by their mode of formation, by exo- and endoglycosidase treatments, and by NMR and MALDI-MS. The presence and number of α1,3-bonded galactoses in octadecamer **4** were further confirmed by chemical degradation experiments and lectin-affinity chromatography. Particularly valuable information was obtained from the present MALDI-MS experiments with the 14- and 18-meric poly-*N*-acetyllactosamine glycans. Oligosaccharides do not protonate by the technique of matrix-assisted laser desorption ionization but run as an $M + Na^+$ species (Mock *et al.*, 1991). The molecular masses observed for the sodium adducts of oligosaccharides **3**, **4**, and **5** were very close to the expected values, the deviations being only of the order of 0.1%. Thus, MALDI-MS established the numbers of hexose and *N*-acetylhexosamine residues in the oligosaccharides with great precision, despite the fact that the instrument calibration depended on an external marker. Even better accuracy can be achieved by using an internal calibrant on the same target as the sample. Another helpful method involved the synthesis of isotopic isomers of radiolabeled poly-*N*-acetyllactosamine glycans, such that they contained the radiolabel in positions that were most helpful

for the interpretation of ensuing degradation studies. So far, only 1D ^{1}H NMR experiments could be performed with the poly-*N*-acetyllactosamine glycans of large size, but the structural reporter groups (Vliegenthart *et al.*, 1983) could be assigned reliably because of the access to a complete series of synthetic intermediates beginning from the monosaccharide GlcNAc.

In conclusion, the present experiments show that it is feasible to synthesize enzymatically relatively large, structurally defined oligosaccharides consisting of interlinked *N*-acetyllactosamine residues and distally located "capping" monosaccharides. It is remarkable that in many of the successful reactions quite crude enzyme preparations were used. When pure sugar nucleotides and pure saccharide acceptors were used as substrates for exhaustive reactions, the problems generated by crude enzymes were limited to hydrolytic side reactions, and their products could be removed by chromatography. The yields of the poly-*N*-acetyllactosamine constructs in the present experiments were tens of nanomoles, large enough for structural analysis, and also for inhibition studies of mouse gamete adhesion (Litscher *et al.*, 1993) that will be described in detail in the accompanying paper (Litscher *et al.*, 1995).

REFERENCES

Blanken, W. M., & van den Eijnden, D. H. (1985) *J. Biol. Chem. 260*, 12927−12934.

Blanken, W. M., Bergh, M. L. E., Koppen, P., & van den Eijnden, D. H. (1985) *Anal. Biochem. 145*, 322−330.

Bleil, J. D., & Wassarman, P. M. (1980) *Cell 20*, 870−882.

Bleil, J. D., & Wassarman, P. M. (1986) *J. Cell Biol. 102*, 1363−1371.

Bleil, J. D., & Wassarman, P. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6778−6782.

Brew, K., Vanaman, T. C., & Hill, R. L. (1968) *Proc. Natl. Acad. Sci. U.S.A. 59*, 491−497.

Brockhausen, I., Williams, D., Matta, K. L., Orr, J., & Schachter, H. (1983) *Can. J. Biochem. Cell Biol. 61*, 1322−1333.

Cummings, R. D., & Mattox, S. A. (1988) *J. Biol. Chem. 263*, 511−519.

Debray, H., Dus, D., Wieruszeski, J.-M., Strecker, G., & Montreuil, J. (1991) *Glycoconjugate J. 8*, 29−37.

DeFrees, S. A., Gaeta, F. C. A., Lin, Y.-C., Ichikawa, Y., & Wong, C.-H. (1993) *J. Am. Chem. Soc. 115*, 7549−7550.

Derevitskaya, V. A., Arbatsky, N. P., & Kochetkov, N. K. (1978) *Eur. J. Biochem. 86*, 423−437.

Feizi, T. (1993) *Curr. Opin. Struct. Biol. 3*, 701−710.

Finne, J., Breimer, M. E., Hansson, G. C., Karlsson, K.-A., Leffer, H., Vliegenthart, J. F. G., & van Halbeek, H. (1989) *J. Biol. Chem. 264*, 5720−5735.

Florman, H. M., & Wassarman, P. M. (1985) *Cell 41*, 313−324.

Florman, H. M., Bechtol, K. B., & Wassarman, P. M. (1984) *Dev. Biol. 106*, 243−255.

Fukuda, M., Carlsson, S. R., Klock, J. C., & Dell, A. (1986) *J. Biol. Chem. 261*, 12796−12806.

Grönberg, G., Lipniunas, P., Lundgren, T., Lindh, F., & Nilsson, B. (1990) *Arch. Biochem. Biophys. 278*, 297−311.

Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J. P., & Vliegenthart, J. F. G. (1992) *Eur. J. Biochem. 209*, 895−915.

Hough, L. (1965) in *Methods in Carbohydrate Chemistry* (Whistler, R. L., BeMiller, J. N., & Wolfrom, M. L., Ed.) Vol. 5, pp 370−377, Academic Press, New York.

Hughes, R. C. (1992) *Curr. Opin. Struct. Biol. 2*, 687−692.

Kamerling, J. P., & Vliegenthart, J. F. G. (1992) *Biol. Magn. Reson. 10*, 1−287.

Karas, M., & Hillenkamp, F. (1988) *Anal. Chem. 60*, 2299−2301.

Karas, M., Bachmann, D., Bahr, U., & Hillenkamp, F. (1987) *Int. J. Mass Spectrom. Ion Proc. 78*, 53−68.

Kinloch, R. A., Mortillo, S., Stewart, C. L., & Wassarman, P. M. (1991) *J. Cell Biol. 115*, 655−664.

Koenderman, A. H. L., Koppen, P. L., & Van den Eijnden, D. H. (1987) *Eur. J. Biochem. 166*, 199−208.

Lamblin, G., Rahmoune, H., Wieruszeski, J.-M., Lhermitte, M., Strecker, G., & Roussel, P. (1991) *Biochem. J. 275*, 199−206.

Lee, Y. C., Townsend, R. R., Hardy, M. R., Lönngren, J., Arnarp, J., Haraldsson, M., & Lönn, H. (1983) *J. Biol. Chem. 258*, 199−202.

Litscher, E. S., & Wassarman, P. M. (1993) *Trends Glycosci. Glycotechnol. 5*, 369−388.

Litscher, E. S., Wassarman, P. M., Juntunen, K., Seppo, A., Niemelä, R., Penttilä, L., & Renkonen, O. (1993) *Mol. Biol. Cell 4*, 140a.

Litscher, E. S., Juntunen, K., Seppo, A., Penttilä, L., Niemelä, R., Renkonen, O., & Wassarman, P. M. (1975) *Biochemistry 34*, 4662−4669.

Miller, D. J., Macek, M. B., & Shur, B. D. (1992) *Nature 357*, 589−593.

Mock, K. K., Davey, M., & Cotterell, J. S. (1991) *Biochem. Biophys. Res. Commun. 177*, 644−651.

Muramatsu, H., Ishihara, H., Miyauchi, T., Gachelin, G., Fujisaki, T., Tejima, S., & Muramatsu, T. (1983) *J. Biochem. 94*, 799−810.

Muramatsu, T. (1988) *Biochimie 70*, 1587−1596.

Muramatsu, T., Gachelin, G., Nicolas, J. F., Condamine, H., Jakob, H., & Jacob, F. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2315−2319.

Murphy, L. A., & Goldstein, I. J. (1977) *J. Biol. Chem. 252*, 4739−4742.

Paulson, J. C., Prieels, J.-P., Glasgow, L. R., & Hill, R. L. (1978) *J. Biol. Chem. 253*, 5617−5624.

Piller, F., & Cartron, J.-P. (1983) *J. Biol. Chem. 258*, 12293−12299.

Piller, F., Cartron, J.-P., Maranduba, A., Veyrières, A., Leroy, Y., & Fournet, B. (1984) *J. Biol. Chem. 259*, 13385−13390.

Rasilo, M.-L., & Renkonen, O. (1982) *Eur. J. Biochem. 123*, 397−405.

Renkonen, O. (1983) *Biochem. Soc. Trans. 11*, 265−267.

Renkonen, O., Penttilä, L., Makkonen, A., Niemelä, R., Leppänen, A., Helin, J., & Vainio, A. (1989) *Glycoconjugate J. 6*, 129−140.

Renkonen, O., Penttilä, L., Niemelä, R., & Leppänen, A. (1991) *Glycoconjugate J. 8*, 376−380.

Sabesan, S., Duus, J., Neira, S., Domaille, P., Kelm, S., Paulson, J. C., & Bock, K. (1992) *J. Am. Chem. Soc. 114*, 8363−8375.

Scudder, P., Lawson, A. M., Hounsell, E. F., Carruthers, R. A., Childs, R. A., & Feizi, T. (1987) *Eur. J. Biochem. 168*, 585−593.

Seppo, A., Penttilä, L., Makkonen, A., Leppänen, A., Jäntti, J., Helin, J., & Renkonen, O. (1990) *Biochem. Cell Biol. 68*, 44−53.

Seppo, A., Penttilä, L., Leppänen, A., Maaheimo, H., Niemelä, R., Helin, J., Wieruszeski, J.-M., & Renkonen, O. (1994) *Glycoconjugate J. 11*, 217−225.

Tai, G. H., Huckerby, T. N., & Nieduszynski, I. A. (1993) *Biochem. J. 291*, 889−894.

Tarrago, M. T., Tucker, K. H., van Halbeek, H., & Smith, D. F. (1988) *Arch. Biochem. Biophys. 267*, 353−362.

van Halbeek, H., Vliegenthart, J. F. G., Winterwerp, H., Blanken, W. M., & van den Eijnden, D. H. (1983) *Biochem. Biophys. Res. Commun. 110*, 124−131.

Velupillai, P., & Harn, D. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 18−22.

Vilkman, A., Niemelä, R., Penttilä, L., Helin, J., Leppänen, A., Maaheimo, H., Lusa, S., & Renkonen, O. (1992) *Carbohydr. Res. 226*, 155−174.

Vliegenthart, J. F. G., Dorland, L., & Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209−374.

Wang, W.-C., Clark, G. F., Smith, D. F., & Cummings, R. D. (1988) *Anal. Biochem. 175*, 390−396.

Weis, W. I., Drickamer, K., & Hendrickson, W. A. (1992) *Nature 360*, 127−134.

Wright, C. S. (1992) *J. Biol. Chem. 267*, 14345−14352.

Wu, A. M., Kabat, E. A., Nilsson, B., Zopf, D. A., Gruezo, F. G., & Liao, J. (1984) *J. Biol. Chem. 259*, 7178−7186.

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

Zielenski, J., & Koscielak, J. (1983) *FEBS Lett. 158*, 164−168.

BI9425563