Structures of Oligosaccharides Produced by Base–Borohydride Degradation of Human Ovarian Cyst Blood Group H, Le^b and Le^a Active Glycoproteins†

Luciana Rovis,‡ Byron Anderson,§ Elvin A. Kabat,* Flavio Gruezo, and Jerry Liao

ABSTRACT: Treatment of two blood group glycoproteins, N-1 (Le^a active) and JS (H, Le^b active), with 0.05 M NaOH and 1 M $NaBH_4$ for 16 hr at 50° liberated about 95% of the carbohydrate as intact reduced oligosaccharides. The oligosaccharides were purified by Bio-Gel P-2 filtration followed by charcoal and paper chromatography. Nine oligosaccharides from N-1 and ten from JS have been isolated and their structures established by analyses, quantitative periodate oxidation, methylation, and Smith degradation. The majority of the compounds have never been described previously; they comprise a heterogeneous population shorter than the majority of the chains and therefore are either products of an incomplete or of a highly regulated biosynthetic process. The similarity of the inner part of the JS and N-1 oligosaccharide structures is striking and indicates a common pathway for the biosynthesis of small chains which in turn suggests some functional significance for the observed heterogeneity. The oligosaccharides from N-1 are simpler than those from JS in that the latter show a wide spectrum of fucosyl substitution on structures obtained in larger quantities from N-1 as expected from the genotypes of the individuals from whom the glycoproteins originated.

Many oligosaccharides and oligosaccharide fragments with blood group A, B, H, Le^a, and Le^b specificity have been isolated by partial acid or alkaline hydrolysis of blood group substances (for reviews *cf.* Watkins, 1972; Kabat, 1970, 1973). Consideration of the mechanism of alkaline borohydride degradation together with the structures of the oligosaccharide fragments isolated led to the formulation of a composite structure for the oligosaccharide portion which accounted for these blood group specificities (Lloyd *et al.*, 1966; Lloyd and Kabat, 1968a,b; Vicari and Kabat, 1970; Lundblad *et al.*, 1972) as well as for the blood group I specificity (Feizi *et al.*, 1971a,b) of a precursor glycoprotein (Morgan, 1960; Vicari and Kabat, 1969) (Figure 1).

Moreover, proposed biosynthetic and genetic schemes (Ceppellini, 1959; Watkins and Morgan, 1959; Ginsburg, 1972) show that the determinants of secreted blood group glycoproteins are built by glycosyl transferases which add specific sugar residues to the terminal nonreducing ends of a common precursor glycoprotein; these glycosyltransferases are under the control of the ABO, Hh, Lele, Sese, and an additional postulated (Lloyd and Kabat, 1968a) and newly described gene (Kobata and Ginsburg, 1969; Watkins, 1970). At a phenotypic level effects of certain of these genes may be expressed as different specificities resulting from the addition of two such sugars to adjacent residues on the precursor structure, for example Le^b specificity arises from the combined action of the H and Le genes (Ceppellini, 1959).

The proposed structure of the complex oligosaccharide moiety (Figure 1) was essentially a statistical representation constructed by assembling the various isolated oligosaccharides none of which was larger than a reduced hexasaccharide, and it was pointed out that many of the compounds isolated were products of incomplete biosynthesis or possibly of degradation in the cyst cavity (Lloyd and Kabat, 1968a,b; Vicari and Kabat, 1970). The oligosaccharides with various blood group activities represented terminal nonreducing fragments resulting from peeling from the reducing end after alkaline cleavage of the GalNAc linkage to the serine and threonine of the polypeptide backbone (Kabat *et al.*, 1965). Many of the fragments isolated were the result of heterogeneity of the original blood group substances; heterogeneity appears to be characteristic of many glycoproteins (Spiro, 1970). The heterogeneity of the blood group active glycoproteins is supported by immunological data in that small numbers of various determinants may be present in the original blood group substances; thus A or B substances may have small numbers of H determinants (Watkins, 1962), Le^a substances may show small numbers of I determinants, etc. (Feizi *et al.*, 1971a,b). Moreover, in the composite oligosaccharide two fucoses were unaccounted for and the nature of the sugar chain linked $\beta(1\rightarrow4)$ to the branched galactose was not certain (in brackets, Figure 1) (Kabat 1970, 1973).

To isolate oligosaccharides larger than those obtained previously, advantage was taken of the procedure of Iyer and Carlson (1971) who used 0.05 M NaOH and 1 M $NaBH_4$ in place of the 0.1 N NaOH and 1% $NaBH_4$ used earlier (Schiffman *et al.*, 1964; Lloyd and Kabat, 1968a; Lloyd *et al.*, 1966, 1968; Vicari and Kabat, 1970) and showed that the peeling reaction could be prevented. This was confirmed in this laboratory on a small scale (Anderson *et al.*, 1972) showing that the oligosaccharides obtained were terminated at the reducing end by *N*-acetyl-D-galactosaminitol[1] and that very

† From the Departments of Microbiology, Human Genetics and Development, and Neurology, College of Physicians and Surgeons, Columbia University and the Neurological Institute, Presbyterian Hospital, New York, New York 10032. *Received July 16, 1973.* Supported by a grant from the National Science Foundation (GB-35243X1) and a General Research Support grant from the U. S. Public Health Service. This is paper LVII in the series Immunochemical Studies on Blood Groups.

‡ Fulbright Senior Scholar.

§ Fellow of the Helen Hay Whitney and the National Cystic Fibrosis Foundations. Present address: Department of Biochemistry, Northwestern University, Medical School, Chicago, Ill. 60201.

[1] Abbreviations used are: Gal, galactose; Fuc, fucose; GlcNAc, 2-acetamido-2-deoxy-D-glucose; GalNAc, 2-acetamido-2-deoxy-D-galactose; *N*-acetylgalactosaminitol, 2-acetamido-2-deoxy-D-galactitol; Ac, mono-*O*-acetyl; Ac_2, di-*O*-acetyl; Me, mono-*O*-methyl; Me_2, di-*O*-methyl, etc.

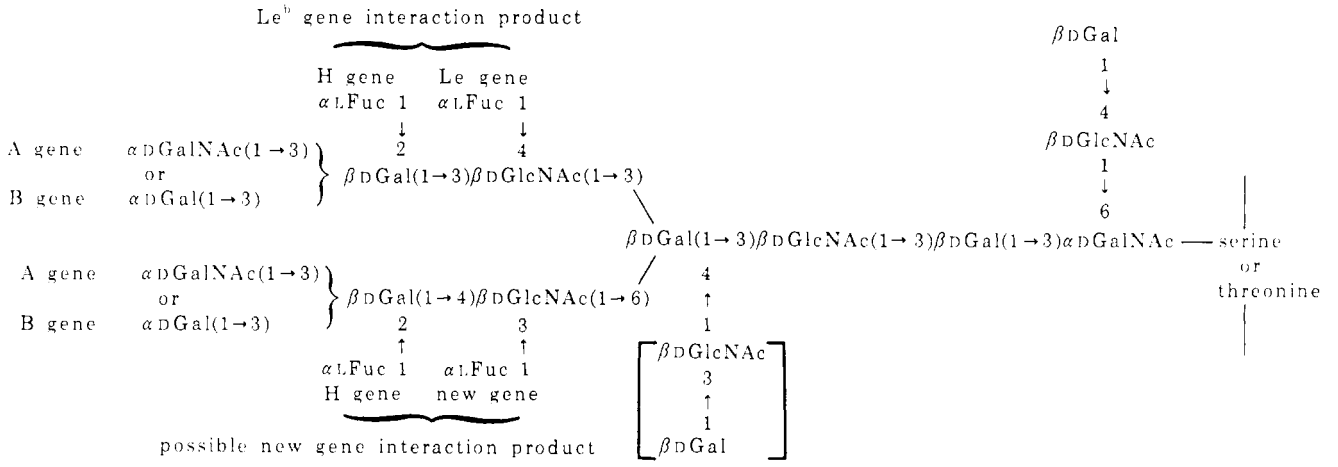

FIGURE 1: Proposed composite oligosaccharide structure showing the relation of the various blood group determinants (from Lloyd and Kabat, 1968a).

little hexenetetrol(s), hexanepentol(s), or galactitol, products of the peeling reaction, resulted.

Accordingly several grams of human ovarian cyst blood group H and Le^a substances were treated separately with 0.05 N NaOH in 1 M $NaBH_4$ and the dialyzable fragments formed were chromatographed on Bio-Gel P-2. With each substance a large excluded peak was obtained followed by a number of included peaks. Since the H substance contained several more fucosyl residues than the Le^a substance, the included peaks generally consisted of fragments somewhat larger in size. The excluded fractions were heterogeneous on paper and on charcoal and ranged in size up to about 13–15 residues, comparable in size to the proposed composite oligosaccharide of the appropriate genetic composition (Rovis *et al.*, 1973a).

The present study reports the purification, characterization, and determination of the structures of 17 oligosaccharides with a terminal *N*-acetylgalactosaminitol at the reducing end included and two excluded on Bio-Gel P-2 from the Le^a (N-1) and the H (JS) substances. They range in size from a reduced disaccharide to a reduced decasaccharide and not only provide additional evidence for the composite oligosaccharide structure (Figure 1) but indicate even greater complexity in that additional fucosyl residues are substituted on the βDGal-(1→4)-βDGlcNAc(1→6) branch linked to the GalNAc which is joined to serine and threonine. The biological activities of the oligosaccharides are reported in the following paper (Rovis *et al.*, 1973b).

Materials and Methods

Analytical Methods. Nitrogen, methylpentose (fucose), hexosamine, *N*-acetylhexosamines, and hexose (galactose) were determined by the colorimetric methods previously described (Kabat, 1961; Lloyd *et al.*, 1966). Galactosamine was estimated by the method of Ludowieg and Benmaman (1967). About 150 μg of each oligosaccharide was also analyzed for sugar components by gas–liquid chromatography (glc) after methanolysis and acetylation with Ac_2O and pyridine (Anderson *et al.*, 1972). The methanolysis procedure was modified to eliminate the neutralization with Ag_2CO_3, the HCl being removed by repeated evaporation with methanol under nitrogen; it is especially important to use methanol dried over magnesium turnings and iodine (Hughes and Clamp, 1972) and distilled. The monosaccharides are converted to fully O-acetylated methyl glycosides, and the alditols to alditol acetates. Galactose and *N*-acetylglucosamine were quantitated on the basis of molar response factors of known samples using inositol or erythritol or both as internal standards. The results, except for fucose, agreed with those obtained by colorimetric analyses and either kind of data will be given. Only colorimetric values for fucose were used because of the relative difficulty of obtaining reliable results by glc. *N*-Acetylgalactosaminitol was determined by glc but was subject to some error due to degradation by the methanolysis. The molar response factor is different for the alditol acetate derivatives obtained on methanolysis of unsubstituted, mono- or disubstituted *N*-acetyl-D-galactosaminitols. Thus it was not possible to use as a standard unsubstituted *N*-acetyl-D-galactosaminitol for the quantitation of the amino alcohol in the structures under study, and two oligosaccharides known to possess respectively a mono- and a disubstituted terminal *N*-acetylgalactosaminitol were used as standards. These compounds are Lewis R_L 1.36, βDGal(1→3)2-acetamido-2-deoxy-D-galactitol, and Lewis R_L 0.44, βDGal(1→4)βDGlcNAc(1→6)[βDGal(1→3)]-2-acetamido-2-deoxy-D-galactitol. The molar response factors found as fully acetylated methyl glycosides or as alditol acetates using these two oligosaccharides as standards relative to erythritol are: galactose in Lewis R_L 1.36, 1.07, in Lewis R_L 0.44, 1.12; standard galactose, 1.20; *N*-acetyl-D-glucosamine in Lewis R_L 0.44, 0.90; standard *N*-acetyl-D-glucosamine sample, 0.99; *N*-acetyl-D-galactosaminitol in Lewis R_L 1.36, 0.49; Lewis R_L 0.44, 0.64; standard *N*-acetyl-D-galactosaminitol, 0.37. Peaks, areas, and retention times were measured by a Hewlett-Packard 3373 B integrator.

Periodate Oxidation. Quantitative periodate consumption was measured by iodimetric titrations; formaldehyde was measured by the chromotropic acid technique and formic acid by a microtitration method as described in Lloyd *et al.* (1966).

Smith Degradation (*cf.* Goldstein *et al.*, 1965). The oligosaccharides (about 0.2–0.3 μmol) were oxidized by 0.5 ml of Rexyn 201 (200–400 mesh, Fisher Scientific Co.), freshly converted to the periodate form, for 10 min with constant shaking; the resin was filtered and washed thoroughly with H_2O. Ten milligrams of $NaBH_4$ was added and reduction allowed to take place overnight (Lee and Scocca, 1972). After desalting on Dowex 50W-X8 (H^+) and removal of borate *in vacuo* as methyl borate, the product was hydrolyzed with 0.5 N HCl for

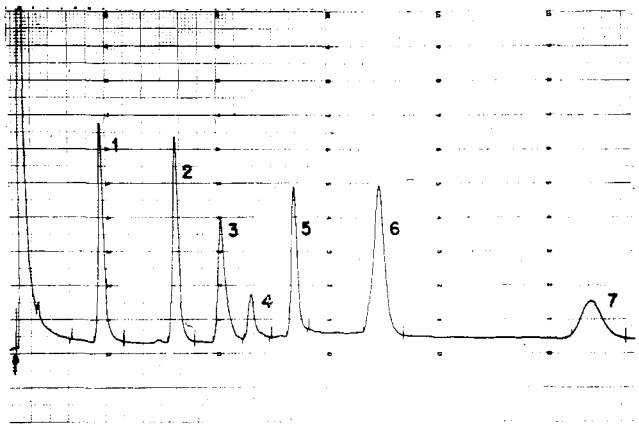

FIGURE 2. Separation of a standard mixture of polyalcohols and monosaccharides by gas–liquid chromatography on ECNSSM column. Peaks are identified as follows: erythritol (internal standard), 1; *N*-acetylserinol, 2; methyl galactoside, 3, 4; *N*-acetylthreosaminitol, 5; *N*-acetylarabinosaminitol and methyl *N*-acetylglucosaminide, 6; *N*-acetylgalactosaminitol, 7. Temperature gradient: hold at 150° for 5 min, then 2° per min to 210°, and hold at this temperature until all peaks were eluted.

24 hr at room temperature and neutralized with Dowex 1-X8 (OH^-). One part of the sample was acetylated and examined by glc for free alditol acetates. A second portion was methanolyzed and acetylated and in addition to free and bound alditol acetates resulting from periodate oxidation of the bound *N*-acetylgalactosaminitol, the surviving sugar residues were detected by glc as methyl glycosides. When the first Smith degradation did not completely destroy the structure and if sufficient oligosaccharide was available, a second Smith degradation was performed in the same manner. A sample of *N*-acetylgalactosaminitol obtained by reduction with $NaBH_4$ of GalNAc was oxidized with the insoluble periodate resin, reduced, and acetylated. The amounts of periodate resin were chosen so as to provide incomplete oxidation to give a mixture of the three possible oxidation products, *N*-acetylserinol, *N*-acetylthreosaminitol, and *N*-acetylarabinosaminitol, for use as a reference. The yield of *N*-acetylarabinosaminitol obtained by this procedure was very low; however, a standard sample (from D. Horton) was also available. βDGal(1→3)DGlcNAc (from A. Gauhe) was reduced with $NaBH_4$; the reduced products and a sample of βDGlc(1→3)-*N*-acetylgalactosaminitol (from D. M. Carlson) were carried through the same procedure and gave *N*-acetylthreosaminitol. A commercial sample of *N*-acetylserinol was also available. Figure 2 shows the separation of a mixture of these polyalcohols and of monosaccharides by glc on ECNSS-M on a Gas Chrom Q column.

As shown, GlcNAc methyl glycosides and *N*-acetylarabinosaminitol eluted in the same peak (peak 6, Figure 2). *N*-Acetylarabinosaminitol would only result from the oxidation of *N*-acetylgalactosaminitol substituted on C-4 or C-3 and C-4. In all of the oligosaccharides studied a major peak corresponding to *N*-acetylthreosaminitol was detected. This amino alcohol can only result from oxidation of either a monosubstituted (on C-3) or a disubstituted (on C-3 and C-6) *N*-acetylgalactosaminitol, and its detection excludes the coexistence, in the same compound, of a substitution on C-4 of the terminal *N*-acetylgalactosaminitol. However, the impossibility of resolving the two peaks does not allow one to exclude the presence, although unlikely, of contaminants containing *N*-acetylgalactosaminitol monosubstituted at C-4 or disubstituted at C-4 and C-3 when one or more periodate-resistant GlcNAc residues are present.

Methylation. The methylation procedure using BaO as catalyst (Kuhn *et al.*, 1958; Etzler *et al.*, 1970) was carried out on less than 1 mg of oligosaccharide. The standard reference compounds described (Anderson *et al.*, 1971) have been used. In addition, 6-*O*-methyl-*N*-methyl-D-glucosamine (from P. A. J. Gorin), 2,3-di-*O*-methyl-D-galactose, 2,4-di-*O*-methyl-D-galactose, and 2-*O*-methyl-D-galactose (from J. K. N. Jones), and 6-*O*-methyl-D-galactose, 3-*O*-methyl-D-galactose, and 4-*O*-methyl-D-galactose (from C. E. Ballou) were used as standards after conversion to their methyl glycosides and acetylation of the free hydroxyls. Lactose (Eastman), lacto-*N*-fucopentaose II, fucosidolactose, and lacto-*N*-tetraose (from R. Kuhn and A. Gauhe) were also methylated and used as reference compounds. The derivatives were identified by gas–liquid chromatography using an F&M (Hewlett-Packard) 810 gas chromatograph with ECNSS-M on Gas Chrom Q using a glass column at 120 and 175°, and an F&M (Hewlett-Packard) 700 gas chromatograph equipped with a metal column filled with NPGS on Chromosorb W at 145 and 190°.

It has already been pointed out that the BaO methylation of oligosaccharides containing GlcNAc gives rise to a variable amount of N-methylation (Anderson *et al.*, 1971). Figure 3 (A–C) shows the characteristic peak patterns due to the methyl glycosides of O-acetylated standard samples of 4,6-Me_2 (A), 3,6-Me_2 (B), and 6-Me (C) *N*-acetyl- and *N*-methylglucosamines. The separations shown are obtained on the NPGS column at 190°; the chromatograms are superimpositions of the peaks obtained separately from the *N*-methyl and *N*-acetyl standard samples. The retention times (T_r) given are relative to Me-3,4,6-Me_3GlcNAc used as reference. The actual chromatograms of oligosaccharides containing GlcNAc residues linked at C-3, at C-4, and at both C-3 and C-4 are also shown in Figure 3 (D–F). The pattern of peaks relative to a C-3-linked internal GlcNAc obtained by methylation, methanolysis, and acetylation of lacto-*N*-tetraose is in Figure 3D. The main peak (T_r 1.17) corresponds to the *N*-acetyl derivative, and a relatively small amount of the methyl *N*-methylglycoside is detected (peaks with T_r 1.6 and 1.9).

Figure 3E is the chromatogram of OG R_L 0.44 (Vicari and Kabat, 1970). This is a reduced tetrasaccharide identical with Lewis R_L 0.44 described later in this paper (see Figure 5 for structure). It had already been studied chromatographically (Anderson *et al.*, 1971) and it is used as reference for the C-3–C-6 disubstituted *N*-acetylgalactosaminitol (peak with T_r 0.77) since another standard is not available. This compound also shows very nicely the six peaks for a C-4 linked GlcNAc. The pattern also shows the methyl Me_4Gal (T_r 0.40). However, permethylated Gal and Fuc methyl glycosides are routinely identified also at lower temperatures (on ECNSS-M at 120° and on NPGS at 145°) as in Anderson *et al.* (1971).

Figure 3F shows the chromatogram of JS R_{IM5} 0.91 (*cf.* Figure 5 for the proposed structure). This compound contains a GlcNAc residue linked at both C-3 and C-4. In addition to the peak with T_r 0.77, due to the disubstituted *N*-acetylgalactosaminitol, three major peaks are seen with T_r 2.14, 2.6, and 3.0. They indicate that the GlcNAc residue had undergone almost complete N-methylation; only a trace of the *N*-acetyl derivative could be seen in the shoulder (T_r 2.41) of the large peak with T_r 2.6.

The amount of N-methylation and N-acetylation occurring is somewhat variable under the same conditions of methyla-

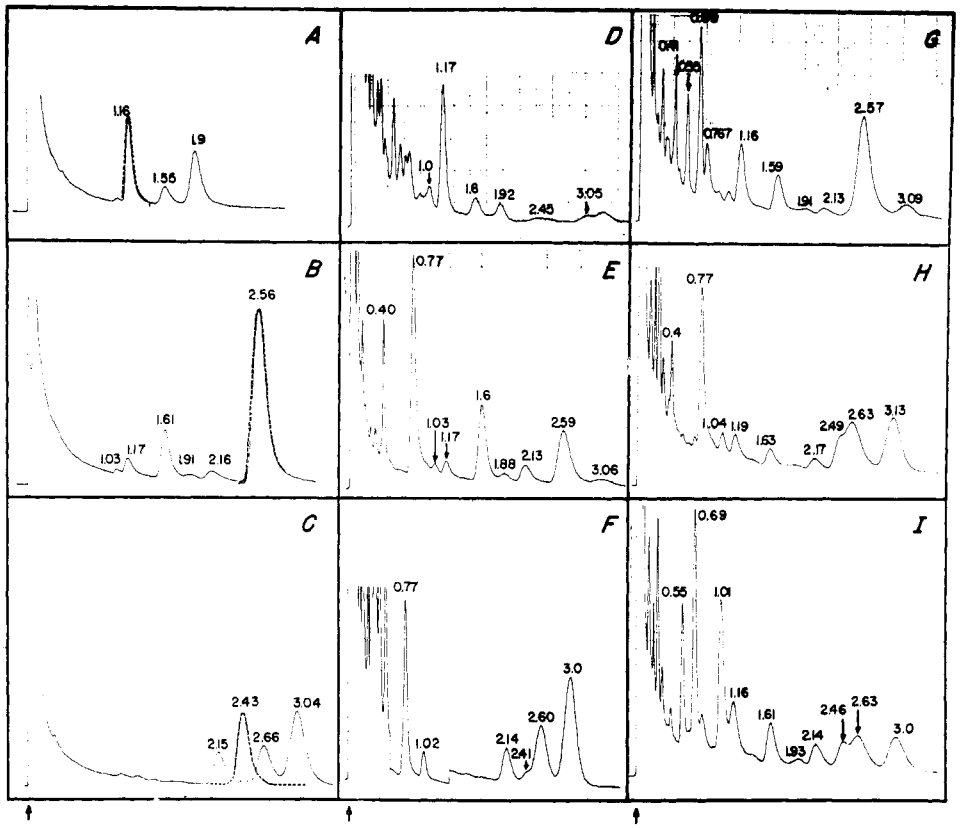

FIGURE 3: Gas chromatographic patterns of methylated and acetylated methyl glycosides of Gal and GlcNAc and of methylated acetylated *N*-acetylgalactosaminitols. All patterns are photographs of actual chromatograms unretouched except for the dotted peaks in A, B, and C which represent superimpositions of the standard *N*-acetamido peaks obtained separately, on the *N*-methyl derivatives. Numbers above peaks denote retention times relative to methyl 3,4,6-Me_3GlcNAc on an NPGS column at 190°. (A) Methyl 3Ac-4,6-Me_2GlcNAc (T_r 1.16) and NMe (T_r 1.55, 1.9); (B) methyl 4-Ac-3,6-Me_2GlcNAc (T_r 2.56) and NMe (T_r 1.03, 1.17, 1.61, 1.91, 2.16); (C) methyl 3,4-Ac_2-6-MeGlcNAc (T_r 2.43) and NMe (T_r 2.15, 2.66, 3.04); methylated, methanolyzed, and acetylated oligosaccharides, (D) lacto-*N*-tetraose; (E) OG R_L 0.44 (Vicari and Kabat, 1970); (F) JS R_{IM5} 0.91; (G) Lewis R_{IM8} 1.95; (H) Lewis R_{IM5} 0.47; (I) Lewis R_{IM8} 3.5. See text for details.

tion. However, we have noted that usually a C-3 linked internal GlcNAc will undergo almost exclusively N-acetylation, a C-4 linked GlcNAc will give approximately equal amounts of both methyl *N*-acetyl-and *N*-methylglycosides, and a residue linked at both C-4 and C-3 will be almost entirely N-methylated.

The last three chromatograms shown in Figure 3G–I were obtained from oligosaccharides isolated from N-1 glycoprotein which are described later in detail. In these compounds (see proposed structures in Figure 5), as with many of the other oligosaccharides studied, several differently substituted GlcNAc residues are present. Thus, after methylation complex chromatograms are obtained; the characteristic presence of several peaks for each type of GlcNAc residue (*N*-methyl and *N*-acetyl) enables identification to be made with greater confidence.

In Lewis T_{IM8} 1.95 (Figure 3G) both C-3 and C-4 linked GlcNAc are identified. The peak with T_r 2.57 is characteristic of the 4-linked GlcNAc (*cf.* Figure 3G with 3B and 3E) and is the major peak. However, in a 4-linked GlcNAc, the peak of T_r 1.17 is relatively minor (Figures 3B and 3E) while in a 3-linked GlcNAc it is the major peak (Figures 3A and D). Thus the pattern in Figure 3G indicates the presence of both and the approximate ratio of peak with T_r 2.57/1.17 is about 2/1. The other peaks occur both in the 3-linked and 4-linked GlcNAc. In the same chromatogram (Figure 3G) the peaks with T_r 0.55 and 0.69 are the two anomers of Me-3,6-Ac_2-2,4-Me_2Gal (standard values, not shown, 0.55 and 0.69). These two peaks are also detected in the chromatogram of Lewis R_{IM8} 3.5 (Figure 3I), which is also an oligosaccharide containing a branched Gal. The peak with T_r 0.767 is the disubstituted *N*-acetylgalactosaminitol (Figures 3D, 3E, 3H). The minor peak with T_r 3.09, although not present in the blank, is always detected after methylation of the oligosaccharides as well as in OG R_L 0.44 (Figure 3E) and in lacto-*N*-tetraose and may either be a product of some side reaction or of impurities. It corresponds in T_r to the major peak of the pattern due to the *N*-methyl derivative of a C-3 and C-4 linked GlcNAc (Figure 3C). However, when this residue is actually present, the area of this peak increases strikingly, as seen in Figures 3H and 3I. The two chromatograms shown in these last two figures are for Lewis R_{IM5} 0.47 and Lewis R_{IM8} 3.5, both of which contain C-4 as well as C-3 and C-4 linked GlcNAc residues (*cf.* Figure 5 for structures). However, their chromatograms differ by the presence, in Lewis R_{IM5} 0.47, of the peak (T_r 0.77) of the disubstituted *N*-acetylgalactosaminitol (Figure 3H), while in Lewis R_{IM8} 3.5 (Figure 3I) the peak of T_r 1.01 is characteristic of *N*-acetylgalactosaminitol substituted only on C-3; the T_r for the standard (not shown) is 1.01. Moreover, R_{IM5} 0.47 does not contain a branched Gal, and accordingly it lacks the the two peaks of T_r 0.55 and 0.69 seen in Lewis R_{IM8} 3.5 (Figure 3I) as well as in Lewis R_{IM8} 1.95 (Figure 3G). The peak with T_r 1.16, detected in Lewis R_{IM8} 3.5 (Figure 3I), could belong to the *N*-methyl derivative of a C-4 linked GlcNAc, as well as to the *N*-acetyl derivative of a C-3 linked GlcNAc (Figures 3A and 3B). It comes very close to the peak with T_r 1.01, so that it is difficult to calculate its actual area. Chromatograms obtained on the same column at lower tempera-

TABLE I: Compounds Identified from Bio-Gel P-2 Fractions of Dialyzable Products of 0.05 M NaOH–1 M $NaBH_4$ Degradation[a] of JS and N-1.

Bio-Gel P-2 Fractions	Yields[b] mg	Yields[b] %[c]	Compounds (Yields (mg))
N-1			
A	490	16.8	Lewis R_{IM8} 0.78 (4.4), Lewis R_{IM8} 1.28 (6.2)
B	315	10.8	Lewis R_{IM5} 0.47 (14.5), Lewis R_{IM8} 1.95 (21), Lewis R_{IM8} 3.5 (4.0)
C	335	11.5	Lewis R_L 0.44 (44), Lewis R_{IM5} 1.30 (5.3), Lewis R_L 0.86 (10.8)
D	145	5	Lewis R_L 1.36 (32)
E	67	2.3	Lewis R_L 1.36 (6.7), *N*-acetyl-D-galactosaminitol (22.7)
F	50	1.7	
G[d]			
JS			
A	661	34.5	
B	491	25.6	JS R_{IM5} 1.80 (2.7), JS R_{IM5} 1.04 (10), JS R_{IM5} 2.35a (2.8), JS R_{IM5} 2.35b (3.6), JS R_{IM5} 1.90 (21) JS R_{IM5} 0.91 (25), JS R_{IM5} 0.92 (23), JS R_{IM5} 1.84 (24)
C	140	5	JS R_L 1.32 (2.9), JS R_L 1.34 (23.7)
D	70	3.6	
E[d]			

[a] These dialyzable products are those described by Rovis *et al.* (1973a). [b] Salt content has been subtracted. [c] Per cent of the total carbohydrate contents of JS and N-1 glycoproteins; 44.4 and 31% of the total carbohydrate was found in nondialyzable fractions of N-1 and JS, respectively. [d] These fractions did not contain carbohydrate and they were the main salt peaks.

ture and on the ECNSS-M column at 175° did not separate the two peaks, so that it cannot be precisely identified.

Methods for charcoal–Celite column chromatography and for filtration on Bio-Gels have been described (Lloyd *et al.*, 1966). Bio-Gel P-2 and Bio-Gel P-4 (minus 400 mesh, 66 × 1.9 cm) columns were calibrated with dextran NRC No. 3 (mol wt 51,000), lactose, a mixture of oligosaccharides of the isomaltose series, generally isomaltose (IM2), isomaltotetraose (IM4), -hexaose (IM6), and -octaose (IM8), and NaCl (Rovis *et al.*, 1973a) and used as the final step of purification for each oligosaccharide.

For paper chromatography the following solvent systems were used: 1-butanol–pyridine–water (35:39:26), solvent 1; 1-butanol–pyridine–water (6:4:3), solvent 2. Galactose (Gal), lactose (L), isomaltopentaose (IM5), and isomaltooctaose (IM8) were reference compounds. Values listed as R_{Gal}, R_L, or R_{IM5} were obtained in solvent 2 while those described as R_{IM8} generally were obtained in solvent 1 or solvent 2. Tests for unsaturation were carried out based on the ability of the oligosaccharides with terminal hexenetetrol(s) to decolorize $KMnO_4$ and Br_2 as described in Lloyd *et al.* (1966). Optical rotations were determined in a Perkin-Elmer polarimeter (Model 141) in water at five different wavelengths.

Experimental Section and Results

Alkaline borohydride degradations of 2.56 g of JS (a blood group H and Le^b active glycoprotein) and of 4.56 g of N-1 (a blood group Le^a active glycoprotein) were carried out in 0.05 N NaOH and 1 M $NaBH_4$ at 50°. Analytical data on the dialyzable and nondialyzable fractions obtained, as well as the fractionation of the dialyzable materials on Bio-Gel P-2, have already been described in detail (Rovis *et al.*, 1973a). The Bio-Gel P-2 filtrate was pooled in seven fractions called A, B, C, D, E, F, and G for N-1, and in five fractions (A, B, C, D, and E) for JS. The pools were complex mixtures and were further fractionated by charcoal–Celite column chromatography and eluted by an ethanol gradient. In general each of the charcoal–Celite fractions containing carbohydrate was examined by analytical paper chromatography in solvents 1 and 2 for various lengths of time. Whenever several components moving on paper were detected, the entire charcoal fraction was subjected to fractionation on paper chromatography in the solvent which analytically gave the best resolution. In some instances, paper chromatographic fractions with similar retention times obtained from adjacent eluates from the charcoal column were pooled and rerun. The main subfractions from paper chromatography were then passed through either a Bio-Gel P-2 or a P-4 column. Usually at this stage the products were pure and eluted in a sharp peak; sometimes, however, they were retested on analytical paper. A large number of minor fractions were obtained and although purification of some of these was tried, the final yields did not allow a precise examination of the material. Table I lists the compounds isolated from the different Bio-Gel P-2 fractions of JS and N-1 with their yields.

No pure oligosaccharides were obtained from fractions A, D, and E of JS. The material present in A was of very large size and not amenable to further fractionation; in D a compound with the chromatographic characteristics of *N*-acetylgalactosaminitol was present, but it was accidentally lost. Fraction E contained mostly salt. From each of the Bio-Gel P-2 fractions of N-1 some purified components have been isolated, except for fractions F and G; however, very little carbohydrate was originally present in F, and the peaks corresponded to the regions at which salt eluted. The fractions which yield adequate amounts of purified oligosaccharides were processed as follows.

Fraction A of N-1 (490 mg). Charcoal–Celite chromatography of this material has already been described (Rovis *et al.*, 1973a). The only fraction which seemed amenable to further purification eluted between 16 and 21% of ethanol (77 mg). Preparative paper chromatography in solvent 1 for 5

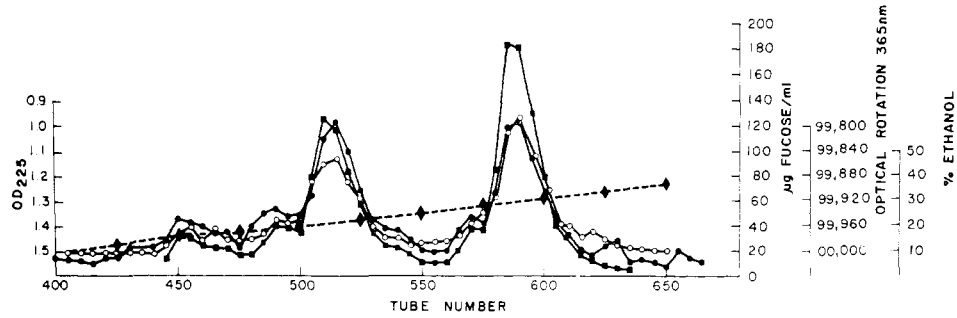

FIGURE 4: Fractionation on a charcoal–Celite column of JS fraction B (354 mg) from Bio-Gel P-2: (●) periodate consumption; (■) fucose; (○) optical rotation; (◆) ethanol concentration.

days gave 13 mg of Lewis R_{IM8} 1.28 and 12.9 mg of Lewis R_{IM8} 0.78. Additional purification through Bio-Gel P-2 gave 6.2 and 4.4 mg of the two compounds, respectively.

Fraction B of N-1 (315 mg). Charcoal chromatography showed two adjacent major peaks; the first containing fucose (57 mg) eluted at 20% ethanol and the second lacking fucose (107 mg) eluted at 23% ethanol. The two peaks were rerun separately on charcoal. Paper chromatography in solvent 2 of the fucose-containing material gave: (i) Lewis R_{IM5} 0.47 (19 mg); this was further purified on Bio-Gel P-4 (14.5 mg); (ii) Lewis R_{IM5} 0.5 (13 mg) was rechromatographed in solvent 2 and gave 6.7 mg of Lewis R_{IM8} 3.5 which was purified through Bio-Gel P-4 (4.0 mg). The non-fucose-containing peak (107 mg) was fractionated on paper in solvent 2 and showed a main subfraction (33 mg) with R_{IM5} 1.28. Rechromatography in solvent 2 and Bio-Gel P-4 filtration of the main component gave 21 mg of R_{IM8} 1.95.

Fraction C of N-1 (335 mg). Two major peaks were separated by charcoal chromatography. The first (39 mg) eluted at 16–18% ethanol and contained a small amount of fucose. It was rechromatographed on charcoal and resolved into two additional peaks, one with and one without fucose. Preparative paper chromatography in solvent 2 of both the fractions and subsequent Bio-Gel P-2 filtration gave 10.8 mg of Lewis R_L 0.86 and 5.3 mg of Lewis R_{IM5} 1.30.

The second main charcoal peak (169 mg) did not contain fucose and eluted at 21% ethanol. Rechromatography on charcoal (87 mg), filtration on Bio-Gel P-2 (68 mg), and paper chromatography in solvent 2 gave 51 mg of Lewis R_L 0.44. Refractionation on Bio-Gel P-2 gave 43.8 mg.

Fraction D of N-1 (145 mg) showed a single major peak on charcoal which eluted at 10% ethanol (74 mg). Preparative paper chromatography in solvent 2 and Bio-Gel P-2 filtration gave 32 mg of Lewis R_L 1.36.

Fraction E of N-1 (67 mg) was fractionated into two major peaks on charcoal, at 6% (26 mg) and at 10% ethanol (29 mg). Paper chromatography of the first peak in solvent 2 and Bio-Gel P-2 filtration gave a crystalline compound, R_{Gal} 1.31 (22.7 mg). The melting point (174–175°) of this compound mixed with *N*-acetyl-D-galactosaminitol showed no depression; it was also identified as *N*-acetyl-D-galactosaminitol by glc of the alditol acetate. The second charcoal peak, after paper chromatography (solvent 2) and Bio-Gel P-2 filtration, gave a compound (6.7 mg) identical with Lewis R_L 1.36 isolated from fraction D of N-1.

Fraction B of JS (491 mg) was rerun on Bio-Gel P-2 and the major fraction obtained (354 mg) was separated into several peaks on charcoal by elution with ethanol. The elution curve is shown in Figure 4; only that portion of the chromatogram with peaks is shown. From the pooled charcoal fractions several compounds were isolated by an identical preparative procedure involving paper chromatography in solvent 2 and subsequent Bio-Gel P-4 filtration. Charcoal fractions pooled and oligosaccharides purified and identified from each by paper chromatography are as shown in Table II. The oligosaccharides were passed through Bio-Gel P-4 and each eluted as a single sharp peak by three criteria: periodate uptake, methylpentose, and optical rotation, with the exception of JS R_{IM5} 2.35. This was resolved on Bio-Gel P-4 into two adjacent peaks; the first, fucose containing, is designated R_{IM5} 2.35a, and the second, without fucose, R_{IM5} 2.35b. The final yields from Bio-Gel P-4 of all the oligosaccharides are given in Table I.

Fraction C of JS (140 mg) gave several small peaks on charcoal chromatography and a major peak eluted at 27% ethanol (39.4 mg). From the fractionation of a pool of the small peaks (58 mg) on paper, many subfractions in amounts too small to be studied and 7.8 mg of JS R_L 1.32 were obtained; after Bio-Gel P-2 it gave 2.9 mg. From the 39.4 mg of material eluted at 27% ethanol, 31 mg of JS R_L 1.34 was isolated by paper chromatography in solvent 2. Purification on Bio-Gel P-2 gave 23.7 mg.

Determination of Structures of Oligosaccharides. Analytical properties, specific optical rotations at two wavelengths, and results of quantitative periodate oxidations are given in Table III. For the periodate oxidation assays, the values of periodate consumed and formaldehyde and formic acid released either after 48 or 80 hr of incubation are given in Table II, corresponding to the final values; however, for each compound a complete oxidation curve was always obtained carrying out determinations at 2, 8, 24, 48, and 80 hr (Lloyd *et al.*, 1966).

All oligosaccharides were tested for their ability to decolorize $KMnO_4$ and Br_2 with negative results, indicating that none contained an unsaturated terminal residue.

Lewis R_{IM8} 0.78 and Lewis R_{IM8} 1.28. These two oligosac-

TABLE II

Charcoal Fractions (Tube Numbers)	Oligosaccharides Isolated
444–475 (24 mg)	JS R_{IM5} 1.80 (4.7 mg)
476–500 (27 mg)	JS R_{IM5} 1.04 (9.4 mg)
	JS R_{IM5} 2.35 (8.5 mg)
501–530 (73 mg)	JS R_{IM5} 0.91 (27 mg)
	JS R_{IM5} 1.90 (24 mg)
576–605 (71 mg)	JS R_{IM5} 0.92 (28 mg)
	JS R_{IM5} 1.84 (30 mg)

N-1 Lewis[a] (4.56 g)

βDGal(1→3)-*N*-acetylgalactosaminitol

Lewis R_L 1.36 38.7 mg

βDGlcNAc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

Lewis R_L 0.86 10.8 mg

βDGal
1
↓
4
βDGlcNAc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

Lewis R_L 0.44 44 mg

βDGal
1
↓
3
βDGlcNAc(4←1)αLFuc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

Lewis R_{IM5} 1.30 5.3 mg

βDGal
1
↓
3
βDGlcNAc(4←1)αLFuc
1
↓
3
βDGal
1
↓
4
βDGlcNAc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

Lewis R_{IM5} 0.47 14.5 mg

JS H, Lewis[b] (2.56 g)

βDGal(1→3)-*N*-acetylgalactosaminitol

JS R_L 1.32 2.9 mg

βDGal
1
↓
4
βDGlcNAc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

JS R_{IM5} 2.35b 2.8 mg

βDGal
1
↓
4
βDGlcNAc(3←1)αLFuc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

JS R_{IM5} 1.80 2.7 mg

βDGal(2←1)αLFuc
1
↓
4
βDGlcNAc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

JS R_{IM5} 1.90 21mg

βDGal(2←1)αLFuc
1
↓
4
βDGlcNAc(3←1)αLFuc
1
↓
6

βDGal(1→3)-*N*-acetylgalactosaminitol

JS R_{IM5} 0.91 25mg

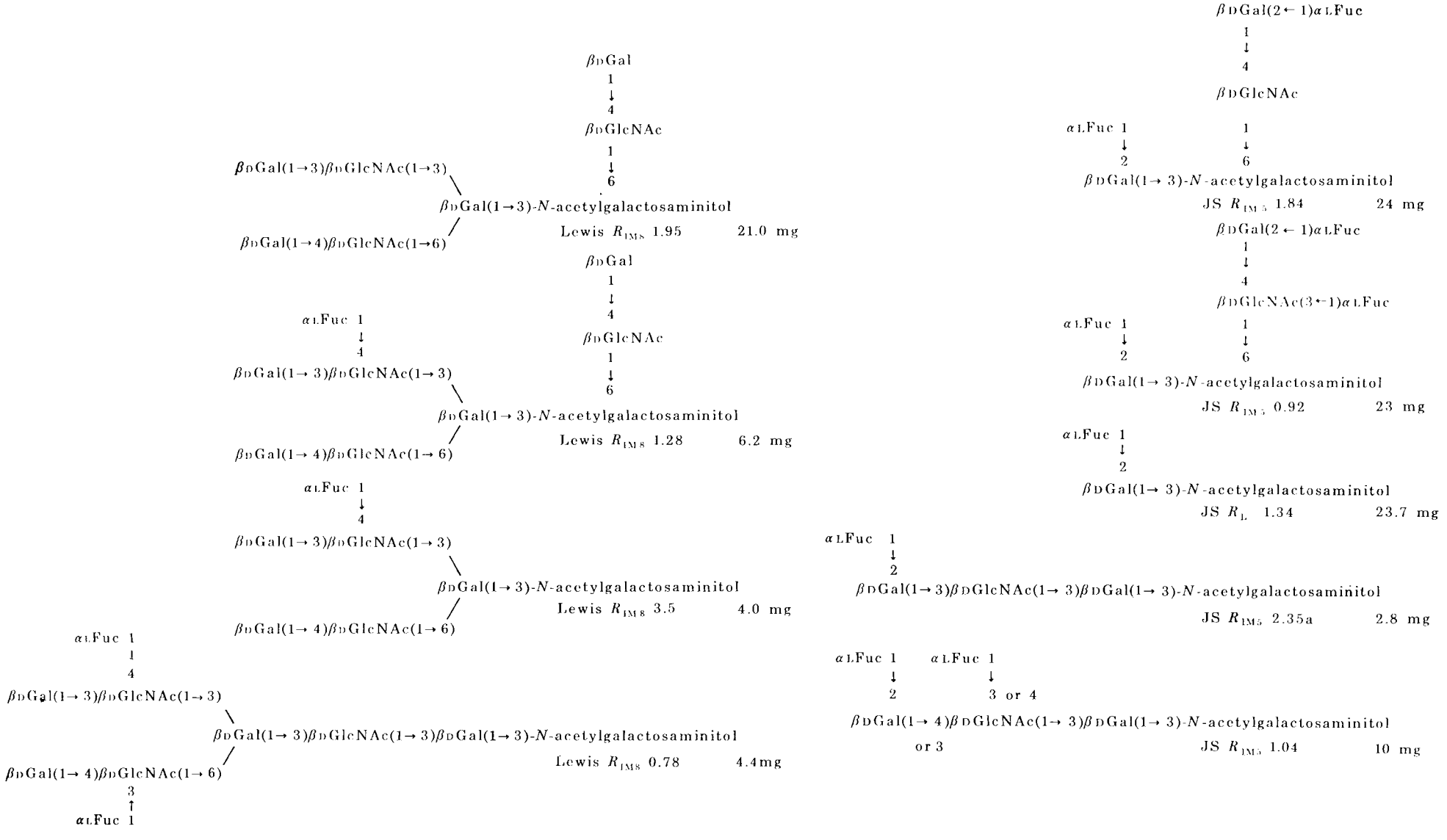

FIGURE 5: Proposed structures and amounts of oligosaccharides isolated from N-1 and from JS.

TABLE III: Analyses, Specific Rotations, and Results of Periodate Oxidations of Isolated Oligosaccharides.

Oligosaccharide[a]	% Composition					Mole Ratios[b]					Opt. Rotation[c] (deg)		Periodate Oxidation		
	N	Fuc	Gal	GlcNAc	*N*-Acetylgalactosaminitol	N	Fuc	Gal	GlcNAc	*N*-Acetylgalactosaminitol	$[\alpha]_D^{27}$	$[\alpha]_{365}^{27}$	Periodate Consumed[d]	Formaldehyde Formed[d]	Formic Acid Formed[d]
Lewis R_{IM8} 0.78	3.2	16.5	39.2	28.8	12.1	4.2	1.9	4.0	2.5	1.0	−66.0	−206	9.2	0.7	5.4
calcd	3.1	18.4	40.5	37.5	12.6	4.0	2.0	4.0	3.0	1.0			10.0	1.0	5.0
Lewis R_{IM8} 1.28[e]	3.1	9.4	36.8	29.0	11.5	4.3	1.1	4.0	3.0	1.0	−43.5	−144	8.7	0.7	4.3
calcd	3.4	10.0	44.1	40.8	13.7	4.0	1.0	4.0	3.0	1.0			9.0	0.0	4.0
Lewis R_{IM5} 0.47	3.4	13.8	44.2	34.3	19.4	2.8	1.0	2.9	1.8	1.0	−64.5	−200	6.8	0.4	3.1
calcd	3.3	12.9	42.7	35.1	17.7	3.0	1.0	3.0	2.0	1.0			7.0	0.0	3.0
Lewis R_{IM8} 1.95	3.7	1.5	47.0	36.6	14.5	4.0	0.01	4.0	2.9	1.0	−29.7	−110	6.7	0.3	3.3
calcd	3.8	0.0	48.5	44.8	15.1	4.0	0.0	4.0	3.0	1.0			7.0	0.0	3.0
Lewis R_{IM8} 3.5	3.3	10.4	39.6	32.7	19.5	2.8	0.7	2.5	1.7	1.0	−52.0	−162	6.7	1.0	3.4
calcd	3.3	12.9	42.7	35.1	17.7	3.0	1.0	3.0	2.0	1.0			8.0	1.0	4.0
Lewis R_L 0.44	3.7	0.8	44.9	24.8	29.8[f]	2.3	0.04	2.1	1.0	1.0	−42.2	−141	4.5	0.2	2.1
calcd	3.7	0.0	47.6	29.5	29.8	2.0	0.0	2.0	1.0	1.0			5.0	0.0	2.0
Lewis R_{IM5} 1.30	2.8	17.3	37.5	21.0	24.9	1.9	1.0	1.9	0.9	1.0	−72.6	−221	6.1	0.3	2.8
calcd	3.1	18.2	40.0	24.7	24.9	2.0	1.0	2.0	1.0	1.0			7.0	0.0	3.0
Lewis R_L 0.86	4.5	0.44	26.9	38.1	40.4	1.8	0.01	0.83	1.0	1.0	−52.4	−163	4.3	0.2	2.0
calcd	4.7	0.0	30.5	37.6	38.0	2.0	0.0	1.0	1.0	1.0			4.0	0.0	2.0
Lewis R_L 1.36	4.6	0.24	45.2	4.2	58.0[f]	1.2	0.0	1.0	0.09	1.0	−53.0	−174	4.0	1.0	2.4
calcd	3.6	0.0	46.6	0.0	58.0	1.0	0.0	1.0	0.0	1.0			4.0	1.0	2.0
JS R_{IM5} 1.80	3.1	21.3	37.4	21.4	23.4	2.0	1.2	2.0	0.93	1.0	−87.0	−274	n.d.[g]		
calcd	3.1	18.2	40.0	24.7	24.9	2.0	1.0	2.0	1.0	1.0					
JS R_{IM5} 1.04[e]	2.0	18.5	20.3	11.9	11.5	2.7	2.2	2.2	1.0	1.0	−57.5	−180	7.1	1.3	3.1
calcd	2.7	31.4	34.4	21.2	21.4	2.0	2.0	2.0	1.0	1.0			7.0	1.0	3.0
JS R_{IM5} 2.35a[e]	2.8	13.0	29.2	17.2	18.0	2.5	1.0	2.2	1.0	1.0	−46.0	−147	5.6	1.2	2.0
calcd	3.1	18.2	40.0	24.7	24.9	2.0	1.0	2.0	1.0	1.0			5.0	1.0	2.0
JS R_{IM5} 2.35b	3.6	1.4	45.8	27.0	25.2	2.2	0.07	2.2	1.1	1.0	−46.3	−159	n.d.[g]		
calcd	3.7	0.0	47.9	29.5	29.8	2.0	0.0	2.0	1.0	1.0					
JS R_{IM5} 1.90	2.5	13.4	34.4	22.0	22.0	1.9	0.84	1.94	1.0	1.0	−54.6	−190	6.0	0.2	2.1
calcd	3.1	18.2	40.0	24.7	24.9	2.0	1.0	2.0	1.0	1.0			6.0	0.0	2.0
JS R_{IM5} 0.91	2.2	28.4	31.6	18.6	19.1	1.9	2.0	2.0	1.0	1.0	−98.0	−304	8.1	0.2	3.2
calcd	2.7	31.4	34.4	21.2	21.4	2.0	2.0	2.0	1.0	1.0			8.0	0.0	3.0
JS R_{IM5} 0.92	2.1	40.0	31.6	15.2	15.5	2.2	3.5	2.5	1.0	1.0	−126.0	−407	8.8	0.1	3.2
calcd	2.4	41.3	30.2	18.6	18.8	2.0	3.0	2.0	1.0	1.0			9.0	0.0	3.0
JS R_{IM5} 1.84	2.7	27.8	31.4	20.9	19.4	2.2	2.0	2.0	1.1	1.0	−90.0	−300	6.8	0.1	2.1
calcd	2.7	31.4	34.4	21.2	21.4	2.0	2.0	2.0	1.0	1.0			7.0	0.0	2.0
JS R_L 1.34	2.7	29.9	33.4	0.0	44.7	0.95	0.91	0.93	0.0	1.0	−124.0	−382	4.9	1.1	1.9
calcd	2.6	30.8	33.8	0.0	42.1	1.0	1.0	1.0	1.0	1.0			5.0	1.0	2.0
JS R_L 1.32[e]	2.8[h]	0.0	32.1	0.0	n.d.	1.0	0.0	0.9	0.0	1.0	−25.0	−127	4.2	1.3	2.2
calcd	3.6	0.0	46.6	0.0	58.0	1.0	0.0	1.0	0.0	1.0			4.0	1.0	2.0

[a] For quantities isolated see Table I and Figure 5. [b] Moles/mole of *N*-acetylgalactosaminitol. [c] Only two values of the five taken are given. [d] Moles/mole of compound. The values of the last determination (usually made after 48 or 72 hr of incubation) are given. [e] Contains inert weight. [f] Theoretical value; this oligosaccharide was used as glc standard (see text). [g] Not determined. [h] N assumed to be *N*-acetylgalactosaminitol.

charides from N-1 were obtained from the same portion (tubes 639–670, 20% ethanol) of the charcoal fractionation of peak A excluded from Bio-Gel P-2 described in Rovis *et al.* (1973a). They were separated by paper chromatography for 5 days in solvent 1. However, structural studies showed that a complete separation had not been achieved and that each compound probably contains some of the other. The small amounts isolated did not allow further purification or more structural studies; therefore, uncertainties remain about the structures of the two compounds.

For Lewis R_{IM8} 0.78 the following criteria are in favor of the tentative structure shown in Figure 5. The analytical data suggest a reduced decasaccharide containing Fuc, Gal, GlcNAc, and *N*-acetylgalactosaminitol in a ratio of 2:4:3:1 (Table III). A first Smith degradation followed by methanolysis and acetylation of the products left GlcNAc and Gal in a ratio of 1.5:1, suggesting that the three GlcNAc and two Gal residues are periodate resistant; *N*-acetylthreosaminitol was also detected. Quantitative periodate oxidation showed consumption of 9.2 mol of periodate (theory, 10) and release of 5.4 (theory, 5.0) mol of formic acid and 0.7 mol (theory, 1.0) of formaldehyde. The value for formaldehyde is lower than expected and this finding in addition to the detection of some 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol among the methylation products suggest a contaminant containing a C-3 and C-6 substituted *N*-acetylgalactosaminitol as in Lewis R_{IM8} 1.28. The peaks for methyl 2,3,4-Me_3Fuc, methyl 2,3,4,6-Me_4Gal, methyl 2,4,6-Me_3Gal, and methyl 3,6-$Ac_2$2,4-Me_2Gal were clearly identified by methylation analysis. The pattern of peaks for the methylated hexosamines was complicated and revealed the presence of a contaminant. The major set of peaks was related to methyl 3,4-Ac_2-6-MeGlcNAc; methyl 3-Ac-4,6-Me_2Glc-NAc and 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol were also identified. However, peaks for methyl 4-Ac-3,6-Me_2GlcNAc and 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol were also present. The methylation analysis is not quantitative, and even a rough quantitation of the derivatives is difficult for a large structure such as this because of the difficulty of obtaining complete methanolysis of all the glycosidic linkages. Therefore it is not possible to decide which of the peaks detected may belong to a contaminant. If the compound contains some Lewis R_{IM8} 1.28, the methyl 4-Ac-3,6-Me_2GlcNAc and the 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol derivatives would come from that structure and therefore the linkage to the more internal GlcNAc in Lewis R_{IM8} 0.78 would be on C-3. Many other variants are compatible with the data available and the structure in Figure 5 must be considered tentative, but it is supported by the biological activities of the oligosaccharide (following paper). For Lewis R_{IM8} 1.28 analytical data suggest a reduced monofucosyl nonasaccharide containing Fuc, Gal, GlcNAc, and *N*-acetylgalactosaminitol in a ratio of 1:4:3:1. After a one-step Smith degradation, methanolysis, and acetylation, GlcNAc and Gal in a ratio of 2.7:1 and *N*-acetylthreosaminitol were detected. Although the values obtained for periodate consumption (8.7 mol) and formic acid (4.3 mol) after periodate oxidation (Table III) are close to the theoretical, the 0.7 mol of formaldehyde (theory 0) suggested a contaminant containing a terminal *N*-acetylgalactosaminitol with a free C-6. The methylation results gave the expected derivatives: methyl 2,3,4-Me_3Fuc, methyl 2,3,4,6-Me_4Gal, methyl 3,6-Ac_2-2,4-Me_2Gal, and methyl 4-Ac-3,6-Me_2GlcNAc as the major hexosamine derivative; methyl 3,4-Ac_2-6-MeGlcNAc and 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol, and in addition small amounts of methyl 2,4,6-Me_3Gal and 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol were detected; they may indicate the presence of some Lewis R_{IM8} 0.78. The analytical data and the biological activities reported in the following paper suggest that the major component of this mixture is a structure identical with Lewis R_{IM8} 1.95 (Figure 5) except for the presence of a fucosyl residue on the C-4 of the GlcNAc of the type 1 chain which is responsible for the Lewis^a activity of this oligosaccharide. The tentative proposal for this compound is in Figure 5 and the same variants as discussed later for Lewis R_{IM8} 1.95 are possible.

Lewis R_{IM5} 0.47 is a reduced heptasaccharide containing Fuc, Gal, GlcNAc, and *N*-acetylgalactosaminitol in a ratio of 1:3:2:1. Methylation gave peaks for methyl 2,3,4-Me_3Fuc, methyl 2,3,4,6-Me_4Gal, and methyl 2,4,6-Me_3Gal, and one of the chromatograms which allowed the identification of 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol, methyl 3,4-Ac_2-6-MeGlcNAc and methyl 4-Ac-3,6-Me_2GlcNAc is shown in Figure 3H. After a first Smith degradation, free *N*-acetylthreosaminitol was detected as the alditol acetate, and after methanolysis and acetylation Gal and GlcNAc in a ratio of 1:1.9 were found. After a second Smith degradation, a compound with Gal and GlcNAc in a ratio of 1:1 remained. The oligosaccharide consumed 6.8 (theory, 7) mol of periodate and released 3.1 (theory, 3.0) mol of formic acid. These data suggest the structure proposed in Figure 5. An alternative structure compatible with the analytical data but not with the biological activity (following paper) would have the Fuc on the GlcNAc linked at C-6 of the *N*-acetylgalactosaminitol.

Lewis R_{IM8} 1.95. Analytical data (Table III) suggest this oligosaccharide to contain 4 mol of Gal and 3 mol of GlcNAc for each *N*-acetylgalactosaminitol. Methylation gave peaks corresponding to methyl 2,3,4,6-Me_4Gal and methyl 3,6-Ac_2-2,4-Me_2Gal. After one Smith degradation *N*-acetylthreosaminitol and Gal and GlcNAc in a ratio of 1:3 were detected after methanolysis and acetylation, indicating that three Gal residues are unsubstituted and one is linked at both C-6 and C-3. By methylation, methyl 3-Ac-4,6-Me_2GlcNAc and methyl 4-Ac-3,6-Me_2GlcNAc were also identified (the chromatogram is in Figure 3G) in the approximate ratio 1:2, indicating that two of the GlcNAc's are substituted at C-4 and one at C-3. A second Smith degradation destroyed all the GlcNAc and only Gal and *N*-acetylthreosaminitol were detected after methanolysis and acetylation by glc. Acetylation of the products of the first and second Smith degradations did not indicate the presence of free *N*-acetylthreosaminitol showing that the βDGal(1→4)DGlcNAc side chain was linked to C-6 rather than to C-3 of the *N*-acetylgalactosaminitol. The peak for the disubstituted *N*-acetylgalactosaminitol was also found by methylation (Figure 3G). The compound reduced 7 mol of sodium periodate (theory, 7) with the release of no formaldehyde (theory, 0) and 3 mol of formic acid (theory, 3). Many structures are compatible with the data available on this reduced octasaccharide, and a proposed one is given in Figure 5. The substituents at C-3 and C-6 of the branched Gal could be as shown in Figure 5 or reversed; in addition a major uncertainty is the location of one of the two Gal-(1→4)GlcNAc residues: it could be directly linked to the C-6 of the terminal *N*-acetylgalactosaminitol (as proposed in Figure 5), or as a second type two chain coming from the branched Gal. However, the structure of Lewis R_{IM8} 3.5, which shows the branched Gal carrying the type 1 and 2 determinants directly linked to C-3 of a monosubstituted *N*-acetylgalactosaminitol, suggests the same situation for Lewis R_{IM8} 1.95. The extent of cross reactivity of this compound with anti-I(Ma) and anti-type XIV pneumococcal serum (Rovis *et al.*, 1973b) is also in favor of the proposed structure.

Lewis R_{IM8} 3.5. Analytical data, periodate oxidation, molar ratios, and optical rotation values suggest that this oligosaccharide is the reduced monofucosyl heptasaccharide shown in Figure 5. The fraction contained about 20% of the same compound lacking the Fuc residue. It consumed 6.7 mol of periodate with release of 3.4 and 1.0 mol of formic acid and formaldehyde, respectively. The optical rotation values are lower than those for Lewis R_{IM5} 0.47 (Table III), which is also a reduced monofucosyl heptasaccharide. Methylation analyses yielded peaks on glc corresponding to methyl 2,3,4-Me_3Fuc, methyl 2,3,4,6-Me_4Gal, methyl 3,6-Ac_2-2,4-Me_2Gal, 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol, methyl 4-Ac-3,6-Me_2-GlcNAc, and methyl 3,4-Ac_2-6-MeGlcNAc. The peaks can be seen on the chromatogram, obtained from the NPGS column separation at 190°, shown in Figure 3I. The same chromatogram shows also a peak with T_r 1.16 which could represent the small amount of methyl 3-Ac-4,6-Me_2GlcNAc present because of the lack of Fuc in 20% of the molecules. The only possible alternative structure for the major component of this mixture would be the one with a reverse substitution at C-3 and C-6 on the branched Gal, although this is not supported by the biological activity of this oligosaccharide (Rovis *et al.*, 1973b).

Lewis R_L 0.44. This reduced oligosaccharide is identical with Lewis R_L 0.41 (Lloyd *et al.*, 1968), OG R_L 0.44 (Vicari and Kabat, 1970), and JS R_{IM5} 2.35b described later in this paper. It contains Gal, GlcNAc, and *N*-acetylgalactosaminitol in a 2:1:1 ratio. After 48 hr at 4° the compound consumed 4.5 mol (theory, 5) of periodate and released 2.1 (theory, 2.0) mol of formic acid. Smith degradation left intact the GlcNAc residue and *N*-acetylthreosaminitol was identified. The glc pattern of the methylation–methanolysis products was identical with the one reported in Figure 3E for OG R_L 0.44 used as a reference.

Lewis R_{IM5} 1.30. Analyses and optical rotation values suggested a reduced pentasaccharide identical with JS R_{IM5} 1.80. Moreover, the two compounds gave the same products after methylation analyses. After one Smith degradation GlcNAc and *N*-acetylthreosaminitol were detected by glc. Quantitative periodate oxidation showed consumption of 6.1 mol (theory, 7.0) of periodate and release of 2.8 (theory, 3.0) mol of formic acid per mol of compound. The activity of this oligosaccharide as inhibitor in the Lewis^a–anti-Lewis^a system (Rovis *et al.*, 1973) indicated that it differs from JS R_{IM5} 1.80 in that the DGal is substituted on C-3 and the LFuc on C-4 of the GlcNAc residue, as shown in Figure 5.

Lewis R_L 0.86 is a reduced trisaccharide containing Gal, GlcNAc, and *N*-acetylgalactosaminitol in equimolar amounts. All three sugars were susceptible to periodate oxidation and *N*-acetylthreosaminitol was identified as one of the products. Methylation analyses showed the presence of methyl 2,3,4,6-Me_4Gal. 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol, and methyl 3,4,6-Me_3GlcNAc. This is the only oligosaccharide isolated in this study containing an unsubstituted GlcNAc residue. Periodate oxidation showed consumption of 4.9 mol (theory, 3.0) of periodate and release of 2 mol of formic acid (theory, 1.0) per mol of compound. The reaction, repeated in parallel with a sample of ethyl βGlcNAc, showed the latter to consume 2 mol of periodate with release of 1 mol of formic acid per mol, suggesting that the unexpected values obtained for this oligosaccharide are due to continued oxidation of the unsubstituted GlcNAc between C-2 and C-3, and not only between C-3 and C-4 as theoretically expected, with the consequent release of an extra mole of formic acid. The atypical behavior of glycosides of GlcNAc toward periodate is known (Kabat, 1961; Jeanloz and Forchielli, 1951). The proposed structure is in Figure 5.

Lewis R_L 1.36. This compound was isolated from fractions E and D of the Bio-Gel P-2 separation (Table III). It contains equimolar amounts of Gal and *N*-acetylgalactosaminitol. Periodate oxidation showed consumption of 3.6 mol (theory, 4) of periodate and release of 2.2 mol of formic acid (theory, 2) and 1 mol of formaldehyde (theory, 1). Smith degradation completely destroyed the structure and *N*-acetylthreosaminitol was detected on glc. Methylation gave the same products as JS R_L 1.32, confirming the identity of the two structures. The same oligosaccharide had been also isolated from OG cyst material and BP_1 substance (*cf.* OG R_{Gal} 0.80 (Vicari and Kabat, 1970) and R_{Gal} 0.87 (Lundblad *et al.*, 1972)).

JS R_{IM5} 1.80. Analyses indicated that this compound was a reduced pentasaccharide containing 2 mol of galactose, 1 mol of fucose, and 1 mol of GlcNAc for every mol of *N*-acetyl-D-galactosaminitol. Glc analysis of the methylated oligosaccharide showed methyl 2,3,4-Me_3Fuc, methyl 2,3,4,6-Me_4Gal, methyl 3,4-Ac_2-6-MeGlcNAc, and 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol. The identification of trimethylfucose and all of the galactose as tetramethylgalactose indicated that the fucose was present as a branch on the GlcNAc. GlcNAc is substituted at C-3 and C-4, so that fucose and galactose could be in either position. However, the absence of Le^a activity of JS cyst and the analogy with the oligosaccharides described later suggested that the fucose is on C-3 and the galactose is on C-4 of the GlcNAc residue, as shown in Figure 4. Only 2.7 mg was isolated, so it could not be studied by quantitative periodate oxidation and Smith degradation.

JS R_{IM5} 1.04. This compound contains about 43% of inert weight. However, the analytical data are proportionately low and did not indicate the presence of more than one compound. The compound is believed to be a reduced hexasaccharide containing 2 mol each of galactose and fucose and 1 mol of GlcNAc for each *N*-acetylgalactosaminitol. Methyl 2,3,4-Me_3Fuc, methyl 3,4,6-Me_3Gal and methyl 3,4-Ac_2-6-MeGlcNAc were detected on methylation and suggest that the two fucoses are bound to a terminal nonreducing galactose and its adjacent GlcNAc residue; methyl 2,4,6-Me_3Gal and methyl 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol were also found, showing that the second galactose is on C-3 of *N*-acetylgalactosaminitol. The oligosaccharide consumed 7.15 (theory, 7) mol of periodate and gave 3.1 (theory, 3) mol of formic acid and 1.3 (theory, 1) mol of formaldehyde after 48 hr of incubation. After a first Smith degradation, Gal and GlcNAc in a ratio of 1:1 and *N*-acetylthreosaminitol were detected by glc of the methanolyzed–acetylated product. These data are consistent with the structure proposed in Figure 5 and resemble JS R_{IM5} 2.35a, isolated from the same charcoal–Celite fraction.

JS R_{IM5} 2.35a. This compound also seems to be a minor type of chain, since only 2.85 mg of the purified compound was obtained, and as with almost all of the minor fractions isolated, the analytical data are proportionally low and show the presence of some inert weight (Lloyd *et al.*, 1966). The analyses and the quantitative periodate oxidation data suggest a reduced pentasaccharide containing 1 mol of fucose, 2 mol of galactose, and 1 mol of GlcNAc for each *N*-acetylgalactosaminitol. The sugar composition is the same as that for JS R_{IM5} 1.80; however, the two compounds eluted at different ethanol concentrations in the charcoal column (Figure 4). JS R_{IM5} 1.80 eluted with fraction 444–475 (16% ethanol) and JS R_{IM5} 2.35a eluted with fraction 475–500 (19% ethanol) together with JS R_{IM5} 1.04, suggesting an analogy in shape of

the former two and a difference from JS R_{IM5} 1.80. Methylation studies showed JS R_{IM5} 2.35a to have a linear structure with a terminal unbranched *N*-acetylgalactosaminitol substituted only at C-3 (3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol was detected). The two galactoses became, after methylation and methanolysis, methyl 3,4,6-Me_3Gal and methyl 2,4,6-Me_3Gal, indicating a terminal galactose substituted on C-2 by fucose and an internal galactose to which GlcNAc is bound on C-3. GlcNAc was detected as methyl 3-Ac-4,6-Me_2GlcNAc. The proposed structure is given in Figure 5.

JS R_{IM5} 2.35b. Analyses and optical rotation values indicate the identity of this reduced tetrasaccharide with Lewis R_L 0.44 described in this paper, and with Lewis R_L 0.41 and OG R_L 0.44 described in Lloyd *et al.* (1968) and Vicari and Kabat (1970). Two moles of galactose and 1 mol of GlcNAc per 1 mol of *N*-acetylgalactosaminitol were calculated from the analytical data. Methylation showed methyl 2,3,4,6-Me_4Gal as the only neutral sugar; methyl 4-Ac-3,6-Me_2GlcNAc and 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol were also identified. In addition, methylation showed two unidentified peaks which may represent some contaminant not found in the OG R_L 0.44 used as a standard. The small amount isolated (2.85 mg) did not allow further study. The cross reaction with type XIV antipneumococcal serum reported in the following paper (Rovis *et al.*, 1973b) supports this structure.

JS R_{IM5} 1.90. The third reduced pentasaccharide isolated from JS has the same composition as JS R_{IM5} 1.80 and JS R_{IM5} 2.35a containing 1 mol of fucose, 2 mol of galactose, and 1 mol of GlcNAc for each *N*-acetylgalactosaminitol. Periodate oxidation and the first Smith degradation destroyed all the Gal and the Fuc leaving intact the GlcNAc and yielding free *N*-acetylthreosaminitol. A second Smith degradation destroyed the GlcNAc, and a peak with the retention time of *N*-acetylserinol resulting from degradation of the free *N*-acetylthreosaminitol was detected by glc. The identification of *N*-acetylserinol was not always possible because of the low molar response of this compound and the small amounts of material surviving the two Smith degradations. The oligosaccharide consumed 6 mol of periodate (theory, 6); no formaldehyde (theory, 0) and 2 mol of formic acid (theory, 2) were released. Thus the GlcNAc was internal and the *N*-acetylgalactosaminitol was substituted on positions 3 and 6, by Gal and GlcNAc, respectively. This was confirmed by methylation with the identification of 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol; methyl 4-Ac-3,6-Me_2GlcNAc showed the GlcNAc to be 4 linked and methyl 2,3,4-Me_3Fuc, methyl 3,4,6-Me_3Gal, and methyl 2,3,4,6-Me_4Gal were compatible with the structure αLFuc-(1→2)βDGal(1→4)βDGlcNAc(1→6)[βDGal(1→3)]-*N*-acetylgalactosaminitol (Figure 5). The other possible structure compatible with methylation and periodate oxidation findings would be substitution of the fucose on C-2 of the galactose which is joined to C-3 of the *N*-acetylgalactosaminitol. The biological properties in the following paper favor the structure in Figure 5.

JS R_{IM5} 0.91. This is a reduced difucosylhexasaccharide containing 2 mol each of fucose and galactose and 1 mol of GlcNAc for each *N*-acetylgalactosaminitol (Table III). It was isolated together with JS R_{IM5} 1.90 from the first major peak (tubes 500–530, 12% ethanol) of the charcoal fractionation shown in Figure 4. The negative rotation of this compound is higher than the values of the monofucosyl reduced pentasaccharide JS R_{IM5} 1.90 and suggests that both the fucosyl residues are α linked (Stanêk *et al.*, 1963). Quantitative periodate oxidation showed the consumption of 8.1 mol (theory, 8.0) of periodate per mol of compound, the release of no formaldehyde as expected, and of 3.2 mol (theory, 3.0) of formic acid. A two-step Smith degradation completely destroyed the structure and gave the same products as did JS R_{IM5} 1.90. Methylation of this oligosaccharide gave the same products as did JS R_{IM5} 1.90 except for the GlcNAc derivative obtained after methylation and methanolysis which was methyl 3,4-Ac_2-6-MeGlcNAc instead of methyl 4-Ac-3,6-Me_2GlcNAc detected in JS R_{IM5} 1.90. Since the two oligosaccharides differ in that one extra fucose is present in JS R_{IM5} 0.91, the methylation data show this fucosyl residue to be on C-3 of the GlcNAc; this position is free in JS R_{IM5} 1.90. This and the biological data (following paper) support the structure proposed in Figure 5. However, other possible structures are compatible with the methylation and periodate oxidation findings; one of the fucoses could be linked to C-2 of either galactose, and both galactose and fucose could be on either C-4 and C-3 of the GlcNAc.

JS R_{IM5} 0.92. This is the largest structure isolated from JS glycoprotein. The molar ratios given in Table III, although based on a per cent composition very close to the calculated theoretical values, are slightly higher than expected, relative to 1 mol of *N*-acetylgalactosaminitol. However, the ratio of methylpentose to N is 1.57 (theory 1.5) and of galactose to N is 1.15 (theory, 1.0) suggesting a reduced heptasaccharide containing three fucoses, two galactoses, and one GlcNAc for each *N*-acetylgalactosaminitol. The higher negative optical rotation for this compound as compared with the mono- and difucosyl reduced oligosaccharides (Table III) supports the presence of a third fucose and also indicates that the fucoses are α linked. GlcNAc and free *N*-acetylthreosaminitol were identified by glc after the first Smith degradation and a second Smith degradation yielded only free *N*-acetylserinol. Complete periodate oxidation of this compound did not take place at 4° after 72 hr of incubation. The reaction, repeated at 27°, gave the expected values for periodate consumption (9 mol) and formic acid (3 mol) and also slight overoxidation occurred after 72 hr at 27°. The rate of the reaction which showed a rapid uptake of 7.4 mol of periodate and gave 2.4 mol of formic acid within the first 8 hr indicated that the bond resistant to rapid oxidation is between C-4 and C-5 of the disubstituted terminal *N*-acetylgalactosaminitol. This was a frequent finding since most of the oligosaccharides studied have a terminal disubstituted *N*-acetylgalactosaminitol and, with many of them, quantitative periodate oxidation was carried out both at 4 and 27°.

Methylation analyses of this compound showed methyl 2,3,4-Me_3Fuc, methyl 3,4,6-Me_3Gal, methyl 3,4-Ac_2-6-MeGlcNAc and 3,6-Ac_2-1,4,5-Me_3-*N*-acetylgalactosaminitol, all supporting the proposed structure (Figure 5). The methylation data allow alternative substitutions on the GlcNAc similar to those discussed for JS R_{IM5} 1.80 and 0.91.

JS R_{IM5} 1.84. This compound shows the same analytical composition and optical rotation values as JS R_{IM5} 0.91 described earlier, and is the third difucosyl-reduced hexasaccharide isolated from JS cyst material. Fucose, galactose, GlcNAc, and *N*-acetylgalactosaminitol are present in a 2:2:1:1 ratio. The compound is eluted from charcoal by 22% ethanol in the same fraction as the reduced heptasaccharide JS R_{IM5} 0.92 (tubes 576–605, Figure 4).

Periodate oxidation after 72 hr of incubation at 4° showed that 6.8 mol (theory, 7.0) of periodate was consumed, and 2.1 mol (theory, 2.0) of formic acid was released. After one Smith degradation the only surviving residue was GlcNAc and *N*-acetylthreosaminitol was detected by glc of both the acetylated and methanolyzed acetylated product, indicating

that the *N*-acetylgalactosaminitol was substituted on C-3 by the Gal and on C-6 by GlcNAc and not *vice versa*. Methylation studies showed the same pattern of derivatives obtained from JS R_{IM5} 0.92, except that methyl 4-Ac-3,6-Me_2GlcNAc was present instead of methyl 3,4-Ac_2-6-MeGlcNAc detected in JS R_{IM5} 0.92. Thus the GlcNAc was not substituted on C-3, the two fucoses are each on the C-2 of the two galactoses, and C-4 of GlcNAc is substituted by an αLFuc(1→2)βDGal residue; the only structure compatible with these data is given in Figure 5.

JS R_L 1.34. This oligosaccharide was shown to be αLFuc-(1→2)βDGal(1→3)-*N*-acetylgalactosaminitol by the following criteria. Analytical data showed equal molar amounts of fucose, galactose, and *N*-acetylgalactosaminitol. The compound consumed 4.9 (theory, 5) mol of periodate with the release of 1.9 (theory, 2.0) mol of formic acid and 1.06 (theory, 1.0) mol of formaldehyde after 24 hr of incubation at 4°. The structure was completely destroyed by Smith degradation and *N*-acetylthreosaminitol was identified by glc. Methylation analyses yielded peaks in glc corresponding to methyl 2,3,4-Me_3Fuc, methyl 3,4,6-Me_3Gal, and 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol.

JS R_{IM5} 1.32. Only 2.95 mg of this compound was isolated and it contains inert weight (Table III), no GlcNAc was detected by analyses and the ratio of Gal to N was 0.9. Methylation showed peaks on Glc corresponding to methyl 2,3,4,6-Me_4Gal and 3-Ac-1,4,5,6-Me_4-*N*-acetylgalactosaminitol. The results of periodate oxidation, corrected for the inert weight, showed consumption of 4.2 mol of periodate (theory, 4.0) and release of 2.2 (theory, 2) and 1.3 (theory, 1.0) mol of formic acid and formaldehyde, respectively. The compound is believed to be βDGal(1→3)-*N*-acetylgalactosaminitol, identical with Lewis R_L 1.36 as described earlier.

Discussion

Alkaline borohydride treatment using 0.05 M NaOH + 1 M $NaBH_4$ at 60° (Iyer and Carlson, 1971) of two immunochemically distinct blood group active glycoproteins with cleavage of the glycosidic linkage of GalNAc to serine and threonine by β elimination yielded more than 95% of the original carbohydrate, most of it as a heterogeneous population of reduced undegraded oligosaccharides (Rovis *et al.*, 1973a). With N-1 61% and with JS 65% of the reduced oligosaccharides (Table III) were either nondialyzable or were excluded from Bio-Gel P-2 indicating the large size of the majority of the chains. The remaining carbohydrate from N-1 (39%) and JS (35%) was included on Bio-Gel P-2 and has been fractionated and studied. Ten different oligosaccharides were isolated from JS and seven from N-1 (Figure 5) comprising less than 20% of the included carbohydrate. Such low yields are a consequence of losses in fractionation, but together with the striking heterogeneity seen on paper chromatographic examination may indicate the presence of many other minor types of chains (as for example JS R_{IM5} 2.35a or Lewis R_{IM5} 1.30) in amounts too small to be isolated. Thus a very restricted population of chains has been fully characterized; some are identical with oligosaccharides previously isolated by different degradative methods. Most are new and fit readily into the structures proposed previously. The majority of the oligosaccharides are pure, by the many criteria already described in Lloyd *et al.* (1966), and the sugars are generally present in proper molar proportions within the precision of the analytical methods. Several, however, contained some inert, noncarbohydrate material, as noted in Table III, a common finding with compounds isolated from chromatographic columns in very small amounts. The following additions and improvements have been made to the systematic methodology of structural characterization already in use in this laboratory: (a) methylation analyses have been improved by the availability of more standards especially for the hexosamine derivatives and the ability to recognize both *N*-acetyl- and *N*-methylhexosamine derivatives, (b) each compound has been analyzed qualitatively and quantatively by glc after methanolysis and acetylation, and (c) almost all the oligosaccharides have been examined after one or two steps of periodate oxidation, reduction, and mild hydrolysis (Smith degradation). Nevertheless several structures are uncertain; the size of some oligosaccharides, the small amounts isolated, and the presence of contaminants were major difficulties and, in certain instances, the structural data available were consistent with more than a single structure. For some substances the biological activities reported in the following paper (Rovis *et al.*, 1973b) permitted a definitive structure to be proposed.

The difficulties of separating oligosaccharides with very similar structures are obvious; however, the number and location of fucosyl residues seemed particularly important in determining chromatographic behavior as seen for the four major oligosaccharides isolated from JS: JS R_{IM5} 0.92 and R_{IM5} 0.91 and JS R_{IM5} 1.84 and R_{IM5} 1.90 have very similar R_F values on paper; however, they could be completely separated from each other by charcoal chromatography because JS R_{IM5} 0.92 and R_{IM5} 1.84 eluted together at 22% ethanol from charcoal (Figure 4) while JS R_{IM5} 0.91 and R_{IM5} 1.90 eluted at 12% ethanol. The presence of GalNAc at the reducing end of all the carbohydrate chains is clearly confirmed by the structures isolated, all of which have a terminal *N*-acetylgalactosaminitol in accord with the alkaline borohydride mechanism. This residue was found to be either disubstituted at C-3 and C-6 or monosubstituted at C-3; no evidence for other linkages has been found. This, of course, does not exclude their presence, and structures with substitutions on C-4 or C-6 or both could occur in the major population of higher oligosaccharide chains left unfractionated. A structure of special interest, obtained from both N-1 (Lewis R_L 0.44) and JS (JS R_{IM5} 2.35b), is βDGal(1→4)βDGlcNAc(1→6)[βDGal-(1→3)]-*N*-acetyl-D-galactosaminitol. Twenty-one and twelve milligrams of this oligosaccharide had been previously isolated by treatment of 860 mg of OG (Vicari and Kabat 1971) and 3000 mg of N-1 (Lloyd *et al.*, 1968), respectively, with 0.2 M NaOD and 1% $NaBD_4$ in D_2O. Since these conditions of alkaline degradation produce peeling of about 50% of Gal linked at C-3 of the terminal GalNAc (Lloyd and Kabat 1968b), the original OG and N-1 glycoproteins would have contained about 50 and 8 mg/g of this type of chain, respectively. In the experiments now reported in which no peeling occurs, amounts corresponding to about 9.5 and 0.1 mg/g of N-1 and of JS have been isolated. Therefore in the precursor OG glycoprotein, this structure is present to the greatest extent, while in N-1 (Lewis[a]) and in JS (H, Le[b]) one-fifth and almost nothing were found, respectively. However, in both, this tetrasaccharide can be recognized as the inner part of the majority of the compounds isolated, indicating that more complex chains are built up on this structure, which seems to be present in all the blood group glycoproteins thus far studied. Five of the JS oligosaccharides (JS R_{IM5} 1.80, 0.91, 1.90, 1.84, and 0.92, Figure 5) each have a different combination of α-L-fucosyl substitutions on this structure. Such substitutions have not previously been reported and thus provide

an indication of the location of the hitherto unaccounted for fucosyl residues previously discussed for JS (Lloyd and Kabat, 1968a). With N-1, from which 44 mg of Lewis R_L 0.44 (Figure 5) had been isolated, a different pattern of substitutions producing longer chains by glycosylation on the two nonreducing terminal Gal residues can be recognized in the structures of Lewis R_{IM5} 0.47, R_{IM8} 1.95, and R_{IM8} 1.28. No comparable structures were obtained from JS perhaps because such oligosaccharides would contain additional fucosyl residues and be in that portion of the JS carbohydrate too large to be fractionated; alternatively a Fuc on C-2 of the terminal nonreducing galactoses might block the action of the enzymes which cause elongation of the carbohydrate chains by addition of more DGlcNAc and DGal residues in accord with the findings that the blood group A, B, and H determinants are always at the nonreducing ends of the chains. The finding of a large amount of JS R_L 1.34 and the structures of JS R_{IM5} 2.35a and R_{IM5} 1.04 would favor this latter possibility. On the other hand such a mechanism, while perhaps convincing for JS, would not explain the large amounts of structures as Lewis R_L 1.36 and Lewis R_L 0.44 in the N-1 and OG glycoproteins, in which, in the absence of an H gene, no fucosyl residues are substituted on the terminal DGal. Thus, these various N-1 and JS oligosaccharides isolated may indicate a general functional significance, perhaps genetically controlled, for small chains occurring along the peptide core, and perhaps of the heterogeneity of the blood group substances themselves. This heterogeneity extends down to the precursor OG glycoprotein which possesses a spectrum of blood group I and i antigenic determinants (Feizi and Kabat, 1972) and which is synthesized by glycosyl transferases acting prior to the A, B, H, and Le^a transferases to give structures identical with Lewis R_L 0.44 (OG R_L 0.44) and Lewis R_L 1.36 (OG R_{Gal} 0.87) in addition to fragments coming from larger chains (Vicari and Kabat, 1970). This heterogeneity would then become more complex by the successive interaction of enzymes determined by the A, B, H, and Lewis genes on the precursor structure. The oligosaccharides described herein comprise the broad spectrum compatible with such a concept. It is still uncertain as to whether the biosynthesis of the carbohydrate chains requires enzymes of very strict specificity, but the diversity of the oligosaccharides listed in Figure 5 suggests either a remarkable flexibility or a very high degree of specificity for the biosynthetic enzymes. Can one assume, for example, that the transferases responsible for substitutions on the branched Gal of Lewis R_{IM8} 0.78 are the same or different from the ones which make the same disubstitution on Lewis R_{IM5} 3.5 or Lewis R_{IM8} 1.95, both of which could well be different substrates?

The isolation of but a single oligosaccharide with a terminal nonreducing GlcNAc residue among the 17 oligosaccharides included on Bio-Gel P-2 argues strongly against these smaller oligosaccharides resulting from incomplete biosynthesis. Were incomplete biosynthesis a random occurrence, one would have expected to find a reasonable number of oligosaccharides which were terminated by GlcNAc. These considerations apply equally to indicate that the blood group substances are not degraded in the cyst cavity.

A structure described in pig submaxillary mucin (Carlson, 1966) and not previously reported in human blood group substances is αLFuc(1→2)βDGal(1→3)-*N*-acetylgalactosaminitol (JS R_L 1.34). The amount of this type of compound isolated may explain the substantial quantities of GalNAc remaining after the first Smith degradation of JS as reported by Lloyd and Kabat (1968a). The enzyme responsible for adding αLFuc to the βDGal residue linked (1→3) to the GalNAc, whether or not it is the same product of the H gene which transfers αLFuc to the βDGal of the type 1 and 2 chains (Figure 1), is not present in N-1, since no substitutions on C-2 of Gal linked to C-3 of *N*-acetylgalactosaminitol were found among the N-1 oligosaccharides (*cf.* Lewis R_L 0.44, R_{IM5} 0.47, R_{IM5} 1.30) and in addition 38 mg of βDGal(1→3)-*N*-acetylgalactosaminitol (Lewis R_L 1.36) was isolated.

As previously noted, more than 60% of the carbohydrate obtained from the two glycoproteins was too large to be fractionated. The material present in the Bio-Gel P-2 excluded (A) fractions of JS and N-1 (Rovis *et al.*, 1973a) appeared to be made up of chains approaching in size the proposed composite oligosaccharide structure of Lloyd and Kabat (1968a) (Figure 1), and indeed a mixture containing Lewis R_{IM8} 0.78 was obtained from N-1 in the excluded fraction. The excluded portion represents about 17 and 35% of N-1 and JS carbohydrate (in JS not all the A fraction was excluded), while 44 and 31%, respectively, were not dialyzable. This material is excluded from Bio-Gel P-4 (unpublished) and contains mostly reduced chains because considerable amounts of *N*-acetylgalactosaminitol were present (Rovis *et al.*, 1973a). The few data available on the intact large size material do not thus far establish an upper limit for the number of sugar residues present in the carbohydrate portion of the blood group substances, or define a complete chain. A significant part of the carbohydrate could be as long or longer and even more complex than predicted on the basis of previous studies with small fragments resulting from β elimination and peeling. The mechanism of stepwise peeling after alkaline elimination, the nature of the degradation products previously found, and the structures now reported suggest that the large nondialyzable chains could represent even more complex variants of the proposed oligosaccharide structure (Figure 1) (Lloyd and Kabat, 1968a).

None of the oligosaccharides thus far isolated from ovarian cyst blood group substances have several GlcNAc residues linked to one another as found with hog gastric mucin blood group A + H active substances (Kochetkov *et al.*, 1970) or two linked GlcNAc residues as in hog gastric mucin sulfated glycoproteins (Slomiany and Meyer, 1972, 1973), although analogous portions of the structure were examined. Such structures could still exist in the large unfractionated oligosaccharides although this is not consistent with the stepwise periodate oxidation results as evaluated both analytically (Lloyd and Kabat, 1968a) and biologically (Vicari *et al.*, 1970) and indicate that the hog gastric mucin and ovarian cyst blood group substances have a different structure despite the isolation of the identical A active pentasaccharide (A R_L 0.52) and H active tetrasaccharide (H R_L 0.75) from both (Lloyd *et al.*, 1966). The isolation of terminal nonreducing α-linked GlcNAc di- and trisaccharides from hog mucin substances (Etzler *et al.*, 1970) substantiates the basic structural differences.

References

Anderson, B., Kabat, E. A., Beychok, S., and Gruezo, F. (1971), *Arch. Biochem. Biophys. 145*, 490.

Anderson, B., Rovis, L., and Kabat, E. A. (1972), *Arch. Biochem. Biophys. 148*, 304.

Carlson, D. M. (1966), *J. Biol. Chem. 241*, 2984.

Ceppellini, R. (1959), Ciba Symposium on the Biochemistry of Human Genetics, Wolstenholme, G. E. W., and O'Connor, G. M., Ed., London, Churchill, p 242.

Etzler, M., Anderson, B., Beychok, S., Gruezo, F., Lloyd,

K. O., Richardson, N., and Kabat, E. A. (1970), *Arch. Biochem. Biophys. 141*, 588.
Feizi, T., and Kabat, E. A. (1972), *J. Exp. Med. 135*, 1247.
Feizi, T., Kabat, E. A., Vicari, G., Anderson, B., and Marsh, W. L. (1971a), *J. Exp. Med. 133*, 39.
Feizi, T., Kabat, E. A., Vicari, G., Anderson, B., and Marsh, W. L. (1971b), *J. Immunol. 106*, 1578.
Ginsburg, V. (1972), *Advan. Enzymol. 36*, 131.
Goldstein, I. J., Hay, G. W., Lewis, B. A., and Smith, F. (1965), *Methods Carbohyd. Chem. 5*, 361.
Hughes, K. W., and Clamp, J. R. (1972), *Biochim. Biophys. Acta 264*, 418.
Iyer, R. N., and Carlson, D. M. (1971), *Arch. Biochem. Biophys. 142*, 101.
Jeanloz, R. W., and Forchielli, E. (1951), *J. Biol. Chem. 188*, 361.
Kabat, E. A. (1961), Kabat and Mayer's Experimental Immunochemistry, 2nd ed, Springfield, Ill., C. C. Thomas.
Kabat, E. A. (1970), *in* Blood and Tissue Antigens, Aminoff, D., Ed., New York, N. Y., Academic Press, p 187.
Kabat, E. A. (1973), *Advan. Chem. Ser. No. 101*, 334.
Kabat, E. A., Bassett, E. W., Pryzwansky, K., Lloyd, K. O., Kaplan, M. E., and Layug, E. J. (1965), *Biochemistry 4*, 1632.
Kobata, A., and Ginsburg, V. (1969), *J. Biol. Chem. 244*, 5496.
Kochetkov, N. K., Derevitskaya, V. A., Lichosherstov, L. M., Martynova, M. D., and Senchenkova, S. N. (1970), *Carbohyd. Res. 12*, 437.
Kuhn, R., Baer, H. H., and Seeliger, A. (1958), *Justus Liebigs Ann. Chem. 611*, 236.
Lee, Y. C., and Scocca, J. R. (1972), *J. Biol. Chem. 247*, 5753.
Lloyd, K. O., and Kabat, E. A. (1968a), *Proc. Nat. Acad. Sci. U. S. 61*, 1470.
Lloyd, K. O., and Kabat, E. A. (1968b), *Carbohyd. Res. 9*, 41.
Lloyd, K. O., Kabat, E. A., Layug, E. J., and Gruezo, F. (1966), *Biochemistry 5*, 1489.
Lloyd, K. O., Kabat, E. A., and Licerio, E. (1968), *Biochemistry 7*, 2976.
Ludowieg, J., and Benmaman, J. D. (1967), *Anal. Biochem. 19*, 80.
Lundblad, A., Hammarström, S., Licerio, E., and Kabat, E. A. (1972), *Arch. Biochem. Biophys. 148*, 291.
Morgan, W. T. J. (1960), *Proc. Roy. Soc., Ser. B 151*, 308.
Rovis, L., Anderson, B., Kabat, E. A., Gruezo, F., and Liao, J. (1973a), *Biochemistry 12*, 1955.
Rovis, L., Kabat, E. A., Pereira, M., and Feizi, T. (1973b), *Biochemistry 12*, 5355.
Schiffman, G., Kabat, E. A., and Thompson, W. (1964), *Biochemistry 3*, 113.
Slomiany, B. L., and Meyer, K. (1972), *J. Biol. Chem. 247*, 5062.
Slomiany, B. L., and Meyer, K. (1973), *J. Biol. Chem. 248*, 2290.
Spiro, R. G. (1970), *Annu. Rev. Biochem. 39*, 599.
Staněk, J., Cerny, M., Kocourek, J., and Pacak, J. (1963), The Monosaccharides, New York, N. Y., Academic Press, p 55.
Vicari, G., and Kabat, E. A. (1969), *J. Immunol. 102*, 821.
Vicari, G., and Kabat, E. A. (1970), *Biochemistry 9*, 3414.
Vicari, G., Sher, A., Cohn, M., and Kabat, E. A. (1970), *Immunochemistry 7*, 829.
Watkins, W. M. (1962), *Immunology 5*, 245.
Watkins, W. M. (1970), *in* Blood and Tissue Antigens, Aminoff, D., Ed., New York, N. Y., Academic Press, p 441.
Watkins, W. M. (1972), *in* Glycoproteins, 2nd ed, Gottschalk, A., Ed., New York, N. Y., Elsevier, p 886.
Watkins, W. M., and Morgan, W. T. J. (1959), *Vox Sang. 4*, 97.