HIGH-RESOLUTION, ^{1}H-NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY AS A TOOL IN THE STRUCTURAL ANALYSIS OF CARBOHYDRATES RELATED TO GLYCOPROTEINS

By Johannes F. G. Vliegenthart, Lambertus Dorland, and Herman van Halbeek

Department of Bio-Organic Chemistry, University of Utrecht, Utrecht, The Netherlands

I. General Introduction

Glycoproteins are biopolymers consisting of a polypeptide backbone bearing one or more covalently linked carbohydrate chains. During the past decade, interest in the structure and function of glycoproteins has increased greatly, as it was found that the carbohydrate chains of these polymers are involved in several important biochemical processes. In particular, their roles in recognition phenomena, in immunological events, and in determining the life-span of cells and glycoproteins must be mentioned. For understanding of the biological function of the glycan chains, detailed knowledge of their structures is a prerequisite.

The carbohydrate chains of glycoproteins may be classified according to the type of linkage to the polypeptide backbone. *N*-Glycosylic chains are attached to the amide group of asparagine (Asn), whereas the *O*-glycosylic chains are linked to the hydroxyl group of such amino acid residues as serine (Ser), threonine (Thr), and hydroxylysine (Hyl). As a whole, the carbohydrate chains show a large variety in pri-

ISBN 0-12-007241-6

mary structure, as has been discussed by Montreuil,[1,2] the Kornfelds,[3] and Sharon and Lis.[4]

The unambiguous determination of the primary structure of carbohydrates is much more cumbersome than for other biopolymers. The large number of glycosylic linkages possible, in conjunction with the occurrence of branching, yields a fantastically large number of theoretically possible isomers, even for a relatively simple oligosaccharide. This demands a high degree of sophistication in methods used for analysis of the structure.

The first analytical problem is encountered at the level of the glycoprotein as such. Owing to natural or artificially introduced (micro)heterogeneity in the carbohydrate chains, it is virtually impossible to obtain the polymer in the form of a single molecular species. Secondly, analysis of the complete primary structure can so far not be conducted on intact glycoproteins, and degradation to glycopeptides, oligosaccharides, or oligosaccharide-alditols is obligatory. Thirdly, the fractionation of more or less complex mixtures of (closely) related, partial structures is difficult. Reliable checks for the purity of the compound isolated are indispensable, because, if a sample is considered to be homogeneous, but, in fact, consists of a mixture of closely related components, incorrect structures may be deduced. This might be one of the reasons why it is not exceptional that, for one and the same carbohydrate side-chain of a certain glycoprotein, more than one structure has been reported.

Over the years, several strategies have been developed for determination of the structure of carbohydrate chains; for concise reviews, see Refs. 3 and 4. In particular, the refinements of methylation analysis,[5,5a] and chemical[6–8] and enzymic degradation methods[9–10] have permitted

(1) J. Montreuil, *Adv. Carbohydr. Chem. Biochem.*, 37 (1980) 157–223.
(2) J. Montreuil, in A. Neuberger (Ed.), *Comprehensive Biochemistry*, Vol. 19 B, Part II, Elsevier, Amsterdam, 1982, pp. 1–188.
(3) R. Kornfeld and S. Kornfeld, in W. J. Lennarz (Ed.), *The Biochemistry of Glycoproteins and Proteoglycans*, Plenum, New York, 1980, pp. 1–34.
(4) N. Sharon and H. Lis, in H. Neurath and R. L. Hill (Eds.), *The Proteins*, 3rd edn., Vol. V, Academic Press, New York, 1982, pp. 1–144.
(5) B. Lindberg and J. Lönngren, *Methods Enzymol.*, 50 (1978) 3–33.
(5a) H. Rauvala, J. Finne, T. Krusius, J. Kärkkäinen, and J. Järnefelt, *Adv. Carbohydr. Chem. Biochem.*, 38 (1981) 389–416.
(6) B. Lindberg, J. Lönngren, and S. Svensson, *Adv. Carbohydr. Chem. Biochem.*, 31 (1975) 185–240.
(7) G. O. Aspinall, *Pure Appl. Chem.*, 49 (1977) 1105–1134.
(8) G. Strecker, A. Pierce-Crétel, B. Fournet, G. Spik, and J. Montreuil, *Anal. Biochem.*, 111 (1981) 17–26.
(9) Y.-T. Li and S.-C. Li, in M. I. Horowitz and W. Pigman (Eds.), *The Glycoconjugates*, Vol. I, Academic Press, New York, 1977, pp. 51–67.
(9a) H. M. Flowers and N. Sharon, *Adv. Enzymol.*, 48 (1979) 29–95.
(10) A. Kobata, *Anal. Biochem.*, 100 (1979) 1–14.

the unraveling of new structures. The enzymic methods were greatly improved by rigorous purification of isolated exo- and endo-glycosidases. However, the classical methods have some inherent limitations, and are time- and material-consuming.

In the past few years, we have had the opportunity to introduce high-resolution, ^{1}H-n.m.r. spectroscopy for determination of the structure of glycan chains of glycoproteins.[11–13] In close collaboration with J. Montreuil and his colleagues (Lille, France), we have shown that this technique, in conjunction with methylation analysis, is extremely suitable for structural studies on *N*-, as well as on *O*-, glycosylic glycans. The present article contains a discussion of the high-resolution, ^{1}H-n.m.r. spectra of compounds, derived from *N*-glycosylically linked carbohydrate chains, comprising the *N*-acetyllactosamine and oligomannoside types of structure.

1. High-resolution, ^{1}H-N.m.r. Spectroscopy

^{1}H-N.m.r. spectroscopy has contributed significantly to extensions of our knowledge on the structure and conformation of biomolecules as well as on intra- and inter-molecular, interaction processes. N.m.r. spectroscopy is, however, an inherently insensitive technique, and the richness of information contained in an n.m.r. spectrum often limits the size and complexity of the molecules that can usefully be studied.

Advances in instrument design have now greatly improved the sensitivity of the spectrometers, so that resonances can readily be observed in aqueous solutions of, for instance, complex carbohydrates, at concentrations of the order[13] of 0.05 m*M*. Increase in the strength of magnetic fields (up to[14] 14 Tesla) has enabled the study of molecules of larger molecular weight and complexity by inducing better spectral-resolution. Moreover, the availability of sophisticated computer-programs allows an artificial resolution-enhancement,[15] for example, by Lorentzian to Gaussian transformation; for general reviews, see

(11) J. Montreuil and J. F. G. Vliegenthart, in J. D. Gregory and R. W. Jeanloz (Eds.), *Glycoconjugate Research, Proc. Int. Symp. Glycoconjugates, 4th*, Vol. I, Academic Press, New York, 1979, pp. 35–78.

(12) J. F. G. Vliegenthart, H. van Halbeek, and L. Dorland, in A. Varmavuori (Ed.), *IUPAC Int. Congr. Pure Appl. Chem., 27th*, Helsinki, 1979, Pergamon, Oxford, 1980, pp. 253–262.

(13) J. F. G. Vliegenthart, H. van Halbeek, and L. Dorland, *Pure Appl. Chem.*, 53 (1981) 45–77.

(14) M. Llinás, A. de Marco, and J. T. J. Lecomte, *Biochemistry*, 19 (1980) 1140–1145; D. G. Davis and B. F. Gisin, *FEBS Lett.*, 133 (1981) 247–251; M. L. Hayes, A. S. Serianni, and R. Barker, *Carbohydr. Res.*, 100 (1982) 87–101.

(15) R. R. Ernst, *Adv. Magn. Reson.*, 2 (1966) 1–135.

Refs. 16–20. Already, a partial interpretation of an n.m.r. spectrum can provide detailed information about molecular structure, as will be shown.

For interpretation of the ^{1}H-n.m.r. spectrum of a carbohydrate chain in terms of primary structural assignments, the concept of "*structural-reporter groups*" was developed.[12,13] This means that the chemical shifts of protons resonating at clearly distinguishable positions in the spectrum, together with their coupling constants and the line widths of their signals, bear the information essential to permit assigning of the primary structure. In high-resolution, ^{1}H-n.m.r. spectra of relatively large, *N*-glycosylic carbohydrate structures, resonances of the following structural-reporter groups can be discerned. (*a*) *Anomeric Protons (H-1 Atoms).* Their chemical shift and coupling constant provide information on the kind of sugar residue, as well as on the type and configuration of its glycosylic linkage. (*b*) *Mannose H-2 and H-3 Atoms.* The pattern of their signals is, as a whole, indicative of the type of substitution of the common, mannotriose branching-core. (*c*) *Sialic Acid H-3 Atoms.* Their chemical shifts are characteristic for the type and configuration of the glycosylic linkage of the sialic acid residue, and, in some cases, for the location of the residue in the chain. (*d*) *Fucose H-5 and CH_3 Atoms.* The chemical shifts of these protons, together with that of H-1 of the residue, are indicative of the type and configuration of its glycosylic linkage, and of the structural environment, in particular of the residue to which the fucose is attached. (*e*) *Galactose H-3 and H-4 Atoms.* Their chemical shifts are, in some cases, useful for characterizing the type and configuration of the glycosylic linkage between galactose and its substituent. (*f*) *Amino Sugar (2-Acetamido-2-deoxyglucose and Sialic Acid) N-Acetyl-CH_3 Protons.* Their chemical shifts are sensitive to even small structural variations, making this region of the spectrum highly informative.

The line widths of the signals of the structural-reporter groups are influenced by the local mobility of the protons. This will be illustrated, in particular, for the anomeric-proton signals and for the *N*-acetyl signals.

(16) R. A. Dwek, *Nuclear Magnetic Resonance (N.M.R.) in Biochemistry: Applications to Enzyme Systems,* Clarendon Press, Oxford, 1973.
(17) P. F. Knowles, D. Marsh, and H. W. E. Rattle, *Magnetic Resonance of Biomolecules,* Wiley, New York, 1976.
(18) K. Wüthrich, *NMR in Biological Research: Peptides and Proteins,* North-Holland, Amsterdam, 1976.
(19) L. J. Berliner and J. Reuben, *Biological Magnetic Resonance,* Vols. 1 and 2, Plenum, New York, 1978.
(20) R. G. Shulman, *Biological Applications of Magnetic Resonance,* Academic Press, New York, 1979.

Besides the aforementioned n.m.r. parameters [chemical shift (δ), coupling constant (J), and line width], spectral integration can give valuable information. The relative intensities of the structural-reporter-group signals in the n.m.r. spectrum can be used as markers for the purity of the compound. Often, from the spectrum, it can be deduced whether or not the sample consists of more than one carbohydrate structure, and in which molar ratios the components of the mixture, and the sugar residues in each of these, occur.

In the ^{1}H-n.m.r. spectrum of a carbohydrate chain, a broad signal is observed, between δ ~3.4 and ~4.0, that is derived from the bulk of nonanomeric, sugar-skeleton protons, but, so far, it could not be resolved into contributions of individual protons. In the case of glycopeptides, additional signals, derived from protons of the amino acid residues, are found in the spectrum.

It should be stressed that a high-resolution, ^{1}H-n.m.r. spectrum of a compound provides a measure of the structural identity which, even if the spectrum cannot be completely interpreted, renders possible a comparison with the spectra of compounds obtained from other sources, allowing a decision as to whether or not the compounds are identical. After recording of this "identity card," the unimpaired compound may be submitted to chemical and enzymic degradation-procedures.

In this article, the 500-MHz, or, sometimes, 360-MHz, ^{1}H-n.m.r. spectra of some seventy *N*-glycosylic carbohydrate chains will be discussed. The spectra were recorded in D_2O at ambient temperature and at pD ~7, unless indicated otherwise; for experimental details, see Section IV. First, the outstanding features of the spectra of fundamental elements of *N*-acetyllactosamine-type structures will be treated in detail (compounds **1–20**); this part also covers the characteristics of the intersecting, GlcNAc residue in this type of structure. Secondly, extensions of these chains with differently linked, sialic acid residues (compounds **21–41**), and thirdly, with fucose residues (compounds **42–54**), will be discussed. Besides the structural-reporter-group signals of the newly introduced residues, the effects of extension on the remaining signals in the spectra will be traced. Next, some unusual, *N*-acetyllactosamine-type, *N*-glycosylic carbohydrate structures (compounds **55–60**) containing a virtually abnormal core-region, namely, βGal(1→4)βGlcNAc(1→N)Asn, will be considered. Finally, the n.m.r. characteristics of oligomannoside-type carbohydrate chains (compounds **61–72**) will be presented; in particular, the second branching-point, and the characteristics and influences of α-(1→2)-linked mannose residues occurring in these structures, are the subjects of discussion.

2. Literature Data on High-resolution, ^{1}H-N.m.r. Spectroscopy of Carbohydrates Derived from Glycoconjugates

The fundamental work of Lemieux and coworkers[21] introduced the successful application of ^{1}H-n.m.r. spectroscopy to structural problems in the carbohydrate field. Since then, numerous studies have been devoted to the (partial) determination of primary structures of carbohydrates and derivatives thereof, as well as to the elucidation of their conformation in solution, by means of ^{1}H-n.m.r. spectroscopy. An exhaustive discussion of all these contributions is beyond the scope of this article. For comprehensive reviews, see Refs. 22–28.

Regarding the applicability of ^{1}H-n.m.r. spectroscopy for elucidation of the structure of the carbohydrate moieties of glycoconjugates, several reports have been published. In 1973, one of the first examples of employment of high-resolution, ^{1}H-n.m.r. spectroscopy for structural studies on intact glycolipids[29] afforded the 220-MHz, ^{1}H-n.m.r.-spectral data for some peracetylated galactocerebrosides, determined in three different solvents. In 1979, Falk and coworkers[30–32a] described 270-MHz, ^{1}H-n.m.r. spectroscopy of non-degraded, permethylated and permethylated-reduced derivatives of (blood-group active) glycosphingolipids (spectra were recorded in chloroform solution, at probe temperatures of ~25 and ~40°), as a suitable approach for the determination of the configuration of the anomeric linkages in the sugar chains. According to Karlsson and coworkers,[33–34a] these data, at most, supplement the structural fingerprinting of lipid-linked

(21) R. U. Lemieux, R. K. Kullnig, H. J. Bernstein, and W. G. Schneider, *J. Am. Chem. Soc.*, 80 (1958) 6098–6105.
(22) L. D. Hall, *Adv. Carbohydr. Chem.*, 19 (1964) 51–93.
(23) T. D. Inch, *Annu. Rev. NMR Spectrosc.*, 2 (1969) 35–82.
(23a) P. L. Durette and D. Horton, *Adv. Carbohydr. Chem. Biochem.*, 26 (1971) 49–125.
(24) B. Coxon, *Adv. Carbohydr. Chem. Biochem.*, 27 (1972) 7–83.
(25) G. Kotowycz and R. U. Lemieux, *Chem. Rev.*, 73 (1973) 669–698.
(26) L. D. Hall, *Adv. Carbohydr. Chem. Biochem.*, 29 (1974) 11–40.
(27) D. B. Davies, *Nucl. Magn. Reson.*, 9 (1980) 182–203.
(28) L. D. Hall, in W. Pigman and D. Horton (Eds.), *The Carbohydrates: Chemistry and Biochemistry*, 2nd edn., Vol. IB, Academic Press, New York, 1980, pp. 1300–1326.
(29) M. Martín-Lomas and D. Chapman, *Chem. Phys. Lipids*, 10 (1973) 152–164.
(30) K.-E. Falk, K.-A. Karlsson, and B. E. Samuelsson, *Arch. Biochem. Biophys.*, 192 (1979) 164–176.
(31) K.-E. Falk, K.-A. Karlsson, and B. E. Samuelsson, *Arch. Biochem. Biophys.*, 192 (1979) 177–190.
(32) K.-E. Falk, K.-A. Karlsson, and B. E. Samuelsson, *Arch. Biochem. Biophys.*, 192 (1979) 191–202.
(32a) K.-E. Falk, K.-A. Karlsson, H. Leffler, and B. E. Samuelsson, *FEBS Lett.*, 101 (1979) 273–276.

oligosaccharides by mass spectrometry.

Since 1980, Dabrowski and coworkers[35–37c] have published 360-MHz, ^{1}H-n.m.r. data for a number of underivatized glycosphingolipids (blood-group active glycosylceramides, as well as more-complex compounds related to the Forssman glycolipid); the spectra were recorded at 65° for solutions in dimethyl sulfoxide-d_6 containing a trace of D_2O. These authors were able to assign all of the anomeric-proton signals, and a number of nonanomeric-proton resonances. The chemical shifts of these protons were found to be dependent on essential, primary, structural parameters.[37–37c] Most of the chemical shifts of the ring protons were determined by spin-decoupling, and nuclear Overhauser, difference spectroscopy.[35,37–37c] Concerning the potential of *J*-resolved, two-dimensional, ^{1}H-n.m.r. spectroscopy for this purpose, Yamada and coworkers[38] drew attention to the fact that, despite the many useful aspects of this method for carbohydrates (see Refs. 28 and 39), these experiments require a fair amount of time and material.

In the field of glycoproteins, the following high-resolution, ^{1}H-n.m.r.-spectral studies for solutions in D_2O on underivatized, glycan chains related to, or identical with, those that will be described herein deserve mention. Wolfe and coworkers published,[40] in 1974 and[41] 1975, 220-MHz, ^{1}H-n.m.r. spectra of two oligosaccharides related to the *N*-acetyllactosamine type of *N*-glycosylic carbohydrate chain.

(33) K.-A. Karlsson, in Ref. 12, pp. 171–183.
(33a) M. E. Breimer, G. C. Hansson, K.-A. Karlsson, and H. Leffler, *Biochim. Biophys. Acta*, 617 (1980) 85–96.
(33b) G. C. Hansson, K.-A. Karlsson, and J. Thurin, *Biochim. Biophys. Acta*, 620 (1980) 270–280.
(34) K.-E. Falk, K.-A. Karlsson, and B. E. Samuelsson, *FEBS Lett.*, 124 (1981) 173–177.
(34a) J. Ångström, M. E. Breimer, K.-E. Falk, I. Griph, G. C. Hansson, K.-A. Karlsson, and H. Leffler, *J. Biochem. (Tokyo)*, 90 (1981) 909–921.
(35) J. Dabrowski, H. Egge, P. Hanfland, and S. Kuhn, in C. C. Sweeley (Ed.), *Cell Surface Glycolipids*, ACS Symp. Ser. 128, American Chemical Society, Washington D. C., 1980, pp. 55–64.
(36) J. Dabrowski, H. Egge, and P. Hanfland, *Chem. Phys. Lipids*, 26 (1980) 187–196.
(37) J. Dabrowski, P. Hanfland, and H. Egge, *Biochemistry*, 19 (1980) 5652–5658.
(37a) J. Dabrowski, P. Hanfland, H. Egge, and U. Dabrowski, *Arch. Biochem. Biophys.*, 210 (1981) 405–411.
(37b) P. Hanfland, H. Egge, U. Dabrowski, S. Kuhn, D. Roelcke, and J. Dabrowski, *Biochemistry*, 20 (1981) 5310–5319.
(37c) J. Dabrowski, P. Hanfland, and H. Egge, *Methods Enzymol.*, 83 (1982) 69–86.
(38) A. Yamada, J. Dabrowski, P. Hanfland, and H. Egge, *Biochim. Biophys. Acta*, 618 (1980) 473–479.
(39) L. D. Hall, G. A. Morris, and S. Sukumar, *Carbohydr. Res.*, 76 (1979) c7–c9.
(40) L. S. Wolfe, R. G. Senior, and N. M. K. Ng Ying Kin, *J. Biol. Chem.*, 249 (1974) 1828–1838.
(41) N. M. K. Ng Ying Kin and L. S. Wolfe, *Biochem. Biophys. Res. Commun.*, 66 (1975) 123–130.

These compounds were isolated from the liver[40] and the urine[41] of patients suffering from GM_1-gangliosidosis type I. Their structures are identical to those of the di- and tri-antennary, asialo oligosaccharides **7** and **10** of this article (see Chart 1). The spectra were recorded at 70°, and the chemical shifts were measured in p.p.m. relative to external tetramethylsilane as the standard. In later studies,[42,43] these authors arrived at a partial interpretation of the 100-MHz, and 90-MHz, ^{1}H-n.m.r. spectra, recorded at 70 and 60°, respectively, of some oligosaccharides stemming from the liver of a patient who died of GM_2-gangliosidosis variant O (Sandhoff–Jatzkewitz disease). The structures of the isolated hexa- and hepta-saccharides are the same as those respectively denoted **6** and **14** herein, and as that of the minor constituent present in the heptasaccharide mixture of which compound **14** forms the major part (see legend to Fig. 15). The linear, GM_2-gangliosidosis tetrasaccharides described in Ref. 43, namely, βGlcNAc(1→2)αMan-(1→3)βMan(1→4)GlcNAc and βGlcNAc(1→4)αMan(1→3)βMan-(1→4)GlcNAc, are not incorporated in the series of compounds discussed herein. In addition, the authors described[43] the 90-MHz, ^{1}H-n.m.r. data for a trisaccharide (compound **5**) isolated from the urine of mannosidosis patients.

In 1977 and 1980, we respectively introduced the application of 360-MHz, and 500-MHz, ^{1}H-n.m.r. spectroscopy for elucidation of the structure of underivatized carbohydrate chains obtained from glycoproteins. Since then, several other research groups have become active in this field. For reasons outlined later, the results of their work will be briefly summarized first. In 1979, Kin and Wolfe[44] published the 220-MHz, ^{1}H-n.m.r. spectra, recorded at 70°, of a mono-antennary glyco-asparagine possessing the same structure as **46** or **47**, and of the corresponding oligosaccharide, having[44a] GlcNAc-2 as the terminal, reducing-sugar residue. Like **46** and **47**, the compounds were isolated from the urine of a fucosidosis patient.

Schachter and coworkers[45] published some 360-MHz, ^{1}H-n.m.r. data for a series of glycopeptides having *N*-glycosylically linked carbohydrate chains of the *N*-acetyllactosamine type, and one having a chain of the oligomannoside type; spectra were recorded at a variety of probe temperatures (20, 25, 30, 70, and 85°). The compounds

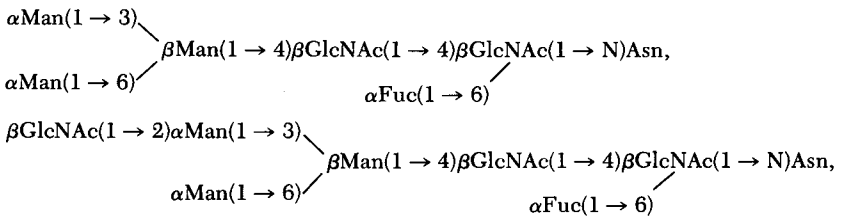
(42) N. M. K. Ng Ying Kin and L. S. Wolfe, *Biochem. Biophys. Res. Commun.*, 59 (1974) 837–844.

(43) N. M. K. Ng Ying Kin and L. S. Wolfe, *Carbohydr. Res.*, 67 (1978) 522–526.

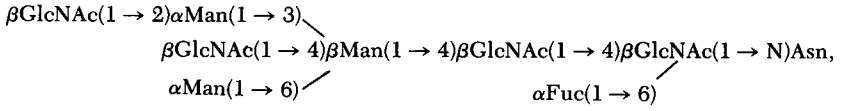
(44) N. M. K. Ng Ying Kin and L. S. Wolfe, *Biochem. Biophys. Res. Commun.*, 88 (1979) 696–705.

(44a) For the system of numbering used for the sugar residues in *N*-glycosylic carbohydrate chains, see the footnote on page 221.

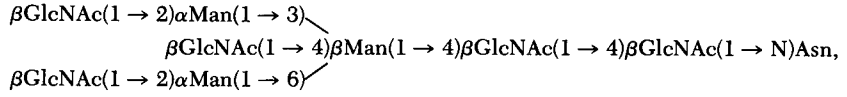
(45) S. Narasimhan, N. Harpaz, G. Longmore, J. P. Carver, A. A. Grey, and H. Schachter, *J. Biol. Chem.*, 255 (1980) 4876–4884.



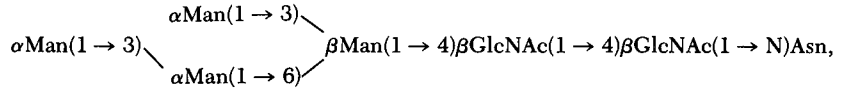
and



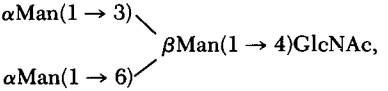
were prepared from a human myeloma IgG; the other two, namely,



and



were obtained from hen-egg albumin. Subsequently, Cohen and Ballou[46] exhaustively discussed the 180-MHz, ^{1}H-n.m.r. data (40°) for a large number of building units (both as oligosaccharides and as glycopeptides) of the oligomannoside type of *N*-glycosylic carbohydrate chain possessing 1 through 8 mannose residues. Most of the compounds treated in Ref. 46 will also be subject to discussion herein (corresponding to structures **5**, **62–65**, and **68–70**). However, some of them, for example, βMan(1→4)GlcNAc and

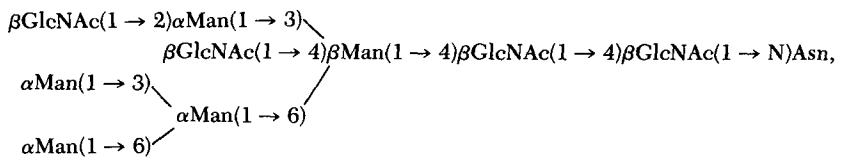


are not included. It should be noted that the pioneering work of Gorin and coworkers[47–49] on yeast mannans and galactomannans provided, for resonances characteristic for a series of oligosaccharides, assignments

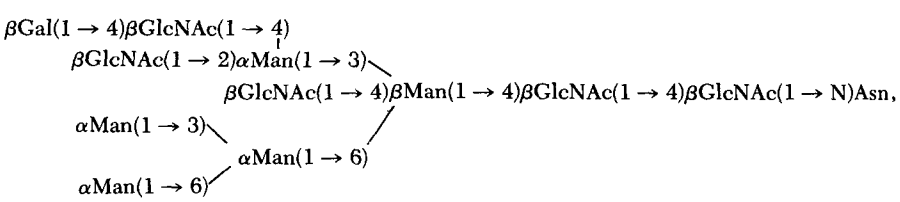
(46) R. E. Cohen and C. E. Ballou, *Biochemistry*, 19 (1980) 4345–4358.
(47) P. A. J. Gorin and J. F. T. Spencer, *Can. J. Chem.*, 46 (1968) 2299–2304.
(48) P. A. J. Gorin, M. Mazurek, and J. F. T. Spencer, *Can. J. Chem.*, 46 (1968) 2305–2310.
(49) P. A. J. Gorin, J. F. T. Spencer, and S. S. Bhattacharjee, *Can. J. Chem.*, 47 (1969) 1499–1505.

that proved to be valuable for high-resolution, ^{1}H-n.m.r.-spectral work.[46]

Finally, in addition to ^{1}H-n.m.r. data for structures identical to **62** and **63**, Carver and coworkers[50] reported the 360-MHz, ^{1}H-n.m.r.-spectral data for two glycopeptides possessing a hybrid type[1] of structure, namely,



and



In this study, n.O.e. difference-spectroscopy was applied to enable the making of some assignments.

The spectral data for the compounds, the detailed structures of which have been presented in Section I, 2, are not compiled in the following Sections, because they were obtained from spectra recorded under more or less different experimental conditions (regarding strength of magnetic field, and reference standard, but, most of all, probe temperature) (compare, our experimental conditions, Section IV). This choice was dictated solely by the impossibility of making these data comparable with ours, and is not intended to imply any criticism of the work performed in other laboratories.

II. High-resolution, ^{1}H-N.m.r. Spectroscopy of Carbohydrates Related to Glycoproteins of the *N*-Glycosylic Type

1. Carbohydrate Chains of the *N*-Acetyllactosamine Type

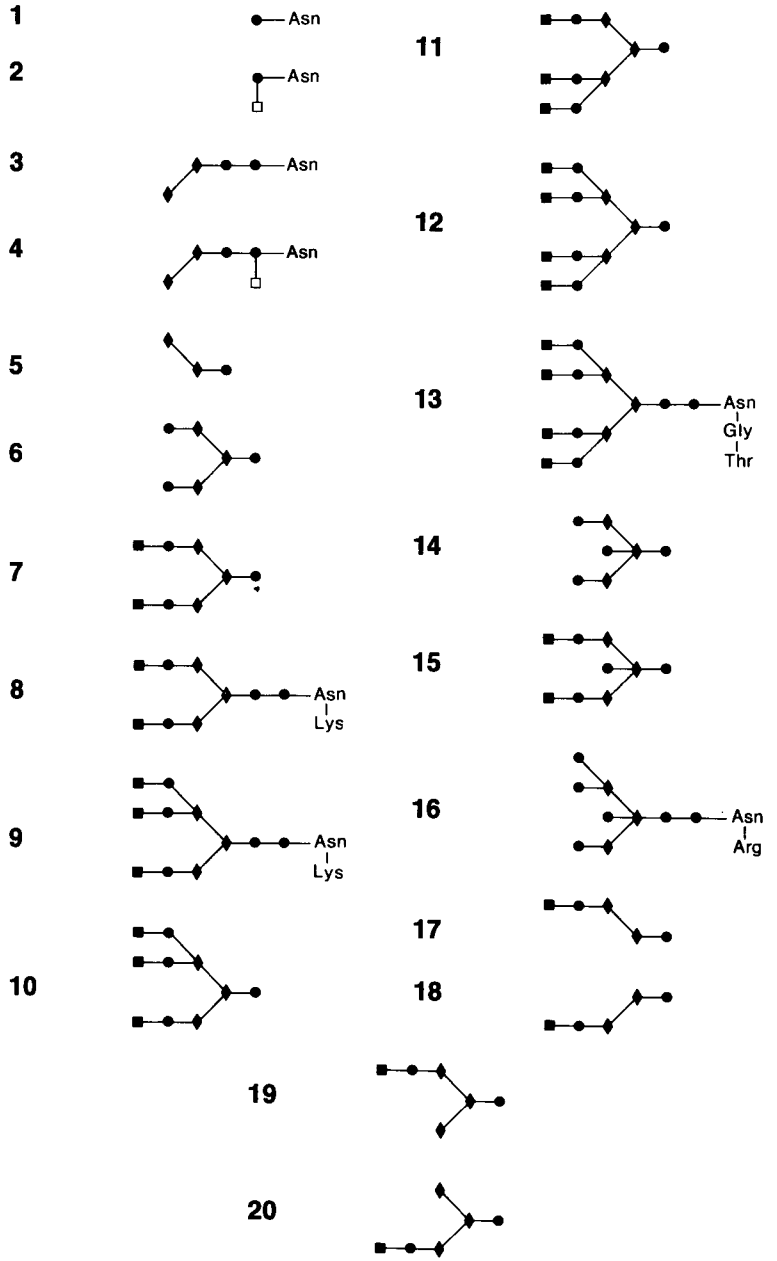
a. Fundamental Elements of Carbohydrate Chains of the *N*-Acetyllactosamine Type (Compounds 1–20).—Symbols employed for compounds **1–20** are depicted in Chart 1.

(50) J. P. Carver, A. A. Grey, F. M. Winnik, J. Hakimi, C. Ceccarini, and P. H. Atkinson, *Biochemistry*, 20 (1981) 6600–6606.

CHART 1. **Symbols Employed for Compounds 1–20**

Key to the symbolic notation:

● = GlcNAc ○ = α-NeuAc-(2→6)
◆ = Man △ = α-NeuAc-(2→3)
■ = Gal □ = Fuc



The simplest element of the linkage region between an *N*-glycosylic carbohydrate chain and the polypeptide backbone to which it is attached is βGlcNAc(1→N)Asn (compound **1**). Among various sources, **1** may be isolated from the urine of patients with aspartylglucosaminuria.[51–53] The 360-MHz, ^{1}H-n.m.r. spectrum of compound **1** is presented in Fig. 1. The resonances in this spectrum may be divided into signals from the following groups of protons.

The structural-reporter groups resonate at clearly distinguishable positions and provide essential information on the primary structure. The GlcNAc anomeric proton resonates at a lower field (see Table I) than could be expected [the normal range of δ for an axial anomeric proton in a free or *O*-glycosylically linked GlcNAc residue in the $^4C_1$(D) conformation is 4.40–4.75 p.p.m.]. It is reasonable to assume that this effect is due to the electron-withdrawing amide group attached to C-1 (compare with Ref. 54). The $J_{1,2}$ value (9.8 Hz) is characteristic for a β-glycosylic linkage. The value is relatively large, due to the *N*-type of glycosylic linkage. The chemical shift of the singlet of the *N*-acetyl group is typical for GlcNAc linked to Asn. As will be shown later, its position may vary, depending on the nature of the peptide moiety.

The nonanomeric protons resonate in the range of 3.4–4.0 p.p.m. For this simple compound, a complete interpretation was achieved, and this was confirmed, and refined, by computer simulation of the spectrum. The n.m.r. data for compound **1** are summarized in Table I. By means of a modified Karplus equation, the $^4C_1$(D) conformation of the GlcNAc group was deduced.[55]

The amino acid protons of Asn resonate apart from the GlcNAc protons.

In glycoproteins, substitution at O-6 of GlcNAc-1 by an α-L-Fuc group frequently occurs.* Compound **2** can be isolated from the urine of patients suffering from fucosidosis.[52,55,56] The 360-MHz, ^{1}H-n.m.r. spectrum of **2** is given in Fig. 2.

(51) J. N. Isenberg, in E. F. Walborg, Jr. (Ed.), *Glycoproteins and Glycolipids in Disease Processes*, ACS Symp. Ser. 80, American Chemical Society, Washington D. C., 1978, pp. 129–131, and references cited therein.
(52) G. Strecker and J. Montreuil, *Biochimie*, 61 (1979) 1199–1246.
(53) In the description of the n.m.r.-spectral features of each compound discussed in this article, only the actual source(s) of the compound used for the n.m.r. investigations is (are) mentioned. For other possible glycoprotein sources of these carbohydrates, the reader is referred to reviews[1–4]; see also, Section IV.
(54) M. Tanaka and I. Yamashina, *Carbohydr. Res.*, 27 (1973) 175–183.
(55) L. Dorland, B. L. Schut, J. F. G. Vliegenthart, G. Strecker, B. Fournet, G. Spik, and J. Montreuil, *Eur. J. Biochem.*, 73 (1977) 93–97.
(56) G. Strecker, B. Fournet, J. Montreuil, L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, and D. Dubesset, *Biochimie*, 60 (1978) 725–734.

TABLE I

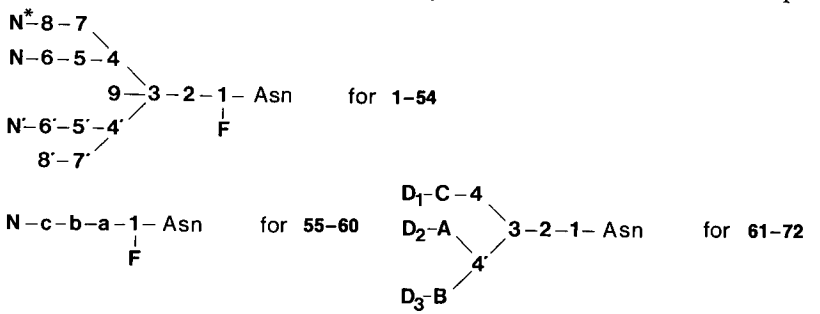
^{1}H Chemical Shifts and Coupling Constants for βGlcNAc(1→N)Asn (Compound 1) and αFuc(1→6)βGlcNAc(1→N)Asn (Compound 2)

		Chemical shift (p.p.m.)[a]			Coupling constant (Hz)[a]	
Residue	**Proton**	**Compound 1**	**Compound 2**		**Compound 1**	**Compound 2**
GlcNAc-1	H-1	5.072	5.095	$J_{1,2}$	9.80	9.85
	H-2	3.821	3.842	$J_{2,3}$	10.60	9.85
	H-3	3.601	3.614	$J_{3,4}$	9.10	9.50
	H-4	3.472	3.519	$J_{4,5}$	10.40	9.50
	H-5	3.527	3.673	$J_{5,6a}$	2.40	2.10
	H-6a	3.876	3.981	$J_{5,6b}$	5.55	5.55
	H-6b	3.739	3.725	$J_{6a,6b}$	−12.70	−11.40
	NAc	2.013	2.023			
αFuc(1→6)	H-1	—	4.901	$J_{1,2}$	—	3.75
	H-2	—	3.770	$J_{2,3}$	—	10.30
	H-3	—	3.889	$J_{3,4}$	—	3.40
	H-4	—	3.805	$J_{4,5}$	—	0.60
	H-5	—	4.117	$J_{5,6}$	—	6.55
	CH_3	—	1.214			
Asn	H-α	3.968	3.981	$J_{\alpha,\beta}$	6.90	6.30
	H-β	2.866	2.881	$J_{\alpha,\beta'}$	4.00	3.80
	H-β′	2.932	2.939	$J_{\beta,\beta'}$	−17.80	−17.25

[a] Chemical shifts and coupling constants measured at 360 MHz and T = 300 K.

Comparison of the spectral data for **1** and **2** shows that the chemical shifts of the GlcNAc structural-reporter groups, that is, H-1 and the *N*-acetyl-CH_3-protons, are slightly, but significantly, influenced, whereas the $J_{1,2}$ value is but little affected by the extension of **1** with a Fuc group. For the Fuc group, the resonances of its H-1, H-5, and CH_3-group protons are characteristic. The $J_{1,2}$ value (3.75 Hz) of Fuc is

* Coding of monosaccharides for the carbohydrate chains described in this Chapter:



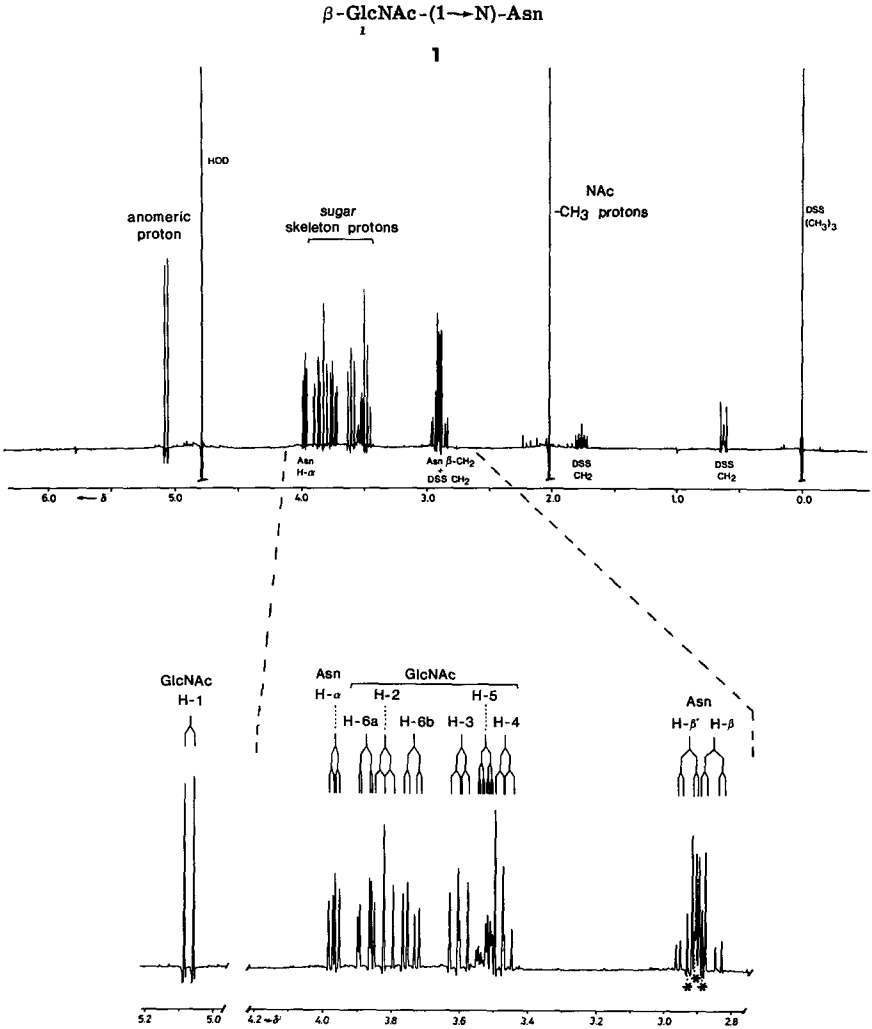


FIG. 1.—Resolution-enhanced, Overall, 360-MHz, ^{1}H-N.m.r. Spectrum of Compound **1** (Upper Trace), Supplied with Assignments in Full Detail, Most of Which are Indicated in the Expanded Regions (Lower Trace). [A small amount of sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) had been added to the D_2O solution, in order to serve as an internal reference for the chemical shifts ($\delta(CH_3)_3$, DSS ≡ 0). (In the expanded pattern for the Asn β-CH_2 protons, the signals marked by asterisks originate from DSS.)]

indicative of an α-L-glycosidic bond between this group and the GlcNAc residue.

The signals of the nonanomeric protons form a bulk that could be completely interpreted. The refined n.m.r. data, obtained after spectral simulation, are given in Table I. For both GlcNAc and Fuc, the normal chair conformations [$^4C_1$(D) and $^1C_4$(L), respectively] were cal-

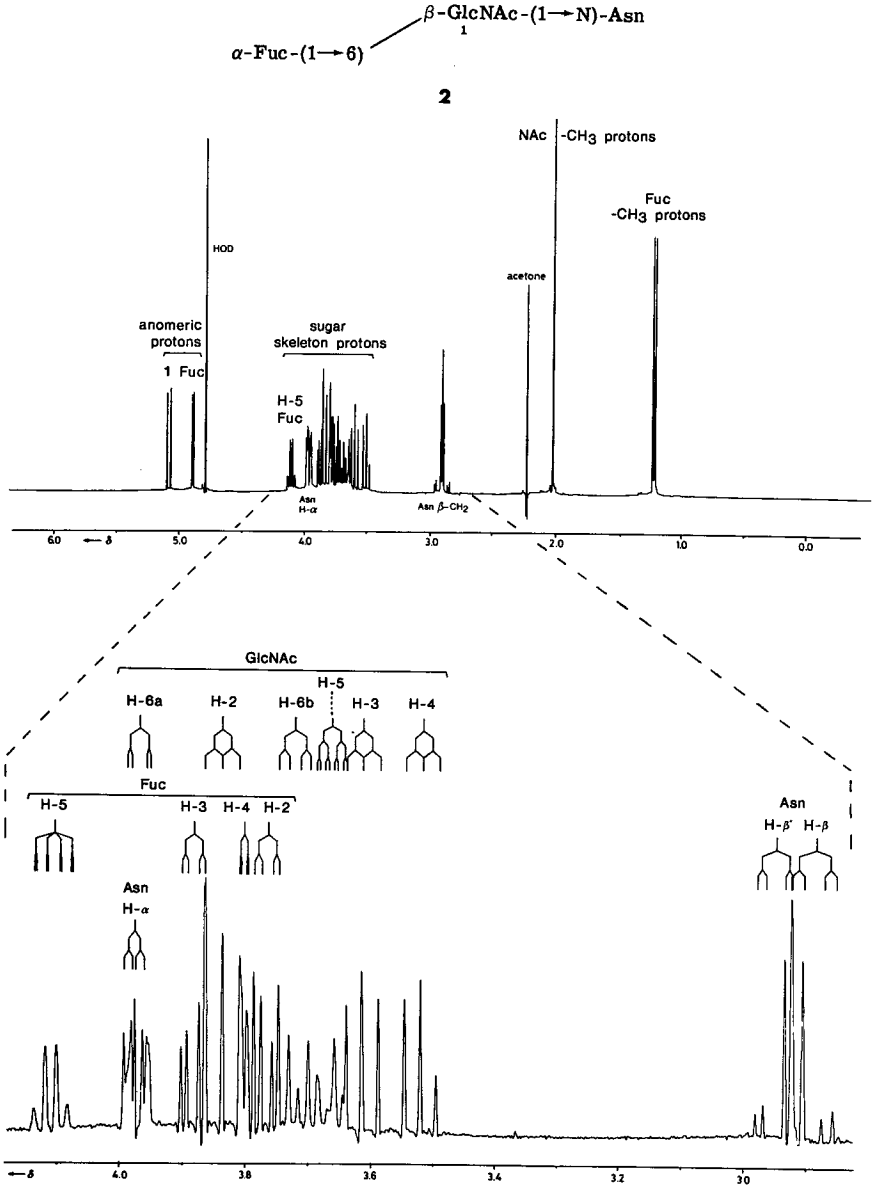


FIG. 2.—Resolution-enhanced, Overall, 360-MHz, ^{1}H-N.m.r. Spectrum of Compound **2** (Upper Trace), Supplied with Assignments in Full Detail, Most of Which are Indicated in the Expanded Region (Lower Trace). [Here, and throughout this article, chemical shifts are measured by reference to internal acetone (δ 2.225, relative to DSS in D_2O at 300 K).]

culated.[55] The attachment of Fuc to O-6 of GlcNAc influences the chemical shifts of the GlcNAc H-4, -5, and -6 atoms, in comparison to those of **1**. The change in the geminal $J_{6,6'}$ value (from −12.7 to −11.4 Hz) can be conceived as a strong indication for a (1→6)-glycosidic linkage. Unambiguous proof for the (1→6) linkage was obtained by ^{13}C-n.m.r. spectroscopy[57] of compound **2**.

(57) The type of linkage between Fuc and GlcNAc was unambiguously determined by ^{13}C-n.m.r. spectroscopy. The ^{13}C-n.m.r.-spectral data for compounds **1** and **2**, acquired at 25 MHz and 33°, are as follows.

^{13}C Chemical Shifts of Constituent Monosaccharides for βGlcNAc(1→N)Asn (Compound 1) and αFuc(1→6)βGlcNAc(1→N)Asn (Compound 2) (in p.p.m. relative to internal DSS)

		Schematic structure of compound		
		1		**2**
		•—Asn		•—Asn
Residue	**Carbon atom**		**^{13}C Chemical shift**	
GlcNAc	C-1	80.85		80.86
	C-2	56.82		56.76
	C-3	76.95		76.86
	C-4	72.25		72.49
	C-5	80.26		79.48
	C-6	63.25		69.93
	—C¹—NH—C(=O)*—CH₂—	175.60[a]		175.66[b]
	—C²—NH—C(=O)*—CH₃	177.57[a]		177.65[b]
	—C²—NH—C(=O)—CH₃*	24.80		24.80
Fuc	C-1	—		101.82
	C-2	—		70.93
	C-3	—		72.25
	C-4	—		74.54
	C-5	—		69.44
	C-6	—		18.00
Asn	C-α	53.68		53.64
	C-β	37.68		37.74
	-CO_2H	175.25[a]		175.27[b]

[a,b] Assignments may have to be interchanged. The increment in chemical shift observed for C-6 of GlcNAc in the step from **1** to **2** (Δδ +6.7 p.p.m.) indicates attachment of a Fuc group at O-6 of GlcNAc. This shift increment is in accord with literature data [P. Colson and R. R. King, *Carbohydr. Res.*, 47 (1976) 1–13]. For ^{13}C-n.m.r.-spectral data on βGlcNAc(1→N)Asn, see also, K. Dill and A. Allerhand, *FEBS Lett.*, 107 (1979) 26–29, and S. Shibata and H. Nakanishi, *Carbohydr. Res.*, 86 (1980) 316–320.

Differences in the chemical shifts of the methylene protons of Asn in the spectrum of **2** in comparison to that of **1** are due to a small variation in the pD for both solutions.

Compound **3** has been isolated from the urine of patients with aspartylglucosaminuria[51,52] or with Gaucher's disease.[58,59] The 500-

(58) H. van Halbeek, L. Dorland, G. A. Veldink, J. F. G. Vliegenthart, J.-C. Michalski, J. Montreuil, G. Strecker, and W. E. Hull, *FEBS Lett.*, 121 (1980) 65–70.

(59) J.-C. Michalski, J. Montreuil, G. Strecker, H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, B. Cartigny, and J.-P. Farriaux, *Eur. J. Biochem.*, (1983) in press.

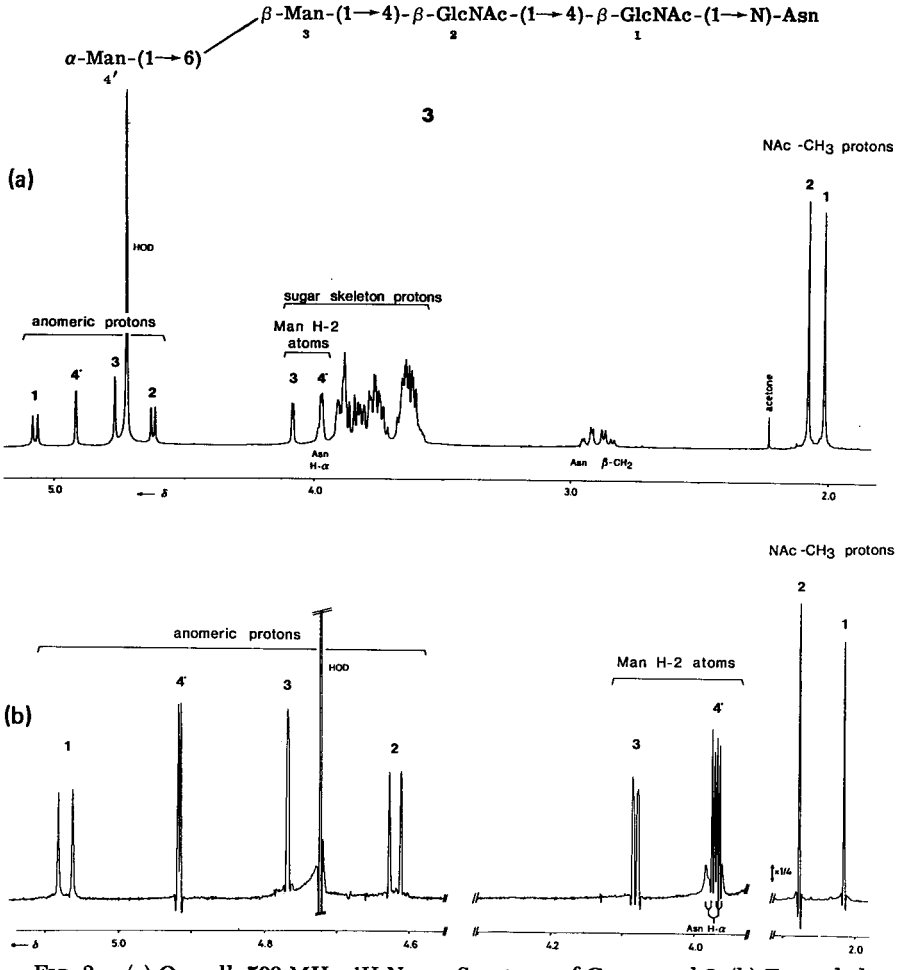


FIG. 3.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **3**. (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **3**. [The bold numbers in the spectra refer to the corresponding residues in the structure. The relative-intensity scale of the expanded, *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

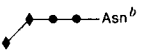
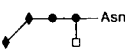
MHz, ^{1}H-n.m.r. spectrum of **3** obtained from the latter source is given in Fig. 3, and the spectral data are presented in Table II.

The substitution at O-4 of GlcNAc-1 by GlcNAc-2 does not significantly affect the chemical shift of its H-1 signal, as compared to **1**. The H-1 atom of GlcNAc-2 resonates at δ 4.618; the $J_{1,2}$ value (8.0 Hz) points to a β-glycosidic linkage between the two GlcNAc residues.

The different configurations of the glycosidic bonds of Man-3 and Man-4′ are reflected in the chemical shifts of their anomeric protons: δ 4.767 (Man-3: β) and δ 4.915 (Man-4′: α). The signal of H-1 of Man-3 appears in the spectrum as a relatively broad singlet due to the small value of $J_{1,2}$ (~0.7 Hz; H-1*a*–H-2*e*). In contrast, the resonance signal of the H-1 atom of Man-4′ is a well resolved doublet, with $J_{1,2}$ 1.8 Hz (H-1*e*–H-2*e*). This difference in coupling constants is even more clearly observable in the H-2 signals of the respective Man residues.

TABLE II

^{1}H Chemical Shifts[a] of Structural-reporter Groups of Constituent Monosaccharides, and of Amino Acid Protons for Glyco-asparagines 3 and 4

		Compound and schematic structure	
		3 	**4** 
Reporter group	**Residue**		
H-1 of	1	5.071	5.076
	2	4.618	4.690
	3	4.767	4.770
	4′	4.915	4.916
H-2 of	3	4.080	4.083
	4′	3.968	3.967
NAc of	1	2.014	2.018
	2	2.076	2.095
H-1 of	αFuc(1→6)	—	4.877
H-5 of	αFuc(1→6)	—	4.125
CH_3 of	αFuc(1→6)	—	1.209
H-α of	Asn	3.971	3.987
H-β of	Asn	2.851	2.871
H-β′ of	Asn	2.931	2.934

[a] N.m.r. data listed in this Table and the other Tables were acquired at 500 MHz, unless indicated otherwise. [b] In this heading and in those of the other Tables, the compounds involved are represented by the following, symbolic notation: ●, GlcNAc; ■, Gal; ◆, Man; ○, αNeuAc(2→6); △, αNeuAc(2→3); and □, Fuc. See also, Chart 1.

The H-2 atom of Man-3 resonates at δ 4.080 [typical for mono-α-(1→6) substitution, see later], and that of Man-4′ at δ 3.968 [typical for a terminal α-(1→6)-linked Man residue, see later]. For both residues, the $J_{2,3}$ value is 3.4 Hz.

The singlet of the *N*-acetyl protons of GlcNAc-1 is found at essentially the same position as in the spectrum of **1** (δ 2.014). The signal at δ 2.076 stems from the corresponding group of GlcNAc-2.

The signals of the remaining, sugar-skeleton protons are found in a very narrow spectral-region (3.55–3.95 p.p.m.), and have not yet been completely assigned.

Compound **4** may be isolated from the urine of patients with fucosidosis.[56] The 500-MHz, ^{1}H-n.m.r. spectrum of **4** is presented in Fig. 4. The spectral data for **4** are compiled in Table II, to enable comparison with those of its afuco analog **3**.

The chemical shift of the anomeric-proton signal of Fuc (δ 4.877) deviates considerably from that in the spectrum of **2**, reflecting the 4,6-disubstitution of GlcNAc-1. However, the H-5 and CH_3-group signals of this Fuc group occupy essentially the same positions as for **2** (see Table I). Therefore, this set of chemical-shift values of H-1, H-5, and the CH_3-group protons may be considered to be characteristic for a Fuc group α-(1→6)-linked to the β-GlcNAc-1 residue (see later).

In turn, the attachment of Fuc to O-6 of GlcNAc-1 has an interesting influence on the chemical shifts of the H-1 (Δδ 0.072 p.p.m.) and the *N*-acetyl protons (Δδ 0.019 p.p.m.) of GlcNAc-2 as compared to **3**.

The spectral region wherein the sugar-skeleton protons are found is unaltered in comparison to that for **3**. The presence of Fuc in **4** makes the bulk more complex.

The methylene resonances of Asn are found at δ 2.934 and 2.871, whereas H-α of the amino acid is observable at δ 3.987, probably due to a small difference in the pD of the solutions of **3** and **4**.

Compound **5** can be isolated from the urine of patients with mannosidosis.[52,60,61] The 500-MHz, ^{1}H-n.m.r. spectrum of the reducing trisaccharide is given in Fig. 5, and the spectral parameters are listed in Table III.

The observed spectrum of **5** is a superposition of the spectra of the two anomeric forms of this trisaccharide, occurring, under the n.m.r. measuring-conditions applied, in the ratio of $\alpha:\beta = \sim 2:1$.

The effect of anomerization is clearly recognizable from the structural-reporter-group signals of the three constituent monosaccharides.

(60) L. Dorland, J. Haverkamp, B. L. Schut, J. F. G. Vliegenthart, G. Spik, G. Strecker, B. Fournet, and J. Montreuil, *FEBS Lett.*, 77 (1977) 15–20.

(61) H. van Halbeek, L. Dorland, G. A. Veldink, J. F. G. Vliegenthart, G. Strecker, J.-C. Michalski, J. Montreuil, and W. E. Hull, *FEBS Lett.*, 121 (1980) 71–77.

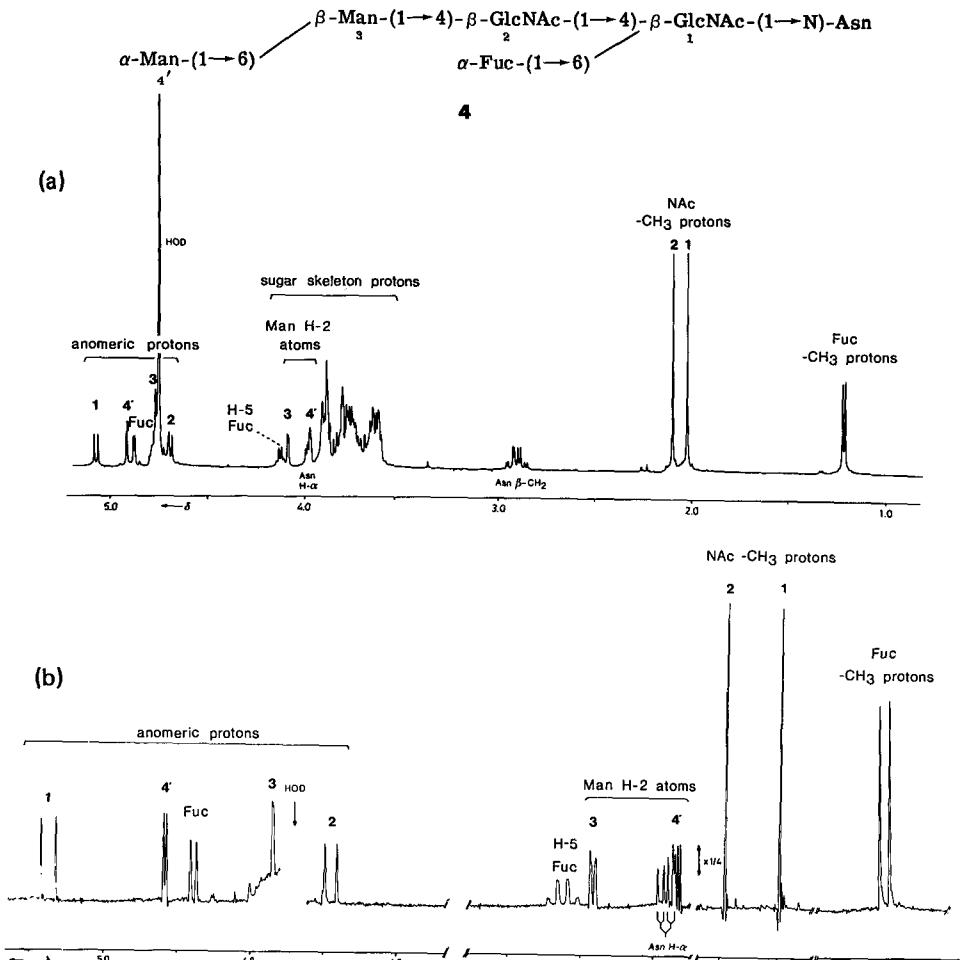


FIG. 4.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound 4. (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound 4. [The bold numbers in the spectra refer to the corresponding residues in the structure. The relative-intensity scale of the expanded, *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the expanded spectrum; its position is indicated by an arrow.]

The H-1 signals of the α and β anomer of GlcNAc-2 are found at δ 5.209 and 4.718, respectively. The intensity of these signals reflects the anomeric ratio. The same ratio is found for the signals of the *N*-acetyl group of GlcNAc-2, at δ 2.043 (α anomer) and 2.041 (β anomer), respectively.

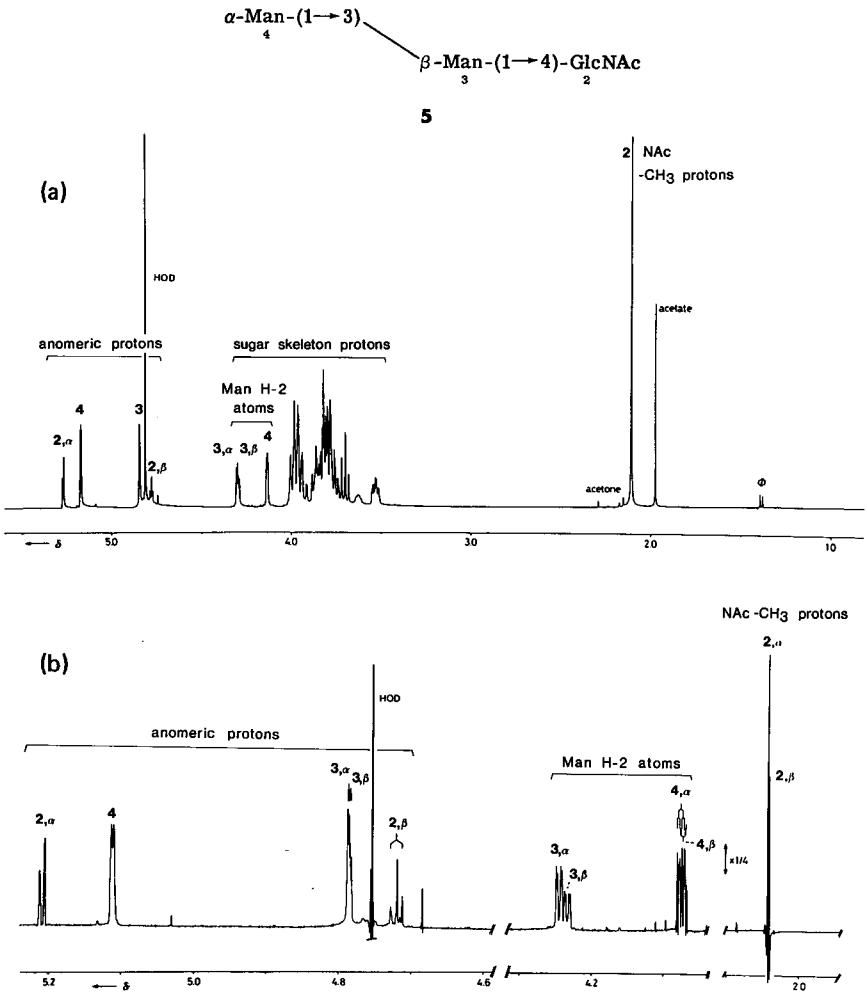


FIG. 5.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **5**. (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **5**. [The bold numbers in the spectra refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **5**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the expanded, *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

The H-1 atom of Man-3 resonates at δ 4.787 for the α anomer of **5**, whereas the chemical shift of this H-1 atom for the β anomer of **5** is 4.783 p.p.m. The latter value deviates from those of the corresponding H-1 atoms of the glyco-asparagines **3** and **4** (see Table II) (GlcNAc-2 β-linked to GlcNAc-1). This effect is due to the α-(1→3) substitution of

TABLE III

^{1}H-N.m.r. Data for Structural-reporter Groups of Oligosaccharide 5, αMan(1→3)βMan(1→4)GlcNAc
4 3 2

Reporter group	Residue	Anomer of compound	Schematic structure: Chemical shift (p.p.m.)	Schematic structure: $J_{1,2}$ (Hz)
H-1 of	2	α	5.209	3.5
		β	4.718	8.1
	3	α	4.787	0.9
		β	4.783	1.0
	4	α,β[a]	5.111	1.8
H-2 of	3	α	4.244	0.9
		β	4.233	1.0
	4	α	4.075	1.8
		β	4.071	1.7
NAc of	2	α	2.043	—
		β	2.041	—

[a] If, for a certain proton, $|\Delta\delta_{\alpha-\beta}| \leq 0.001$ p.p.m., only an average δ value has been listed in this and other Tables in this article, as such a difference in chemical shift due to anomerization, if any, is not detectable at 500 MHz, unless the lines of the signal are extremely sharp.

Man-3 (see later). The doublets for this H-1 for both the α and β form of **5** are well resolved, reflecting the relatively large $J_{1,2}$ value (0.95 Hz) for this compound. For the H-2 atom of Man-3, this large $J_{1,2}$ value leads to two doublets of doublets, at δ 4.244 (α anomer of **5**) and at δ 4.233 (β anomer of **5**), respectively, instead of to two doublets having relatively broad lines. Nevertheless, the pattern of the H-2 resonance of an α-linked Man residue is clearly distinguishable from that of a β-linked one.

The signal at δ 5.111 is ascribed to H-1 of Man-4. The value of its coupling constant ($J_{1,2}$ 1.8 Hz) indicates an α-linkage between the two Man residues. The influence of anomerization on the chemical shift of H-2 of the nonreducing-end sugar, Man-4 (δ 4.075 and 4.071 for the α and β anomer, respectively) is apparent. The chemical shifts of H-1 and H-2 of Man-4 in the β anomer of **5** differ from those of H-1 and H-2 of Man-4′ in **3** and **4**. Obviously, the type of linkage of the terminal Man residues [α-(1→3) or α-(1→6), respectively] has a large effect upon the chemical shifts of their H-1 and H-2 atoms. Especially, the chemical shift of H-2 of Man-3 reflects the position of substitution of 3

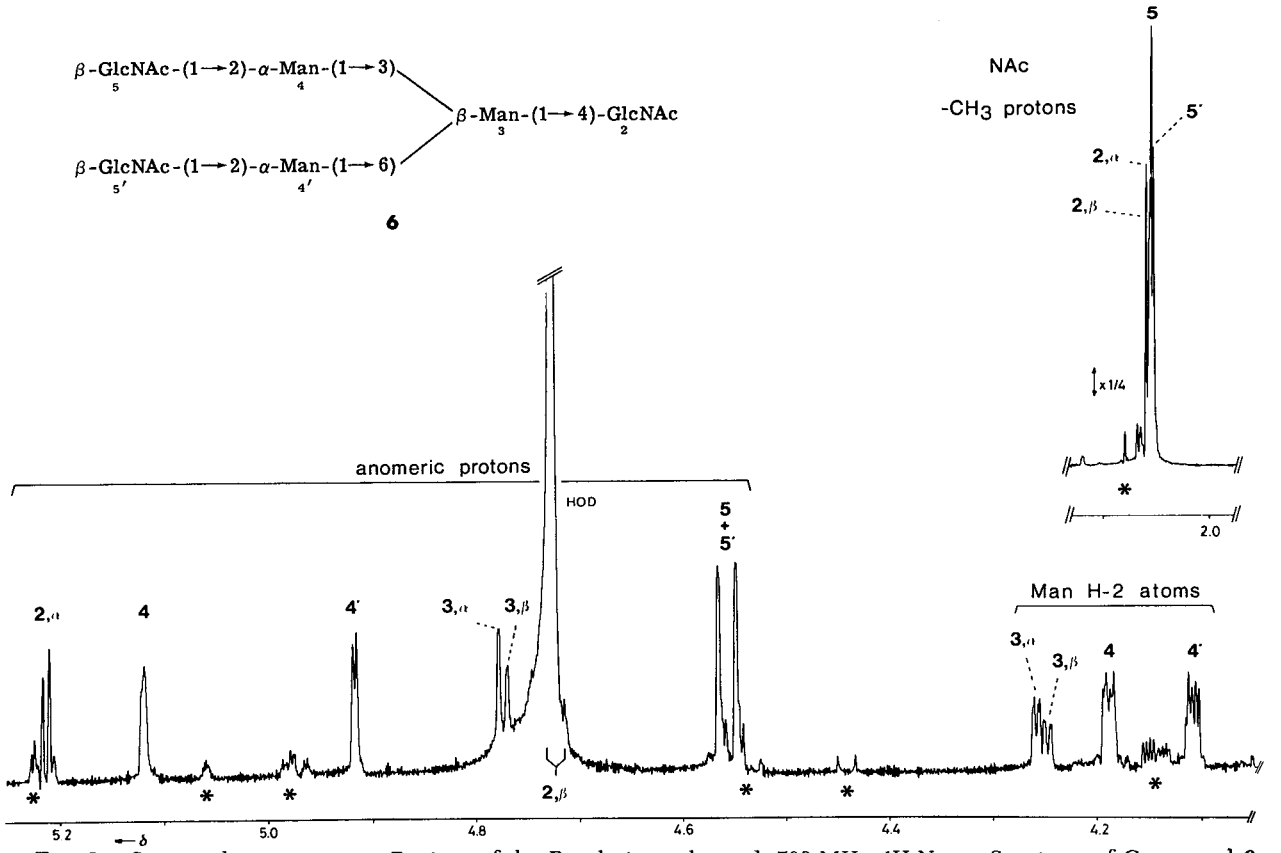
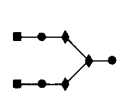
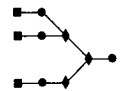
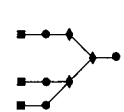
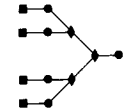


FIG. 6.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **6**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **6**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. In addition to **6**, the sample contained a small proportion of oligosaccharide **14**, as can be inferred from the signals marked by asterisks (compare Fig. 15).]

TABLE IV

^{1}H Chemical Shifts of Structural- reporter Groups of Constituent Monosaccharides for Asialo-agalacto Di-, and Asialo Di-, Tri-, Tri'-, and Tetra-antennary Oligosaccharides (Compounds 6, 7, and 10–12)

			Compound and schematic structure				
			6	7	10	11	12
							
Reporter group	**Residue**	**Anomer of compound**					
H-1 of	2	α	5.213	5.212	5.211	5.206	5.205
		β	4.724	4.721	4.723	4.721	4.719
	3	α	4.779	4.775	4.767	4.776	4.768
		β	4.770	4.765	4.757	4.770	4.761
	4	α	5.119	5.123	5.118	5.133	5.131
		β		5.121		5.131	5.129
	4′	α	4.919	4.927	4.923	4.874	4.870
		β		4.930	4.925		
	5	α,β	4.558	4.582	4.569	4.585	4.572
	5′	α,β	4.558	4.582	4.584	4.592	4.592

	6	α,β	—	4.468	4.468	4.469	4.469
	6′	α	—	4.471	4.471	4.469	4.469
		β		4.473	4.473	4.471	4.471
	7	α,β	—	—	4.546	—	4.547
	7′	α,β	—	—	—	4.555	4.554
	8	α,β	—	—	4.463	—	4.464
	8′	α,β	—	—	—	4.481	4.480
H-2 of	3	α	4.257	4.259	4.223	4.263	4.214
		β	4.247	4.249	4.212	4.252	4.203
	4	α	4.189	4.193	4.218	4.201	4.228
		β				4.198	4.224
	4′	α,β	4.108	4.114	4.114	4.099	4.099
H-3 of	4	α	<4.0	<4.0	4.050	<4.0	4.056
		β			4.048		4.053
	4′	α,β	<4.0	<4.0	<4.0	<4.0	<4.0
NAc of	2	α	2.059	2.060	2.060	2.060	2.059
		β	2.055	2.057	2.057	2.057	2.056
	5	α,β	2.054	2.052	2.050	2.057	2.054
	5′	α	2.052	2.048	2.048	2.043	2.041
		β		2.046	2.046	2.046	2.045
	7	α,β	—	—	2.078	—	2.079
	7′	α	—	—	—	2.039	2.038
		β				2.040	2.040

in the case of mono-substitution by another Man residue: δ ~4.24 for **5**, and δ ~4.08 for **3** and **4**.

Concerning the remaining skeleton-protons of the three residues of **5**, it is very probable that, due to anomerization, doubling of signals from protons other than those from structural-reporter groups occurs; however, this could not be traced in the bulk.

Compound **6** can be isolated[62] from the urine of patients with Sandhoff's disease (GM_2-gangliosidosis variant O). The 500-MHz, ^{1}H-n.m.r. spectrum of this reducing oligosaccharide is depicted in Fig. 6; its spectral parameters are compiled in Table IV. Compound **6** contains the 3,6-disubstituted Man-3, characteristic for diantennary structures. Each branch is terminated with a β-(1→2)-linked GlcNAc residue.

As described for compound **5**, subspectra of the two anomeric forms constitute the spectrum observed. Anomerization comes to expression in the structural-reporter-group signals of GlcNAc-2 and Man-3. The H-1 atoms of the two α-linked Man residues resonate at δ 5.119 (Man-4) and δ 4.919 (Man-4′), respectively. The H-1 signal of Man-4 shows broader lines than that of Man-4′; this may reflect the effect of anomerization, although other influences cannot be excluded. Comparison with compounds **3** and **5** reveals that the substitution at O-2 of Man-4 and -4′ by GlcNAc residues causes only small increments in shift for their anomeric protons. However, these substitutions give rise to significant shift-increments for the H-2 atoms. Furthermore, the δ values of H-1 and H-2 of Man-3 are significantly influenced by these substitutions (Δδ −0.008 and +0.014 p.p.m., respectively). As will be shown later, for compound **18**, the latter effects are attributable to the attachment of GlcNAc-5′ rather than of -5.

The signals of the H-1 atoms of GlcNAc-5 and -5′ are found at δ 4.558. The complexity of the latter signal(s) is probably due to differences in the chemical shifts of the H-1 atoms in the α and β anomer of **6**. The β-type of the linkages between GlcNAc and Man is evident from the $J_{1,2}$ value (8.6 Hz).

The *N*-acetyl signals of the amino sugars of **6** are well separated; for GlcNAc-2, two singlets are observed, at δ 2.059 and at 2.055, respectively corresponding to the α and β anomer of **6**. The difference in peak height of the other two *N*-acetyl signals does not reflect an intensity ratio of 2:1, due to anomerization. In fact, this difference results from a variation in line width. Therefore, each signal belongs to one

(62) G. Strecker, M.-C. Herlant-Peers, B. Fournet, J. Montreuil, L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, and J.-P. Farriaux, *Eur. J. Biochem.*, 81 (1977) 165–171.

GlcNAc residue. Their assignment is based on spectra of more-complex compounds (for example, **7**, **10**, **25**, and **26**).

Compound **7** is a so-called diantennary oligosaccharide; it can be isolated from the urine of patients suffering from GM_1-gangliosidosis,[52,63–64a] as well as from[64,65] Morquio syndrome type B. The 500-MHz, ^{1}H-n.m.r. spectrum of **7** is presented in Fig. 7, and the chemical shifts of the structural-reporter groups are summarized in Table IV.

As compared to the corresponding agalacto compound **6**, the presence of Gal residues 6 and 6′ introduces a change in the chemical shift of the H-1 atoms of the peripheral GlcNAc residues 5 and 5′ from δ 4.558 to 4.582. However, this extension has virtually no effect upon the chemical shifts of the H-1 atoms of Man residues 3, 4, and 4′. Interestingly, the H-1 atom of Man-4 and also that of Man-4′ each yields two separated doublets which must be ascribed to the two anomeric forms of **7**. From the relative intensities of these doublets, it may be concluded that the anomers of this octasaccharide occur in the ratio of $\alpha:\beta = 2:1$. Apparently, the chain length does not affect the anomeric-equilibrium constant.

The anomeric protons of Gal-6 and -6′ resonate at δ ~4.47. Three doublets are observed; $J_{1,2}$ of each of these has a value of 7.9 Hz, indicating β bonds between Gal and GlcNAc. The highest-field doublet, at δ 4.468, representing one proton, is attributed to Gal-6. This assignment is based on spectral data for monosialo diantennary structures (compounds **27** and **28**). Consequently, the two lower-field doublets, at δ 4.471 and 4.473, which together represent one proton, belong to H-1 of Gal-6′ in the α and β anomer of **7**, respectively.

Doubling of the signals, as observed for the H-1 atoms of Gal-6′ and Man-4 and Man-4′, also appears for the H-1 and H-2 atoms of Man-3 and for H-1 of GlcNAc-2. No doubling is detectable for the anomeric signal of Gal-6, nor for the H-1 signals of GlcNAc-5 and -5′. However, a possible doubling of the latter signals cannot be completely excluded due to the relatively broad lines.

The total pattern of the H-2 signals of Man-3, -4, and -4′ is essentially identical to that for **6**.

(63) E. G. Brunngraber, in Ref. 51, pp. 135–149.

(64) H. van Halbeek, L. Dorland, G. A. Veldink, J. F. G. Vliegenthart, G. Strecker, J.-C. Michalski, J. Montreuil, and W. E. Hull, *Eur. J. Biochem.*, (1983) submitted for publication.

(64a) K. Yamashita, T. Ohkura, S. Okada, H. Yabuuchi, and A. Kobata, *J. Biol. Chem.*, 256 (1981) 4789–4798; T. Ohkura, K. Yamashita, and A. Kobata, *ibid.*, 256 (1981) 8485–8490.

(65) J.-C. Michalski, G. Strecker, H. van Halbeek, L. Dorland, and J. F. G. Vliegenthart, *Carbohydr. Res.*, 100 (1982) 351–363.

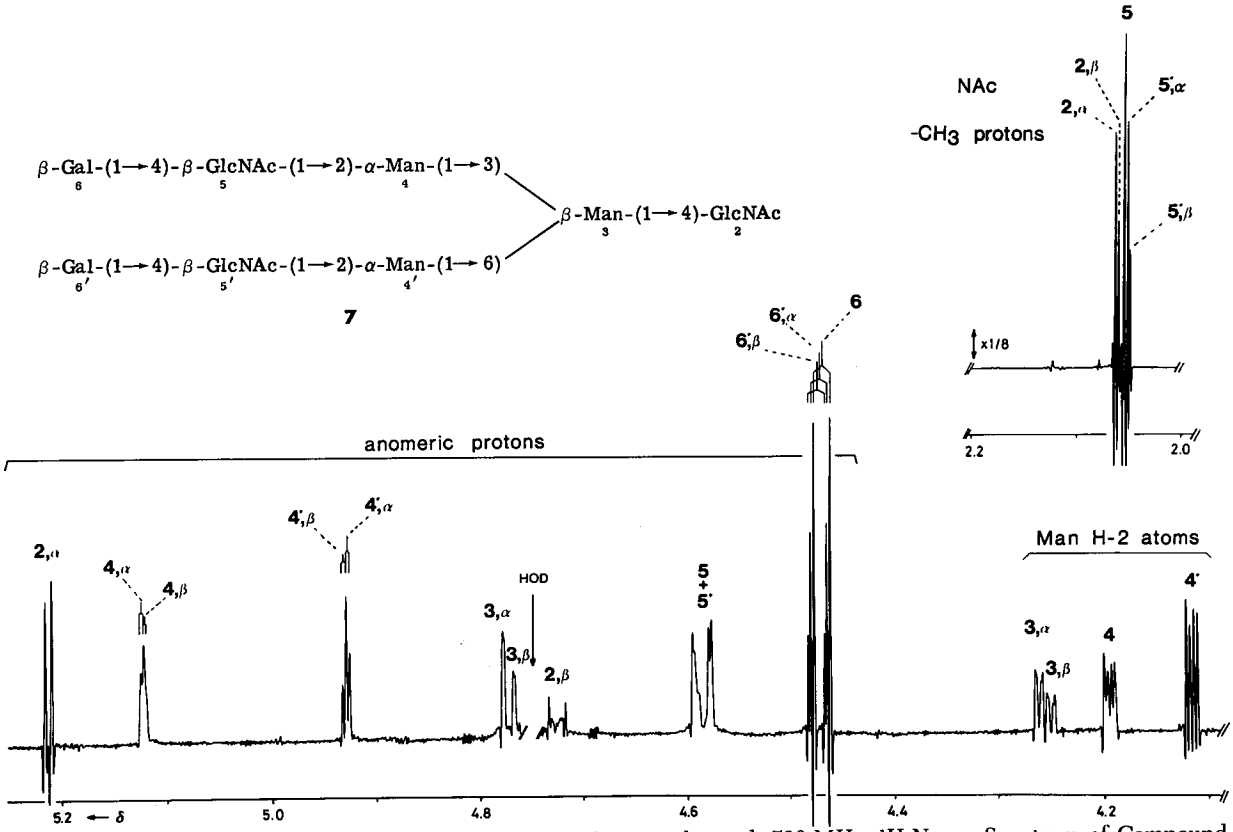


FIG. 7.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **7**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **7**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow.]

The *N*-acetyl signals of GlcNAc-5 and -5′ were assigned by comparison with compounds **27** and **28**; only the *N*-acetyl signal of GlcNAc-5′ undergoes doubling.

These results demonstrate that there is a large difference in the effect of anomerization on the upper- and the lower-branch signals; the effect is far more pronounced for the lower branch.

Compound **8** is a diantennary, asialo glycopeptide that can be isolated from various desialylated glycoproteins.[1–4] The 500-MHz, ^{1}H-n.m.r. spectrum of this nonreducing compound, obtained from α_1-acid glycoprotein,[66–70] is given in Fig. 8, and the n.m.r. parameters are listed in Table V. Comparison of the spectrum of **8** with that of **7** demonstrates that, for the glycopeptide, the spectrum is less complex, despite the additional residues.

The additional, H-1 signal of GlcNAc-1 is found at δ 5.094. This value is similar to the corresponding values for compounds **1–4**. Although the resonance position of the anomeric proton of GlcNAc-1 can vary considerably (5.05–5.10 p.p.m.),[67] due to differences in the pD of the glycopeptide solution and in the amino acid composition of the peptide moiety, the influence of the presence of Lys is only small. A similar phenomenon is observed for the *N*-acetyl signal of GlcNAc-1; its chemical shift may vary[67] from 2.004 to 2.014 p.p.m. Despite these variations, the signals can be used for identification purposes, because they occupy unique spectral-positions. The chemical shift of H-1 of GlcNAc-2 is not sensitive to the aforementioned influences; this proton resonates at δ 4.616, which is in perfect agreement with its position in the spectrum of glyco-asparagine **3**.

The substitution pattern of the Man residues 3, 4, and 4′ is reflected in the chemical shifts of their H-1 and H-2 atoms (compare, compounds **6** and **7**; see Table VI). The H-1 signals of GlcNAc-5 and -5′ coincide at δ 4.582, whereas the H-1 signals of Gal-6 and -6′ are distinct from each other. The assignment of the Gal H-1 doublets is based on the spectrum of a diantennary, monosialo glycopeptide (**31**), and is consistent with that for oligosaccharide **7**.

(66) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, K. Schmid, J. Montreuil, B. Fournet, and W. E. Hull, *FEBS Lett.*, 114 (1980) 11–16.
(67) B. Fournet, J. Montreuil, G. Strecker, L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, J. P. Binette, and K. Schmid, *Biochemistry*, 17 (1978) 5206–5214.
(68) L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, B. Fournet, G. Strecker, G. Spik, J. Montreuil, K. Schmid, and J. P. Binette, *FEBS Lett.*, 89 (1978) 149–152.
(69) K. Schmid, J. P. Binette, L. Dorland, J. F. G. Vliegenthart, B. Fournet, and J. Montreuil, *Biochim. Biophys. Acta*, 581 (1979) 356–359.
(70) L. Dorland and J. F. G. Vliegenthart, in R. Balian, M. Chabre, and P. F. Devaux (Eds.), *Membranes and Intercellular Communication* (Les Houches, 1979), North-Holland, Amsterdam, 1981, pp. 183–192.

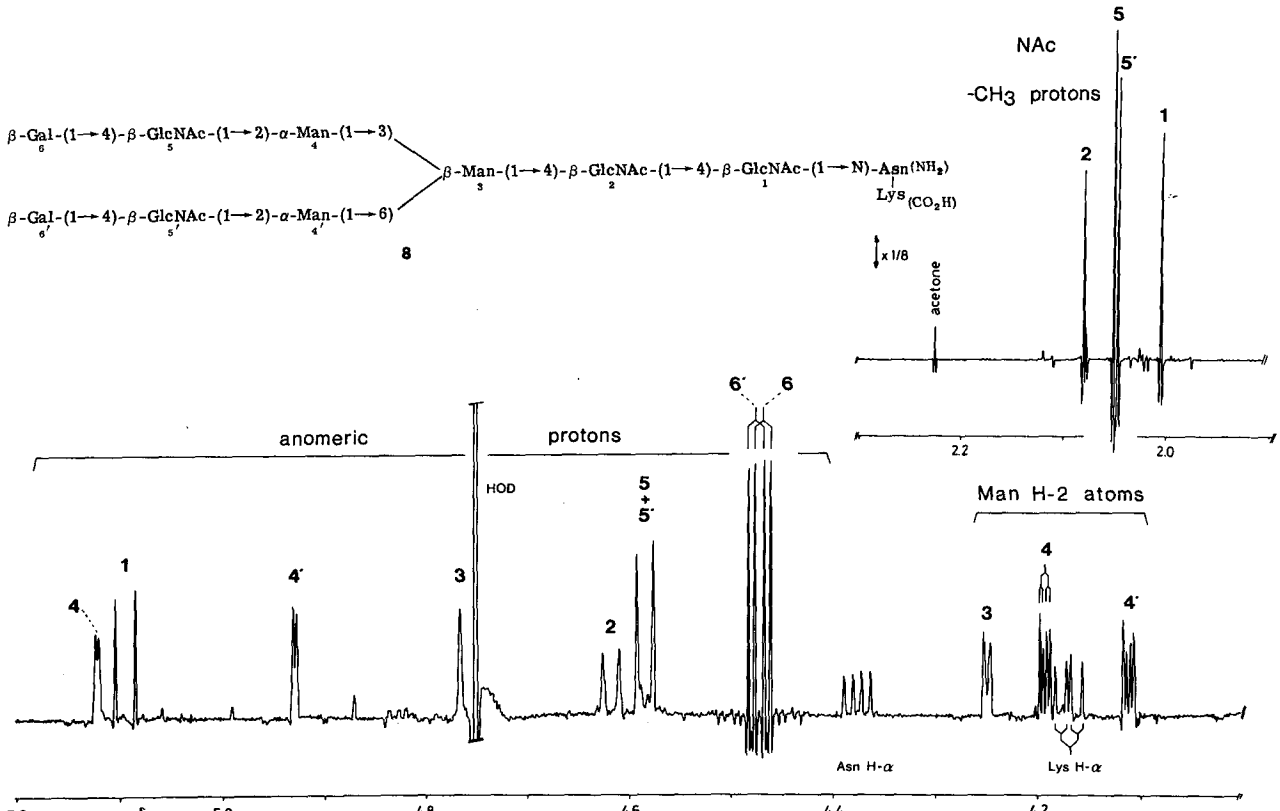
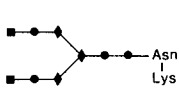
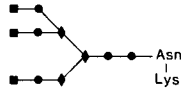
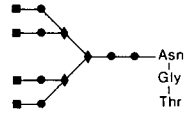


FIG. 8.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound 8. [The bold numbers in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated.]

TABLE V

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Asialo Di-, Tri-, and Tetra-antennary Glycopeptides of the *N*-Acetyllactosamine Type (Compounds 8, 9, and 13)

		Compound and schematic structure		
		8	9	13
				
Reporter group	**Residue**			
H-1 of	1	5.094	5.092	5.053
	2	4.616	4.614	4.614
	3	4.765	4.755	4.757
	4	5.121	5.120	5.129
	4′	4.928	4.924	4.868
	5	4.582	4.570	4.573
	5′	4.582	4.580	4.596
	6	4.467	4.468	4.470
	6′	4.473	4.473	4.472
	7	—	4.545	4.547
	7′	—	—	4.553
	8	—	4.462	4.465
	8′	—	—	4.481
H-2 of	3	4.246	4.209	4.210
	4	4.190	4.218	4.224
	4′	4.109	4.108	4.092
H-3 of	4	<4.0	4.045	4.052
	4′	<4.0	<4.0	<4.0
NAc of	1	2.004	2.003	2.008
	2	2.079	2.078	2.078
	5	2.050	2.048	2.054
	5′	2.046	2.045	2.042[a]
	7	—	2.075	2.079
	7′	—	—	2.041[a]

[a] Assignments may have to be interchanged.

The spectrum of **8** shows four, separated, *N*-acetyl signals, at δ 2.004 (GlcNAc-1, see earlier), 2.046 (GlcNAc-5′), 2.050 (GlcNAc-5), and 2.079 (GlcNAc-2), respectively. The assignments of the *N*-acetyl signals of **1** and **2** were made with the aid of the spectrum of compound **3**, and those of the signals of 5 and 5′ on the basis of specific shift-increments, introduced by the presence of sialic acid in the monosialo compound **31** (see also, Ref. 66). An interesting, spectral feature is the line width of the *N*-acetyl signals and of some anomeric-proton resonances

in relation to the position of the corresponding monosaccharide in the carbohydrate chain. The line widths of the H-1 doublet and the *N*-acetyl singlet of GlcNAc-1 vary upon changing the pD of the glycopeptide solution; consequently, this parameter is of only limited value for the identification of this residue. The doublet of H-1 and the *N*-acetyl signal of GlcNAc-2 show relatively broad lines in comparison to those of residues 5 and 5′, which is ascribed to the rigidity of the intact, core structure of trimannosyl-*N*,*N*′-diacetylchitobiose.

The line widths of the Gal H-1 signals are small compared to those of the H-1 doublet of the peripheral GlcNAc residues 5 and 5′. This observation is in accord with the relatively large mobility of terminal residues. The H-1 signals of GlcNAc-5 and -5′ in the spectra of compounds **7** and **8** show broader lines than those in the spectrum of compound **6** (see Fig. 6), wherein GlcNAc-5 and -5′ occupy terminal positions.

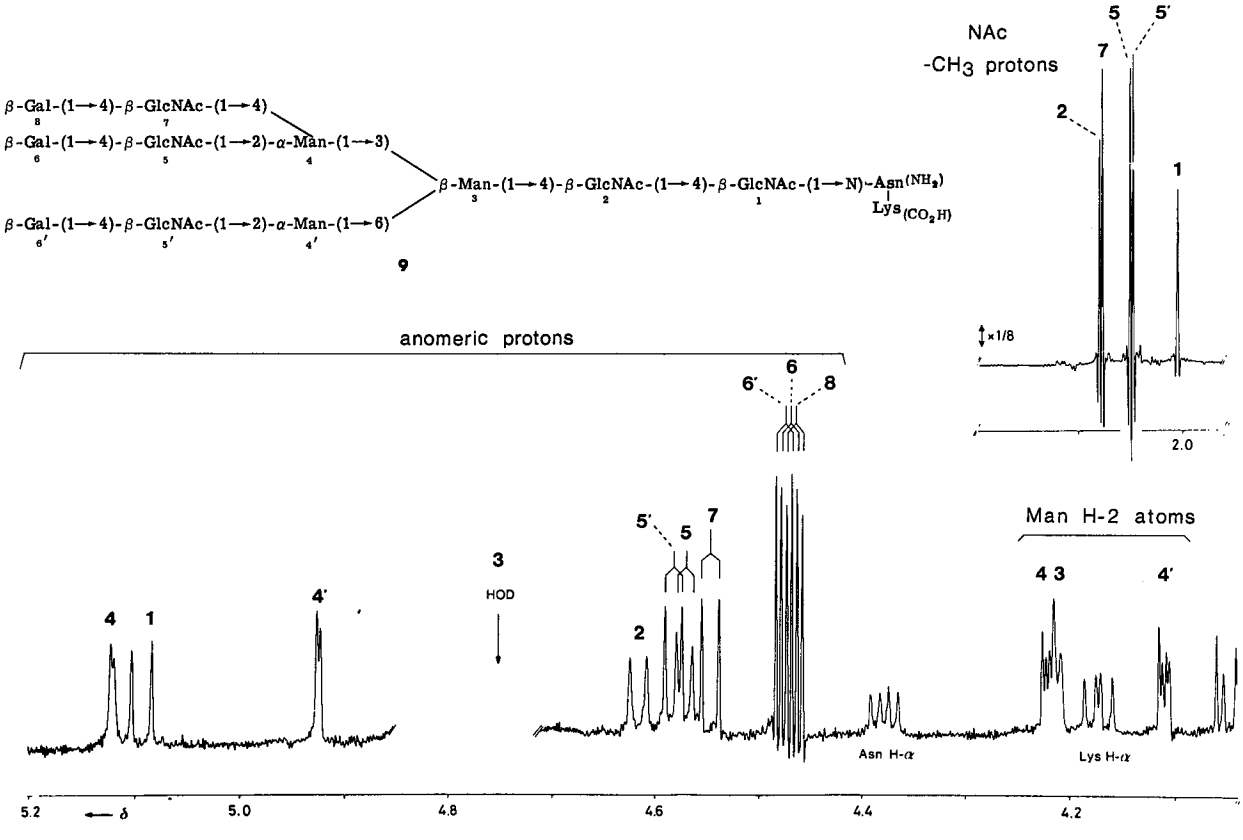
The somewhat narrower lines of the H-1 doublet of Man-4′ compared to that of 4 may indicate that the conformational freedom of Man-4′, occurring in α-(1→6) linkage, is larger than that of Man-4, which is present in an α-(1→3) linkage. This may ultimately lead to differences in mobility for the upper and lower branches.

Compound **9** is a triantennary, asialo glycopeptide that can be isolated from various desialylated glycoproteins.[1–4] The 500-MHz, ^{1}H-n.m.r. spectrum of **9**, derived from α_1-acid glycoprotein,[66–69] is given in Fig. 9, and the n.m.r. parameters are compiled in Table V. The spectral features of this glycopeptide bear a high similarity to those of compound **8** (see Table V).

The presence of the third *N*-acetyllactosamine unit gives rise to additional signals for the anomeric protons of GlcNAc-7 (at δ 4.545) and of Gal-8 (at δ 4.462), and to an extra *N*-acetyl signal, at δ 2.075, belonging to GlcNAc-7.

The presence of the additional branch causes some specific changes in the chemical shifts of the H-2 atoms of Man-3 and -4 in comparison to those of **8**. The H-2 signal of Man-3 undergoes an upfield shift (Δδ −0.037 p.p.m.), whereas that of H-2 of Man-4 shifts downfield (Δδ 0.028 p.p.m.). These effects are rather large, so that their relative posi-

FIG. 9.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **9**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3 (which, at 300 K, is almost hidden under this line), have been omitted from the spectrum; their position is indicated by an arrow.]


β-Gal-(1→4)-β-GlcNAc-(1→4)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn(NH2)
Lys(CO2H)
9
NAc
-CH3 protons
anomeric protons
Man H-2 atoms
HOD
×1/8
2.0
Asn H-α
Lys H-α
δ
5.2
5.0
4.8
4.6
4.4
4.2

tions are interchanged. The spectral resolution at 360 MHz is not sufficient to permit observation of these H-2 signals separately.[67] Furthermore, an upfield shift is observed for H-1 of Man-3 ($\Delta\delta$ −0.010 p.p.m.) as compared to **8**. The set of chemical shifts for the H-1 and H-2 atoms of the Man residues are characteristic for triantennary, asialo glycopeptides (see Table VI). Another interesting feature is the appearance of the H-3 signal of Man-4 at a position that is well resolved from the bulk of the skeleton protons.

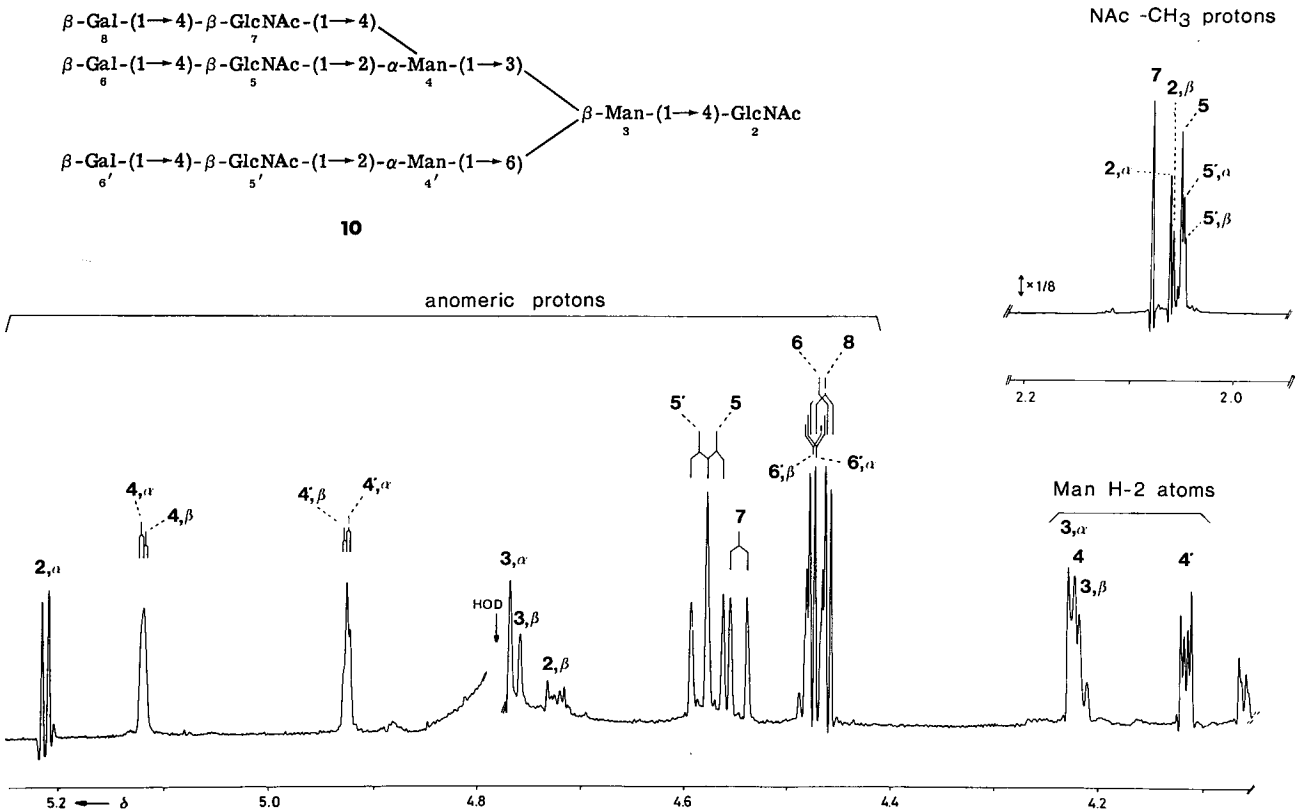
The H-1 signal of GlcNAc-5 shifts upfield ($\Delta\delta$ −0.012 p.p.m.) and undergoes a significant line-broadening upon substitution of Man-4 with *N*-acetyllactosamine in a β-(1→4) linkage. The chemical shifts and the line widths of the other GlcNAc anomeric signals (1, 2, and 5′) remain unchanged with respect to **8**.

Attachment of the third branch causes only slight differences in the chemical shifts of the *N*-acetyl signals of GlcNAc residues 1, 2, 5, and 5′ in comparison to **8**, but the small decrement in shift for GlcNAc-5 ($\Delta\delta$ −0.002 p.p.m.) is typical (see later). The assignment of the *N*-acetyl signal at δ 2.075 to GlcNAc-7 is in agreement with the line width of this singlet, because a peripheral residue will give rise to a sharper line than an internal residue (GlcNAc-2: δ 2.078).

Compound **10** is a triantennary, asialo oligosaccharide that has been isolated from the urine of a patient with GM_1-gangliosidosis.[52,64,64a] The 500-MHz, ^{1}H-n.m.r. spectrum of **10** is given in Fig. 10, and the n.m.r. parameters are summarized in Table IV. For spectral interpretation, **10** may be conceived of either as an extension of oligosaccharide **7**, with an *N*-acetyllactosamine unit linked to Man-4, or as compound **9** lacking the βGlcNAc(1→N)Asn-Lys moiety. The attachment of the third *N*-acetyllactosamine unit to Man-4 of the diantennary oligosaccharide **7** has the same influence upon the structural-reporter groups of its β anomer as was described for the extension of **8** to **9**.

A difference between the chemical shifts of H-1 of Man-3 for the α and β anomer of **10** is clearly visible. This effect is somewhat less pronounced for Man-4′ in comparison to the spectrum of **7** (see Fig. 7). The doublets of the H-1 atoms of GlcNAc-5, -5′, and -7 are well separated, showing a total pattern similar to that observed for the corre-

FIG. 10.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **10**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **10**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow.]


β-Gal-(1→4)-β-GlcNAc-(1→4)
8 7
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'
10
NAc -CH3 protons
7
2,β
5
2,α
5',α
5',β
×1/8
2.2
2.0
anomeric protons
6
8
5'
5
6',β
6',α
4,α
4,β
4',β
4',α
3,α
HOD
3,β
2,β
7
2,α
Man H-2 atoms
3,α
4
4'
3,β
5.2 ← δ
5.0
4.8
4.6
4.4
4.2

sponding signals of **9**. Among the anomeric protons of the Gal residues 6, 6′, and 8, only H-1 of Gal-6′ gives rise to two doublets (due to anomerization).

The substitution of the mannotriose branching core by three *N*-acetyllactosamine units can be recognized on the basis of the chemical shifts of the H-2 signals of the three Man residues (see Table VI). Owing to anomerization, two signals are observed for H-2 of Man-3. The complexity of the H-2 signals of Man-3 and -4, in comparison to **9**, is caused by partial overlap of the H-2 signal of Man-3 of the α anomer with the H-2 signal of Man-4 of both anomers of **10**.

As discussed for **9**, H-3 of Man-4 can be conceived of as being a structural-reporter group typical for the presence of the 7–8 branch. Its signal is doubled; the intensities reflect the anomeric ratio.

The pattern of the *N*-acetyl proton signals of 2, 5, and 5′ is identical to that of oligosaccharide **7**. As described for **9**, the attachment of the third branch to O-4 of Man-4 induces a shift decrement of the *N*-acetyl signal of 5 ($\Delta\delta$ −0.002 p.p.m.). The *N*-acetyl signal of GlcNAc-7 has a relatively narrow line-width (compare **9**, Fig. 9).

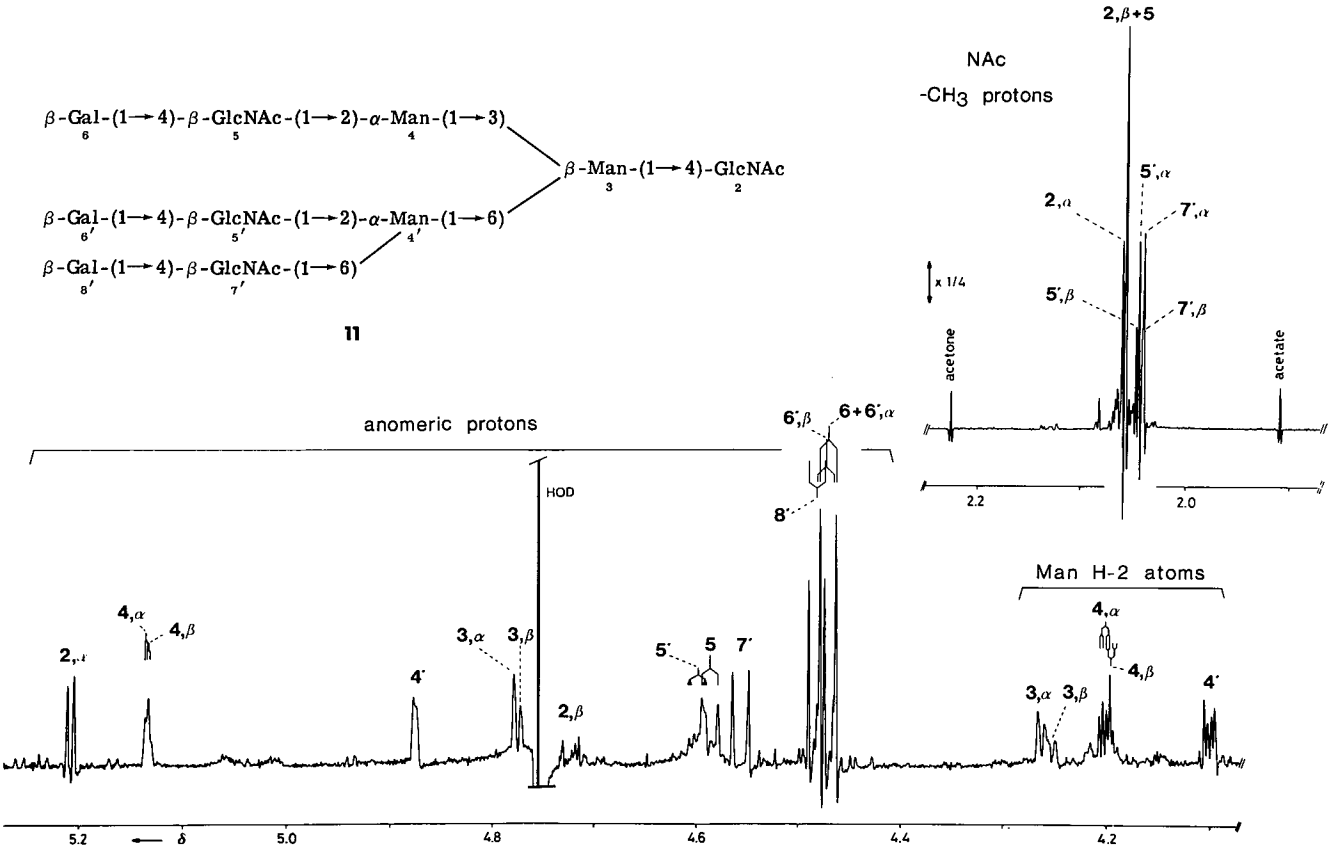
The chemical shifts of the signals of the structural-reporter groups of the additional *N*-acetyllactosamine unit are identical for the α and β anomer of **10**.

Compound **11** has been isolated from the urine of patients suffering from GM_1-gangliosidosis[52,64,64a] or from[64,65] Morquio syndrome type B. This reducing decasaccharide is isomeric with oligosaccharide **10**. It contains three *N*-acetyllactosamine units, the third unit being attached to O-6 of Man-4′. This structure is denoted as a tri′-antennary oligosaccharide.

In the 500-MHz, ^{1}H-n.m.r. spectrum of **11** (see Fig. 11), the presence of the third *N*-acetyllactosamine unit β-(1→6)-linked to Man-4′ is revealed in the pattern of the H-1 and H-2 signals of the Man residues (see Table IV). In comparison to **7**, relatively large shift-decrements of the H-1 ($\Delta\delta \sim -0.054$ p.p.m.) and H-2 signals ($\Delta\delta$ −0.015 p.p.m.) of Man-4′ are observed (see Table IV), whereas the chemical shifts of H-1 and H-2 of Man-4 undergo smaller, but significant shift-increments ($\Delta\delta$ 0.010 and 0.007 p.p.m., respectively). The positions of the H-1 and H-2 signals of Man-3 remain unchanged.

The third *N*-acetyllactosamine unit gives rise to anomeric-proton signals at δ 4.555 (H-1 of GlcNAc-7′) and at δ 4.481 (H-1 of Gal-8′). The

FIG. 11.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **11**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **11**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated.]


β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′
β-Gal-(1→4)-β-GlcNAc-(1→6)
8′ 7′
11
NAc
-CH3 protons
2,β+5
5′,α
2,α
7′,α
x 1/4
5′,β
7′,β
acetone
acetate
2.2
2.0
anomeric protons
6′,β
6+6′,α
8′
HOD
Man H-2 atoms
4,α
4,β
3,α
3,β
2,α
4′
2,β
5′
5
7′
5.2
δ
5.0
4.8
4.6
4.4
4.2

attachment of the βGal(1→4)GlcNAc moiety to O-6 of Man-4′ causes a downfield shift of the H-1 signal of GlcNAc-5′ (Δδ 0.010 p.p.m.) and a considerable line broadening of this doublet, when compared with **7** (compare with the chemical shift and line width of the H-1 signal of 5 in **10**).

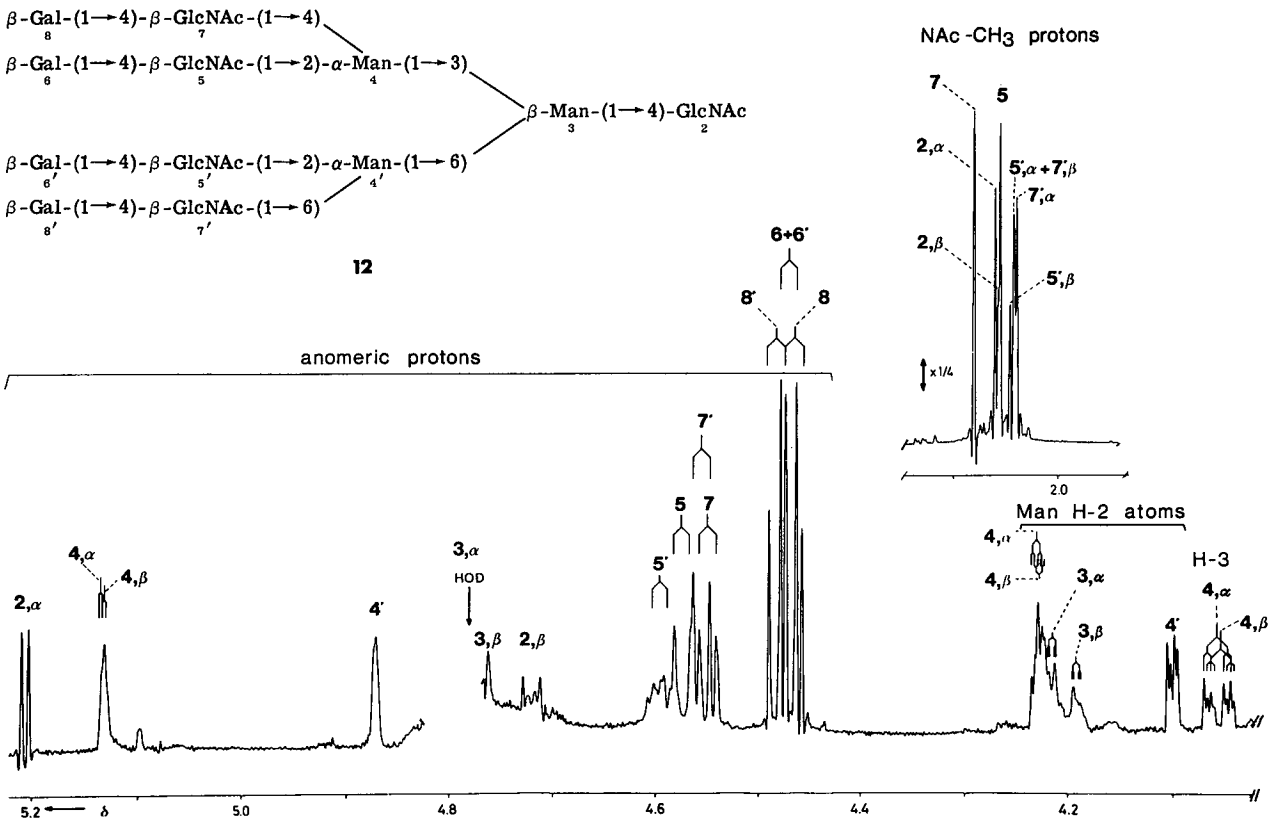
The pattern of the *N*-acetyl proton signals is rather complex. The signal at δ 2.060 is attributed to the α anomer of GlcNAc-2, and that at δ 2.057 belongs partially to the β anomer of GlcNAc-2. The remaining part of the latter singlet belongs to GlcNAc-5 (compare with **7** and **10**, see Table IV). Furthermore, four signals occur in the relatively high-field part of this spectral region, at δ 2.039, 2.040, 2.043, and 2.046, with the relative intensity ratios of 2:1:2:1. Therefore, they can be divided into two pairs, the intensities within each pair reflecting the anomeric ratio. The former pair of signals, separated by $|\Delta\delta_{\alpha-\beta}|$ 0.001 p.p.m., stems from GlcNAc-7′ in the α and β anomer of compound **11**, respectively. The latter pair of singlets, having $|\Delta\delta_{\alpha-\beta}|$ 0.003 p.p.m., is ascribed to GlcNAc-5′ in the α and β form of **11** (see also, Fig. 14b). The assignment is based on the more pronounced influence of the anomeric configuration of GlcNAc-2 on the chemical shifts of the structural-reporter groups of the 5′–6′ branch, in comparison to those of the 7′–8′ branch.

Although the influence of attachment of the third *N*-acetyllactosamine unit, to O-6 of Man-4′, on the chemical shift of the *N*-acetyl singlet of GlcNAc-5 (Δδ 0.005 p.p.m.) might be unexpected, it will later be shown that it is consistent with that for comparable extensions.

The overall effects of the 7′–8′ branch make it tempting to suggest that this branch is in close proximity to the 4–5–6 branch, whereas the anomerization effects suggest that the 4′–5′–6′ branch is in the sphere of influence of GlcNAc-2.

Compound **12** is a tetra-antennary oligosaccharide that was isolated from the urine of a patient with GM_1-gangliosidosis.[52,64,64a] This structure can be conceived of as an extension of the tri- (compound **10**) or tri′-antennary (compound **11**) oligosaccharide with an *N*-acetyllactosamine unit. The 500-MHz, ^{1}H-n.m.r. spectrum of **12** is presented in Fig. 12, and its n.m.r. data are given in Table IV.

FIG. 12.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **12**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **12**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3 for the α anomer of **12**, have been omitted from the spectrum; their position is indicated by an arrow.]


β-Gal-(1→4)-β-GlcNAc-(1→4)
8 7
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'
β-Gal-(1→4)-β-GlcNAc-(1→6)
8' 7'
12
NAc -CH3 protons
7
5
2,α
5',α +7',β
7',α
2,β
5',β
x1/4
2.0
6+6'
8'
8
anomeric protons
7'
5
7
5'
3,α
HOD
3,β
2,β
4,α
4,β
2,α
4'
Man H-2 atoms
4,α
4,β
3,α
3,β
H-3
4,α
4,β
4'
5.2
δ
5.0
4.8
4.6
4.4
4.2

For the spectral interpretation, the spectrum of the triantennary oligosaccharide **10** is used as a reference. The attachment of the *N*-acetyllactosamine unit in β-(1→6) linkage to Man-4′ does not influence the chemical shifts of the H-1 and H-2 signals of Man-3 in the spectrum of the α and β anomers of the oligosaccharide. Shift increments are observed for the H-1 and H-2 signals of Man-4 (Δδ 0.012 and 0.008 p.p.m., respectively), and shift decrements for Man-4′ (Δδ −0.054 and −0.015 p.p.m., respectively). The complete set of chemical shifts of the H-1 and H-2 signals of the Man residues reflects the substitution pattern of the mannotriose branching core. The H-3 signal of Man-4 is unaffected by this substitution.

The signals of the anomeric protons of GlcNAc-5 and -7 resonate at the same positions as for **10**. The H-1 signals of GlcNAc-5′ and -7′ are found at δ 4.592 and 4.554, respectively. The anomeric-proton signals of the four terminal Gal residues are all separated, and could be assigned. The signals of Gal-6, -6′, and -8 are found at positions similar to those for **10**, and the additional doublet at δ 4.480 is ascribed to Gal-8′.

The pattern in the *N*-acetyl region of the spectrum of **12** is similar to that of **10** with respect to the signals of GlcNAc-2, -5, and -7. The four remaining singlets, at δ 2.045, 2.041, 2.040, and 2.038, are ascribed to GlcNAc-5′ in the β and α anomer of **12**, and to GlcNAc-7′ in the β and α anomer of **12**, respectively. This assignment is in accordance with that for oligosaccharide **11** (see Table IV and Fig. 14b).

Considering **12** as an extension of **11**, an analogous reasoning can be applied for the interpretation of its spectrum. From the latter step, the effects of introduction of GlcNAc-7 and Gal-8 (for example, the characteristic shifts for the H-2 signals of Man-3 and -4, and for the *N*-acetyl signal of GlcNAc-5) can be derived.

It turns out that, starting from the diantennary oligosaccharide **7**, the influences of attachment of the third and fourth *N*-acetyllactosamine unit are independent and additive. The doubling of signals due to anomerization, as observed for **7**, remains unaltered upon extension to tri-, tri′-, and tetra-antennary structures. The newly introduced *N*-acetyllactosamine units do not show doubling of signals (except the *N*-acetyl signal of GlcNAc-7′), indicating that they are beyond the sphere of influence of the anomeric center of GlcNAc-2.

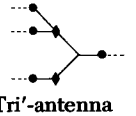
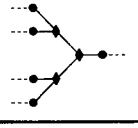
Compound **13** is one of the tetra-antennary, asialo glycopeptides, isolated from asialo α_1-acid glycoprotein,[66–68] that differ only in amino acid composition. It may be conceived as oligosaccharide **12** having an extension of βGlcNAc(1→N)Asn-Gly-Thr. The 500-MHz, ^{1}H-n.m.r. spectrum of **13** is given in Fig. 13a, and the n.m.r. parameters are shown in Table V.

The presence of GlcNAc-1, *N*-glycosylically linked to Asn, is evident from the doublet of its H-1 atom at δ 5.053 and the singlet of its *N*-acetyl protons at δ 2.008. The resonance positions of these protons of GlcNAc-1 are influenced by the structure of the peptide moiety, as already mentioned (compare **8** and **9**; see Table V). In order to illustrate this influence once again, the spectrum of an analog of **13**, ending with βGlcNAc(1→N)[Thr-Pro-]Asn-Lys, is given for comparison in Fig. 13b [GlcNAc-1, δ 5.063 (H-1), δ 2.011 (NAc)]. The H-1 and *N*-acetyl signals of GlcNAc-2 are found at the same positions as described for the triantennary glycopeptide **9**, indicating that GlcNAc-2 is linked β-(1→4) to GlcNAc-1 (see Table II).

The resonance positions of all other structural-reporter groups of **13**, namely, the H-1 doublets of residues 3 to, and including, 8 and 8′, the Man H-2 signals, the H-3 signal of Man-4, and the *N*-acetyl singlets of the peripheral GlcNAc residues, remain unchanged as compared to those of the β anomer of the tetra-antennary oligosaccharide **12**. It may

TABLE VI

Chemical-shift Data for the Mannose H-1 and H-2 Atoms in Asialo Di-, Tri-, and Tetra-antennary Glycopeptides, and in Asialo Di-, Tri-, Tri′-, and Tetra-antennary Oligosaccharides

Schematic structure	Compound		Chemical shift (δ) of					
			H-1 of Man residue			H-2 of Man residue		
			3	4	4′	3	4	4′
Diantenna	**8**		4.765	5.121	4.928	4.246	4.190	4.109
	**7**	α	4.775	5.123	4.927	4.259	4.193[a]	4.114
		β	4.765	5.121	4.930	4.249		
Triantenna	**9**		4.755	5.120	4.924	4.209	4.218	4.108
	**10**	α	4.767	5.118[a]	4.923	4.223	4.218	4.114
		β	4.757		4.925	4.212		
Tri′-antenna	**11**	α	4.776	5.133	4.874	4.263	4.201	4.099
		β	4.770	5.131		4.252	4.198	
Tetra-antenna	**13**		4.757	5.129	4.868	4.210	4.224	4.092
	**12**	α	4.768	5.131	4.870	4.214	4.228	4.099
		β	4.761	5.129		4.203	4.224	

[a] See footnote to Table III.

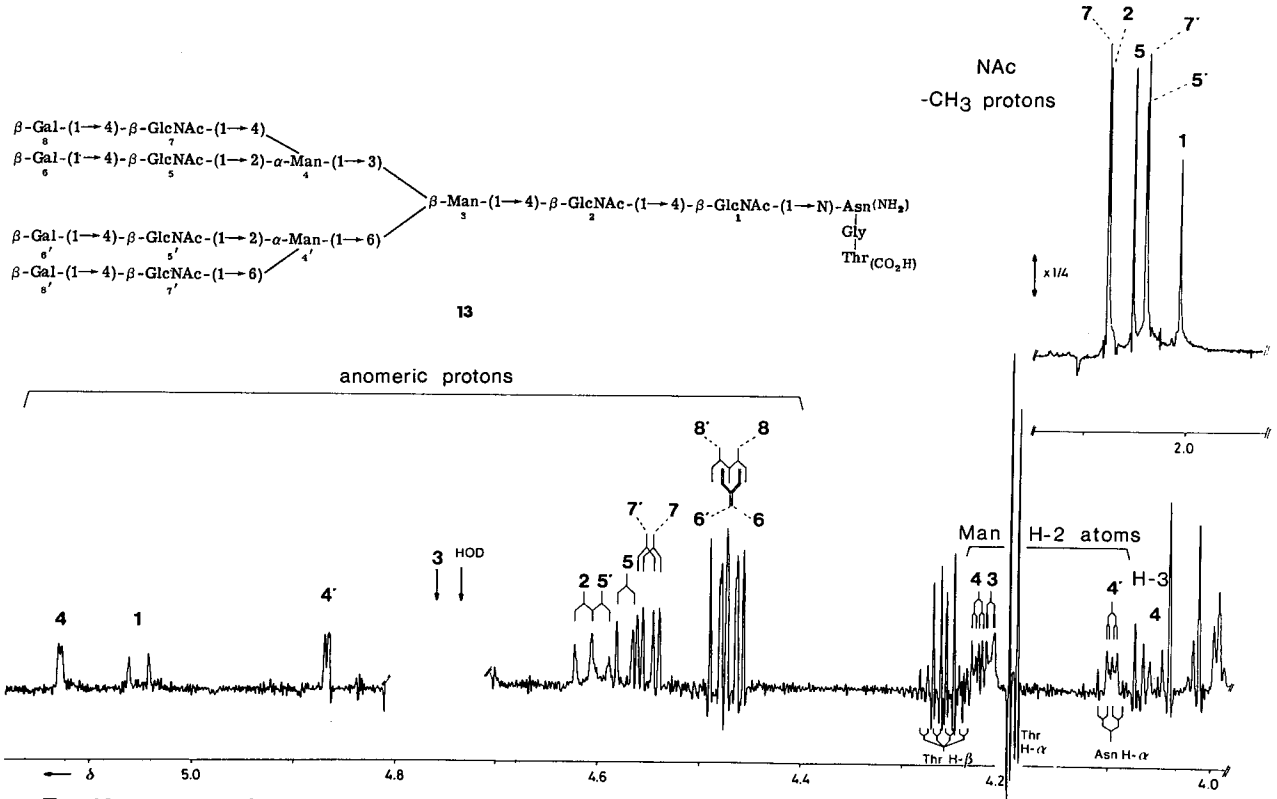


FIG. 13a.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound 13 (peptide moiety: Asn-Gly-Thr).

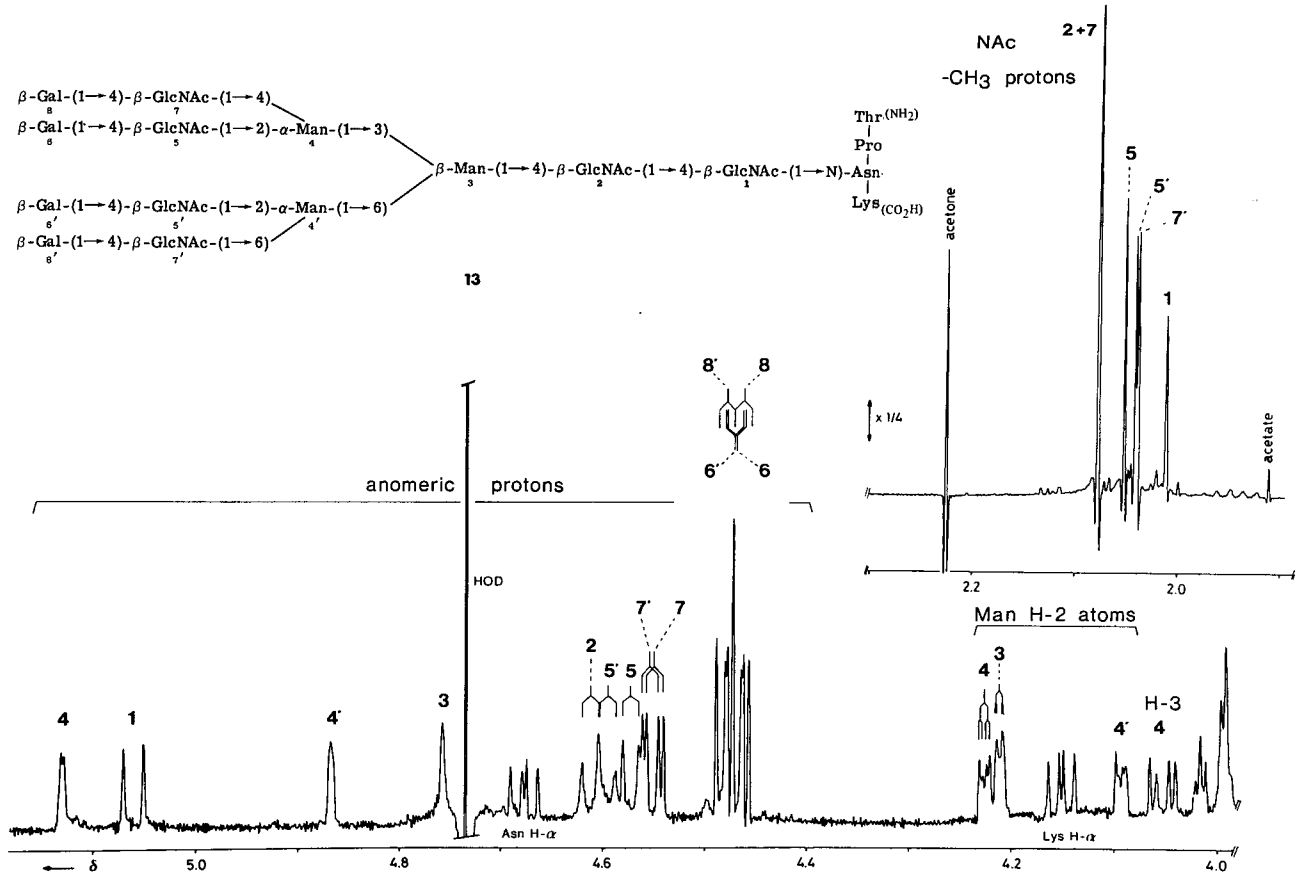


FIG. 13b.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **13** (peptide moiety: Thr-Pro-Asn-Lys). [The bold numbers in the spectra refer to the corresponding residues in the structures. For both spectra, the relative-intensity scale of the *N*-acetyl-proton region (see insertions) differs from that of the other part as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from spectrum (a); their positions are indicated by arrows.]

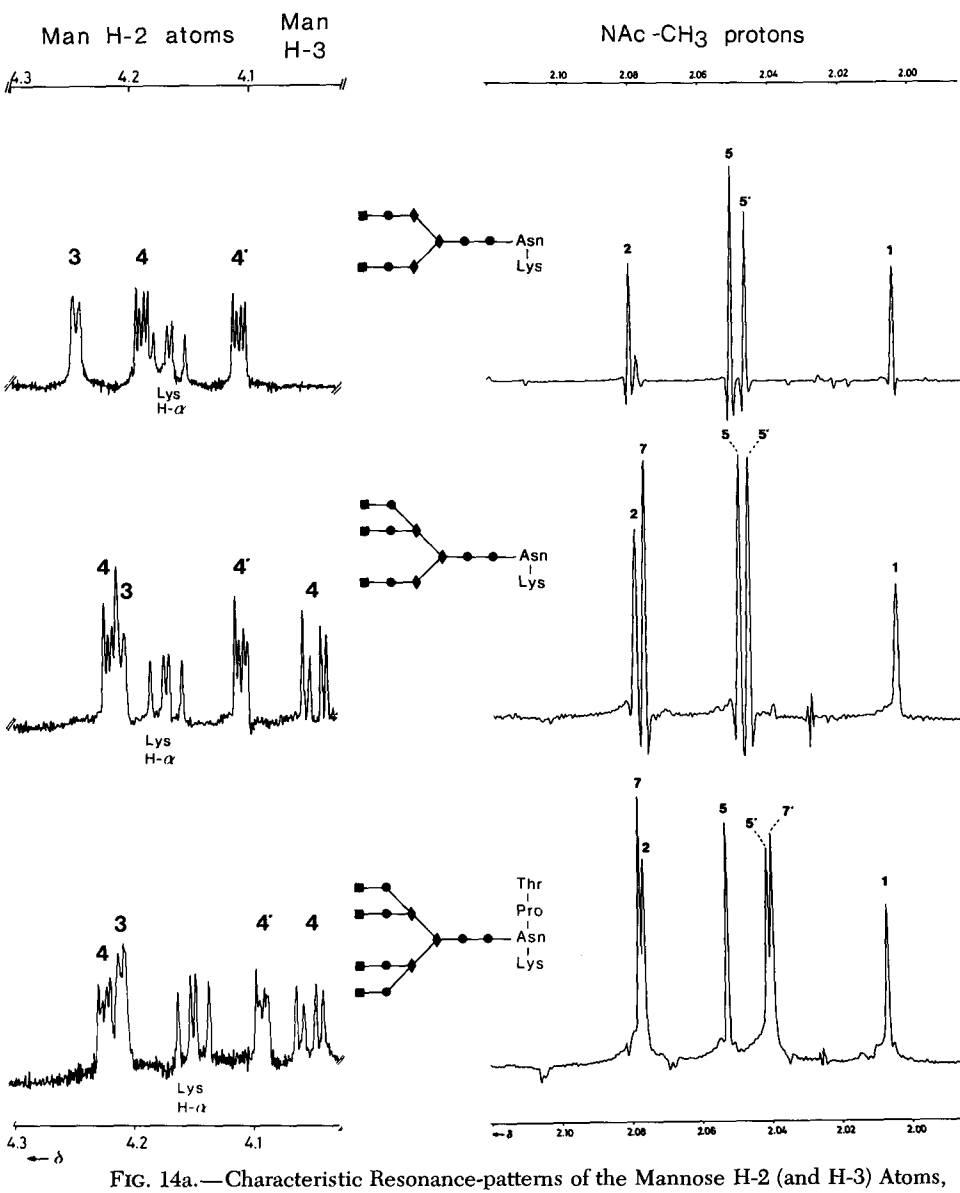
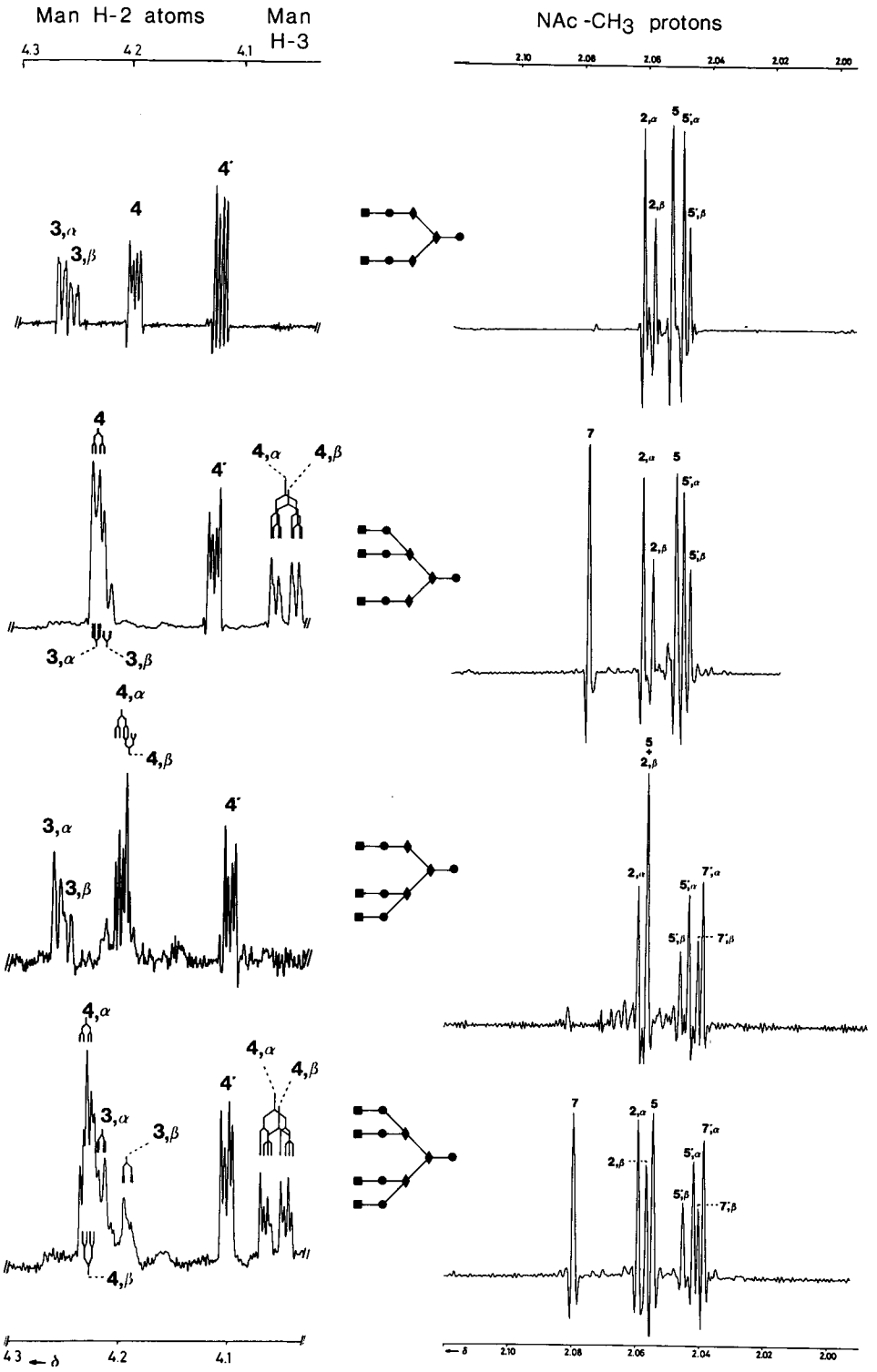


FIG. 14a.—Characteristic Resonance-patterns of the Mannose H-2 (and H-3) Atoms, and of the 2-Acetamido-2-deoxyglucose *N*-Acetyl Methyl Protons, in the 500-MHz, ^{1}H-N.m.r. spectra for Di-, Tri-, and Tetra-antennary Glycopeptides of the *N*-Acetyllactosamine type.

FIG. 14b.—Characteristic Resonance-patterns of the Mannose H-2 (and H-3) Atoms, and of the 2-Acetamido-2-deoxyglucose *N*-Acetyl Methyl Protons, in the 500-MHz, ^{1}H-N.m.r. Spectra for Di-, Tri-, Tri'-, and Tetra-antennary Oligosaccharides of the *N*-Acetyllactosamine Type. Signals for Corresponding Protons in the α and β Anomer of Such an Oligosaccharide, Occurring in the Ratio of 2:1, Coincide, Unless Otherwise Indicated. [The bold numbers in the Figure refer to the corresponding Man (◆) and GlcNAc (●) residues. The carbohydrate chains are represented by the symbolic notation, defined in Chart 1 (see also footnote on page 221).]


Man H-2 atoms
Man H-3
4.3
4.2
4.1
NAc -CH3 protons
2.10
2.08
2.06
2.04
2.02
2.00
4'
4
3,α
3,β
2,α
5
5',α
2,β
5',β
4
4'
4,α
4,β
3,α
3,β
7
2,α
5
5',α
2,β
5',β
4,α
4,β
3,α
4'
3,β
5
2,β
2,α
7',α
5',α
5',β
7',β
4,α
4,α
4,β
4'
3,α
3,β
4,β
7
2,α
5
7',α
2,β
5',α
5',β
7',β
δ
4.3
4.2
4.1
δ
2.10
2.08
2.06
2.04
2.02
2.00

be mentioned that the chemical-shift values of the Man H-1 and H-2 signals unambiguously point to a tetra-antennary type of structure.

The sets of chemical shifts of the Man H-1 and H-2 signals, which are highly characteristic for the type of branching of the *N*-acetyllactosamine type of carbohydrate chain of glycopeptides and related oligosaccharides, are summarized in Table VI. The typical, resonance patterns of the Man H-2 atoms for di-, tri-, tri'-, and tetra-antennary chains are compared in Fig. 14.

An independent criterion for recognition of the asialo di-, tri, tri'-, and tetra-antennary structures is found in the resonance patterns of the *N*-acetyl-proton signals of the glycopeptides and oligosaccharides. For comparison, these characteristic patterns, which may additionally be used as "fingerprints" for recognition of the type of branching of these complex carbohydrate chains, are also depicted in Fig. 14.

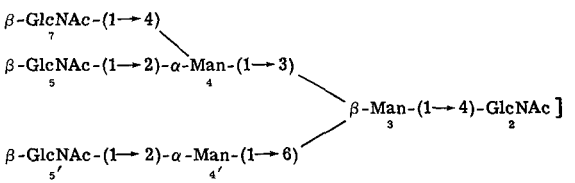
Interesting differences in line width for various anomeric-proton resonances can be observed in a series of di- (**8**), tri- (**9**), and tetra-antennary (**13**) glycopeptides. In particular, the introduction of the 7–8 branch causes a line broadening for H-1 of GlcNAc-5; the effect of the 7′–8′ branch on H-1 of GlcNAc-5′ is even more pronounced. The line-broadening effects in a homologous series of oligosaccharides (**7, 10, 11,** and **12**) are more difficult to define, owing to anomerization effects.

Compound **14** may be isolated from the urine of patients with Sandhoff's disease (GM_2-gangliosidosis variant O).[62] Fundamentally, it is a diantennary oligosaccharide bearing an additional, so-called intersecting GlcNAc residue (9), β-(1→4)-linked to Man-3. The 500-MHz, ^{1}H-n.m.r. spectrum of **14** is given in Fig. 15, and the n.m.r. data are presented in Table VII.

Comparison with the spectrum of **6** shows that the additional

---

FIG. 15.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **14**. [The bold-face numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **14**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. In addition to **14**, the sample contained a small proportion of a positionally isomeric oligosaccharide that possesses GlcNAc-7 instead of -9, as follows. This can be inferred from the signals marked by asterisks.



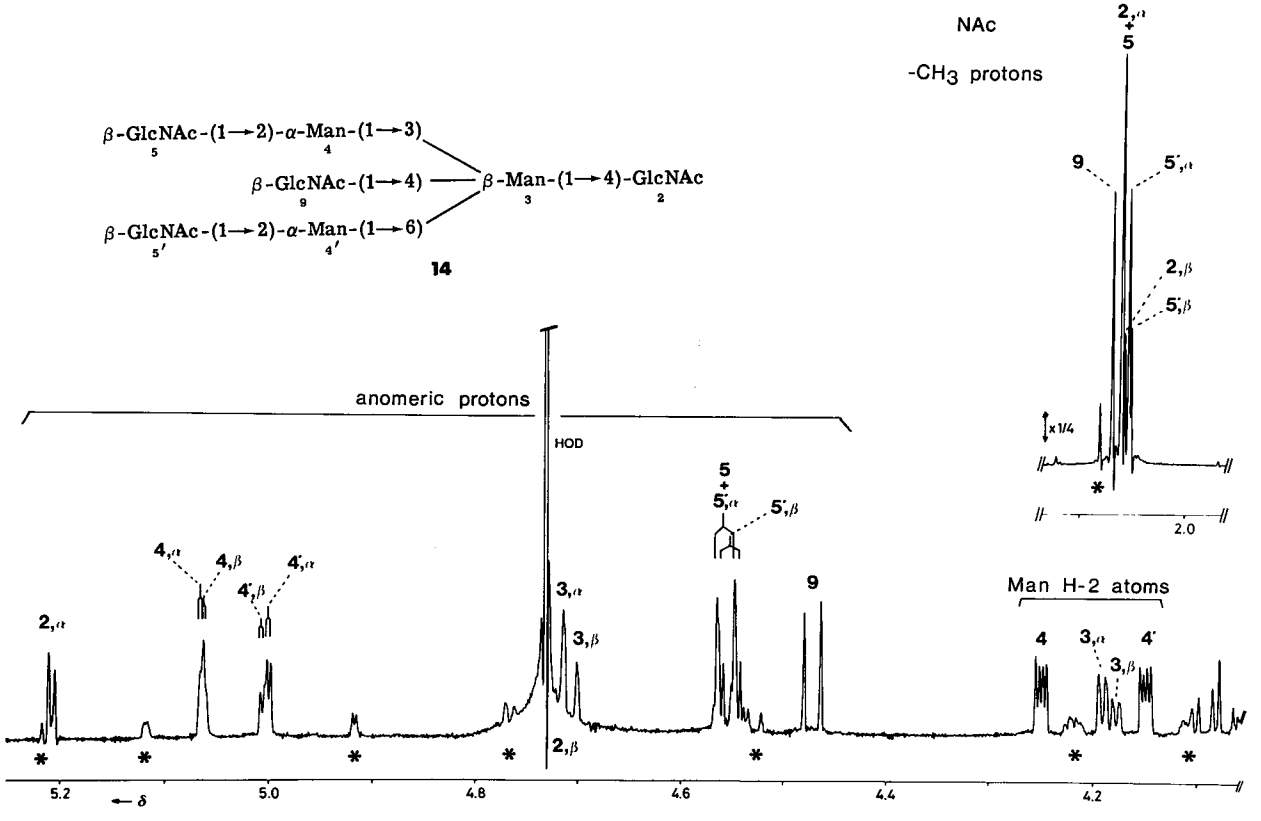
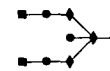
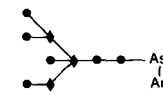

β-GlcNAc-(1→2)-α-Man-(1→3)
β-GlcNAc-(1→4)
β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-GlcNAc
14
NAc
-CH3 protons
anomeric protons
HOD
Man H-2 atoms
x1/4
2.0
5.2
4.8
4.6
4.4
4.2
δ

TABLE VII

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Two Oligosaccharides (14 and 15) and a Glycopeptide (16) Containing the Intersecting GlcNAc Residue (9)

			Compound and schematic structure		
			14	15	16
					
Reporter group	**Residue**	**Anomer of oligosaccharide**			
H-1 of	1		—	—	5.056
	2	α	5.206	5.204	4.613
		β	4.722	4.721	
	3	α	4.713	4.704	4.687
		β	4.699	4.686	
	4	α	5.063	5.059	5.057
		β	5.060	5.057	
	4′	α	4.999	5.008	4.998
		β	5.003	5.020	
	5	α,β	4.555	4.583	4.537
	5′	α	4.555	4.579	4.545
		β	4.549	4.591	
	6	α,β	—	4.468	—
	6′	α	—	4.473	—
		β		4.477	
	7		—	—	4.520
	9	α	4.469	4.466	4.463
		β		4.468	
H-2 of	3	α	4.190	4.190	4.146
		β	4.177	4.175	
	4	α,β	4.247	4.258	4.282
	4′	α	4.148	4.148	4.141
		β		4.142	
H-3 of	4	α,β	<4.0	<4.0	4.044
	4′	α,β	<4.0	<4.0	<4.0
NAc of	1		—	—	2.008
	2	α	2.059	2.064	2.077
		β	2.056	2.060	
	5	α,β	2.059	2.053	2.057
	5′	α	2.052	2.044	2.048
		β	2.050	2.039	
	7		—	—	2.082
	9	α,β	2.068	2.064	2.063

GlcNAc-9 has a profound influence on the resonance positions of several reporter groups. For H-1, as well as H-2, of Man-3, two signals are observed, all four of which are shifted upfield as compared to **6** (compare Table IV). The signals of H-1 of Man-4 shift to higher field ($\Delta\delta$ −0.057 p.p.m.), whereas the signal of H-2 of this residue shifts downfield ($\Delta\delta$ +0.058 p.p.m.). The latter signal may be recognized very readily as belonging to Man-4, because of the specific pattern of the well resolved doublet of doublets for H-2 of an α-linked Man (see earlier). This eliminates any confusion with H-2 of Man-3, which resonates at δ ~4.25 in normal, diantennary structures (**6**, **7**, and **8**). The H-1 doublets, as well as the H-2 signal, of Man-4′ are shifted downfield.

The resonance positions of the anomeric protons of GlcNAc-2, -5, and -5′ are essentially unchanged with respect to **6**. The relatively sharp, H-1 doublet of GlcNAc-9 is found at δ 4.469; its $J_{1,2}$ value (8.4 Hz) is indicative of a β-glycosylic linkage.

The presence of GlcNAc-9 makes the subspectra of the two anomers of **14** more different than those of **6**; this is especially illustrated by the relatively large differences in chemical shift for H-1 of Man-4, as well as H-1 of Man-4′, in both anomers of **14**. It seems as if the steric requirements of GlcNAc-9 push the two branches towards the sphere of influence of the anomeric center of GlcNAc-2. GlcNAc-9 itself is apparently remote from this center, as no doubling of its signals is observed.

In the *N*-acetyl region of the spectrum, the singlet at δ 2.068 is ascribed to GlcNAc-9. The α and β anomer of GlcNAc-2 give rise to two signals, at δ 2.059 and 2.056, respectively, as described for other reducing oligosaccharides. Considering the strong influence of anomerization on the chemical shifts of the structural-reporter groups of the lower branch (for example, H-1 of Man-4′ in **14**), the set of signals at δ 2.052 and 2.050 is assigned to the *N*-acetyl-group protons of GlcNAc-5′ in the α and β form of **14**, respectively. Consequently, the rest of the broad singlet at δ 2.059 is attributed to GlcNAc-5; this means that the presence of GlcNAc-9 causes a downfield shift ($\Delta\delta$ 0.005 p.p.m.) for this signal.

Compound **15** is another diantennary oligosaccharide, having the intersecting GlcNAc-9 linked β-(1→4) to Man-3. It has been isolated from the urine of a patient with[64,65] Morquio syndrome type B. The 500-MHz, ^{1}H-n.m.r. spectrum of **15** is presented in Fig. 16, and the chemical shifts of the structural-reporter groups are included in Table VII.

As compared to the corresponding agalacto compound **14**, Gal-6 and -6′ introduce downfield shifts of the H-1 atoms of GlcNAc-5 and -5′

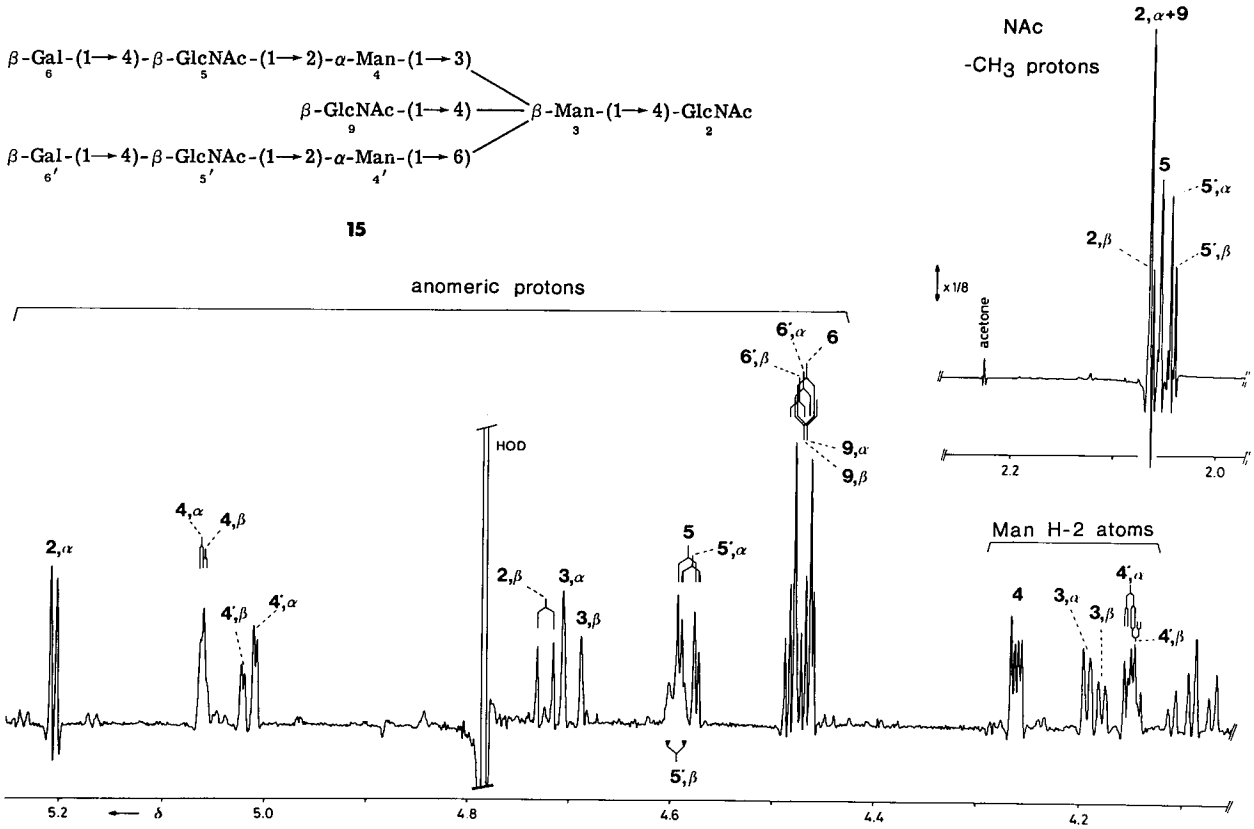


FIG. 16.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **15.** [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **15**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated.]

(compare the step from **6** to **7**), whereas the chemical shift of the anomeric proton of GlcNAc-9 is essentially unaffected. However, for **15**, the latter proton gives rise to two doublets. The doublet of this H-1 atom for the β anomer of compound **15** coincides with that of H-1 of Gal-6, at δ 4.468, whereas, for the α form, it is found at δ 4.466. This assignment is based on the relative-intensity ratio of the aforementioned doublets, namely, 2:1. The H-1 resonance of Gal-6′ is also doubled, due to anomerization. This feature is in accord with the relatively strong (for example, $|\Delta\delta_{\alpha-\beta}|$ 0.012 p.p.m. for H-1 of Man-4′, as well as for H-1 of GlcNAc-5′), anomerization effects present on lower-branch residues (compare **7**), especially when GlcNAc-9 is present.

The set of chemical shifts of the H-1 and H-2 signals of Man-3, -4, and -4′ reflects the substitution pattern of the branching point, as outlined for the agalacto analog **14**.

In the *N*-acetyl region of the spectrum, five separate singlets are observed. At δ 2.064, two signals coincide that are attributed to GlcNAc-9 and to the α anomer of GlcNAc-2. The resonance at δ 2.060 stems from GlcNAc-2 in the β anomer of **15**. The *N*-acetyl signals of GlcNAc-5 and -5′ are shifted slightly upfield, as compared to **14** (compare the step from **6** to **7**). The relatively large $|\Delta\delta_{\alpha-\beta}|$ value for the *N*-acetyl signals of GlcNAc-5′ is worth mentioning, as it is typical for the presence of GlcNAc-9 in combination with that of Gal-6′.

Starting from the spectrum of **7** as a reference, a detailed interpretation of the spectrum of **15** can also be achieved. Essentially the same characteristic influences of GlcNAc-9 may be traced along this route, as described for the step from **6** to **14**.

Compound **16** is a triantennary glycopeptide containing GlcNAc-9. It may be obtained by pronase digestion of chicken ovotransferrin.[70,71] The de-Arg analog of **16** was also available for investigation. The 500-MHz, ^{1}H-n.m.r. spectrum of **16** is presented in Fig. 17, and its n.m.r. data are summarized in Table VII. For the spectral interpretation, **16** is conceived as **14** extended with GlcNAc-7, as well as with βGlcNAc(1→N)Asn-Arg.

The chemical shift (δ 5.056) of the H-1 signal of GlcNAc-1 in **16** is distinguishable from that (δ 5.070) for its de-Arg analog, illustrating the effect of the peptide moiety on this residue. A similar feature occurs for the *N*-acetyl signal of GlcNAc-1. However, the H-1 signal and the relatively broad *N*-acetyl signal of GlcNAc-2 are found at δ 4.613 and 2.077, respectively, for both compounds. This shows that the influence of Arg is restricted to GlcNAc-1.

(71) L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, G. Spik, B. Fournet, and J. Montreuil, *Eur. J. Biochem.*, 100 (1979) 569–574.

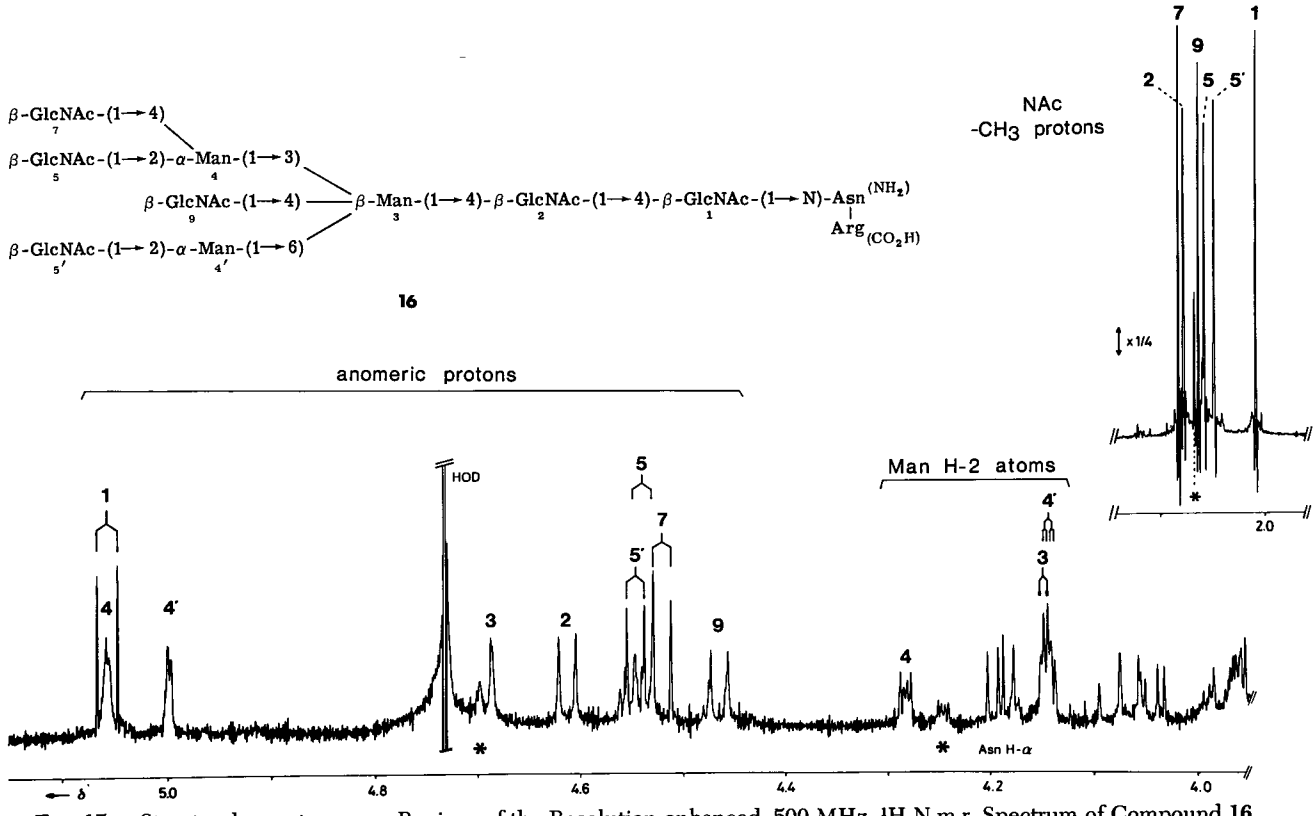


FIG. 17.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **16**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. In addition to **16**, the sample contained a small proportion of a similar glycopeptide lacking GlcNAc-7 (as compared to **16**), that is, a glycopeptide analog of compound **14**. This can be inferred, for example, from the signals marked by asterisks.]

Owing to the presence of GlcNAc-7, the H-1 signal of Man-3 is shifted upfield ($\Delta\delta$ −0.012 p.p.m.) with respect to the H-1 signal of Man-3 in the β anomer of oligosaccharide **14**. This effect has the same magnitude as described for this signal in the steps from a di- (**8**) to a tri-antennary (**9**) glycopeptide, from a di- (**7**) to a tri-antennary (**10**) oligosaccharide, and also from a tri'- (**11**) to a tetra-antennary (**12**) oligosaccharide (compare Tables IV and V). In consequence, the presence of Gal-8 is not a requirement for the shift effect observed; it can be fully attributed to the attachment of GlcNAc-7 to Man-4 in β-(1→4) linkage. Similarly, GlcNAc-7 is responsible for the specific changes in the chemical shifts of the H-2 signals of Man-3 and -4 in comparison to **14**. The resulting set of chemical shifts for the H-1 and H-2 signals of the Man residues of **16** is characteristic for this kind of substitution of the mannotriose core.

The H-1 signal of GlcNAc-5 shifts upfield ($\Delta\delta$ −0.018 p.p.m.) upon substitution of Man-4 with GlcNAc-7 in a β-(1→4) linkage. Furthermore, the doublet of this anomeric proton has broader lines than the doublet of the H-1 atoms of the other terminal GlcNAc residues. This might reflect the lessened conformational freedom of GlcNAc-5, due to the presence of GlcNAc-7 (compare, for instance, Fig. 9). GlcNAc-7 also exerts some influence on the chemical shift and the line width of the H-1 signal of GlcNAc-9.

In the *N*-acetyl region, the assignment of the signals of GlcNAc-1, -2, -5', and -7 is straightforward (compare **9**). The signal at δ 2.057 is ascribed to GlcNAc-5. It is shifted upfield ($\Delta\delta$ −0.002 p.p.m.) in comparison to **14**, which corresponds with the known effect of attachment of GlcNAc-7 (compare **9**). In consequence, the signal at δ 2.063 belongs to GlcNAc-9. Apparently, this singlet undergoes an upfield shift ($\Delta\delta$ −0.005 p.p.m.) due to the introduction of GlcNAc-7.

Comparison of **14** with **15** shows that, for the H-1 signals of GlcNAc-5 and -5', the expected "agalacto" upfield shifts ($\Delta\delta \sim -0.025$ p.p.m.) occur. Assuming that this effect is independent of the location of the *N*-acetyllactosamine unit in the structure, the chemical shift of the H-1 signal of GlcNAc-7 for **16** may be derived from that for **9**, correcting for the absence of Gal-8.

Compound **17** is a mono-antennary, upper-branch, asialo oligosaccharide that may be regarded as an extension of trisaccharide **5** with an *N*-acetyllactosamine unit β-(1→2)-linked to Man-4. This pentasaccharide has been isolated, from the urine of a patient with GM_1-gangliosidosis,[52,64,64a] in a mixture of oligosaccharides consisting of **17** and **18** in the ratio of 3:2. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is given in Fig. 18a, and the n.m.r. data are compiled in Table VIII.

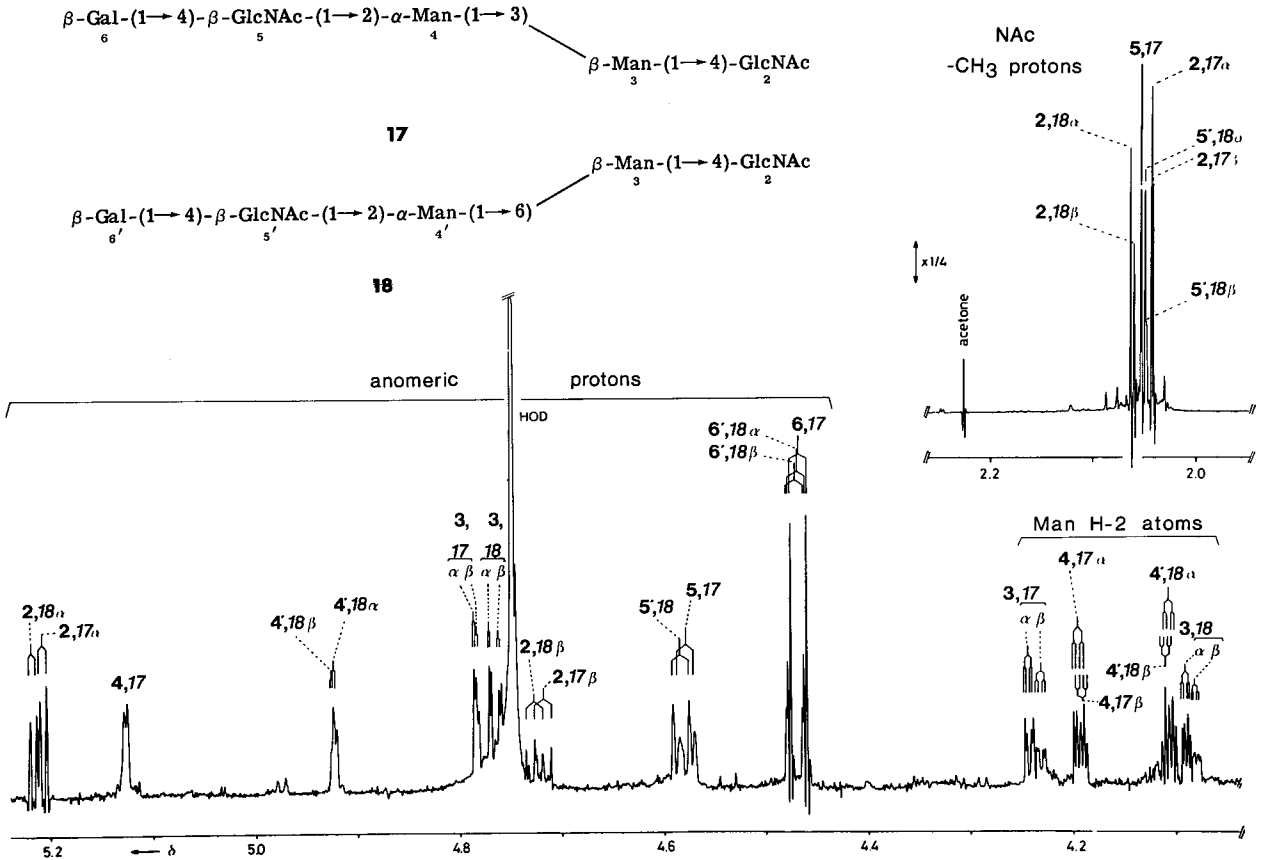

β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
17
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'
β-Man-(1→4)-GlcNAc
3 2
18
NAc
-CH3 protons
5,17
2,17α
2,18α
5',18α
2,17β
2,18β
5',18β
x1/4
acetone
2.2
2.0
anomeric
protons
HOD
6',18α
6',18β
6,17
3, 3,
17 18
α β α β
2,18α
2,17α
4',18β
4',18α
2,18β
2,17β
4,17
5',18
5,17
Man H-2 atoms
4,17α
3,17
α β
4',18α
3,18
α β
4',18β
4,17β
5.2
δ
5.0
4.8
4.6
4.4
4.2

TABLE VIII

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Asialo Mono-*N*-acetyllactosaminyl Oligosaccharides (Compounds 17–20)

			Compound and schematic structure			
			17	18	19	20
Reporter group	Residue	Anomer of compound				
H-1 of	2	α	5.207	5.216	5.213	5.213
		β	4.719	4.726	4.725	4.725
	3	α	4.785	4.769	4.788	4.779
		β	4.782	4.760	4.779	4.768
	4	α	5.127	—	5.122	5.106
		β				5.102
	4′	α	—	4.920	4.920	4.925
		β		4.923	4.925	4.929
	5	α,β	4.577	—	4.579	—
	5′	α,β	—	4.582	—	4.583
	6	α,β	4.468	—	4.467	—
	6′	α	—	4.470	—	4.470
		β		4.472		4.472
H-2 of	3	α	4.242	4.089	4.263	4.263
		β	4.231	4.079	4.253	4.253
	4	α	4.194	—	4.193	4.071
		β	4.190			
	4′	α	—	4.104	<4.0	4.112
		β		4.107		
NAc of	2	α	2.043	2.063	2.058	2.061
		β	2.041	2.060	2.055	2.058
	5	α,β	2.053	—	2.052	—
	5′	α	—	2.048	—	2.048
		β		2.047		2.046

FIG. 18a.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **17** and **18** in the Ratio of 3:2. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. (Anomeric-proton signal designated as 4′, **18**α means: H-1 of Man-4′ in the α anomer of compound **18**). Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated.]

As with **5**, the reducing character of compound **17** gives rise to two doublets for H-1 of GlcNAc-2, at δ 5.207 ($J_{1,2}$ 3.1 Hz) and δ 4.719 ($J_{1,2}$ 8.1 Hz), for the α and β anomer, respectively. For the *N*-acetyl protons of GlcNAc-2, two singlets are observed, also in the anomeric ratio, at δ 2.043 (α) and 2.041 (β anomer of **17**). The set of chemical shifts of the H-1 doublets, in conjunction with that of the *N*-acetyl singlets of GlcNAc-2, is indicative of a mono-α-(1→3) substitution of Man-3 by another Man residue (compare **5** in Table III, and **18** in Table VIII). The aforementioned chemical shifts are not influenced by the presence of the *N*-acetyllactosamine unit.

Equally, this extension has hardly any influence on the chemical shifts and coupling constants ($J_{1,2}$ and $J_{2,3}$) of the H-1 and H-2 signals of Man-3 in both anomers of the oligosaccharide. The $\Delta\delta_{\alpha-\beta}$ values (0.003 p.p.m.) for H-1 of Man-3 in the α and β anomers of **17** and **5** are smaller than those observed for all other related oligosaccharides ($\Delta\delta_{\alpha-\beta}$ 0.01 p.p.m., compare Tables IV and VII).

The H-1 atom of Man-4 resonates, as one doublet with relatively broad lines, at a position that is rather low-field for an asialo upper branch (δ 5.127). The H-2 signal of Man-4 appears as two doublets of doublets in the anomeric ratio. The shift increment (Δδ 0.12 p.p.m.) of H-2 of Man-4 in comparison to **5** reflects the substitution at O-2 of Man-4 by GlcNAc-5, as already described for the step from **5** to **6**.

The single doublet of H-1 of Gal-6 is found at δ 4.468, which is in accord with the diantennary oligosaccharide **7** (see Table IV). The relatively narrow lines of this signal correspond with the terminal position of Gal-6.

The doublet at δ 4.577 has relatively broad lines, and must be ascribed to H-1 of GlcNAc-5. Its resonance position deviates by 0.005 p.p.m. from that observed for this proton in the diantennary oligosaccharide **7**, and turns out to be characteristic for the peripheral GlcNAc-5 in an asialo, mono-*N*-acetyllactosamine, upper-branch oligosaccharide (see later; compound **19**, Table VIII). As already described, the resonance position of the signal of the *N*-acetyl group of GlcNAc-5 is not influenced by anomerization of the oligosaccharide.

Compound **18** is a mono-antennary, lower-branch, asialo oligosaccharide that has been isolated from the urine of patients suffering from GM_1-gangliosidosis[52,64,64a] or from[64,65] Morquio syndrome type B. The 500-MHz, ^{1}H-n.m.r. spectrum of the relatively pure pentasaccharide is presented in Fig. 18b. The signals corresponding to **18** in Fig. 18a can readily be traced by comparison of spectra 18a and 18b. The chemical shifts of its structural-reporter groups are summarized in Table VIII.

The resonance positions of the H-1 doublets of GlcNAc-2 for the α and β anomer of **18**, in combination with those of the *N*-acetyl signals

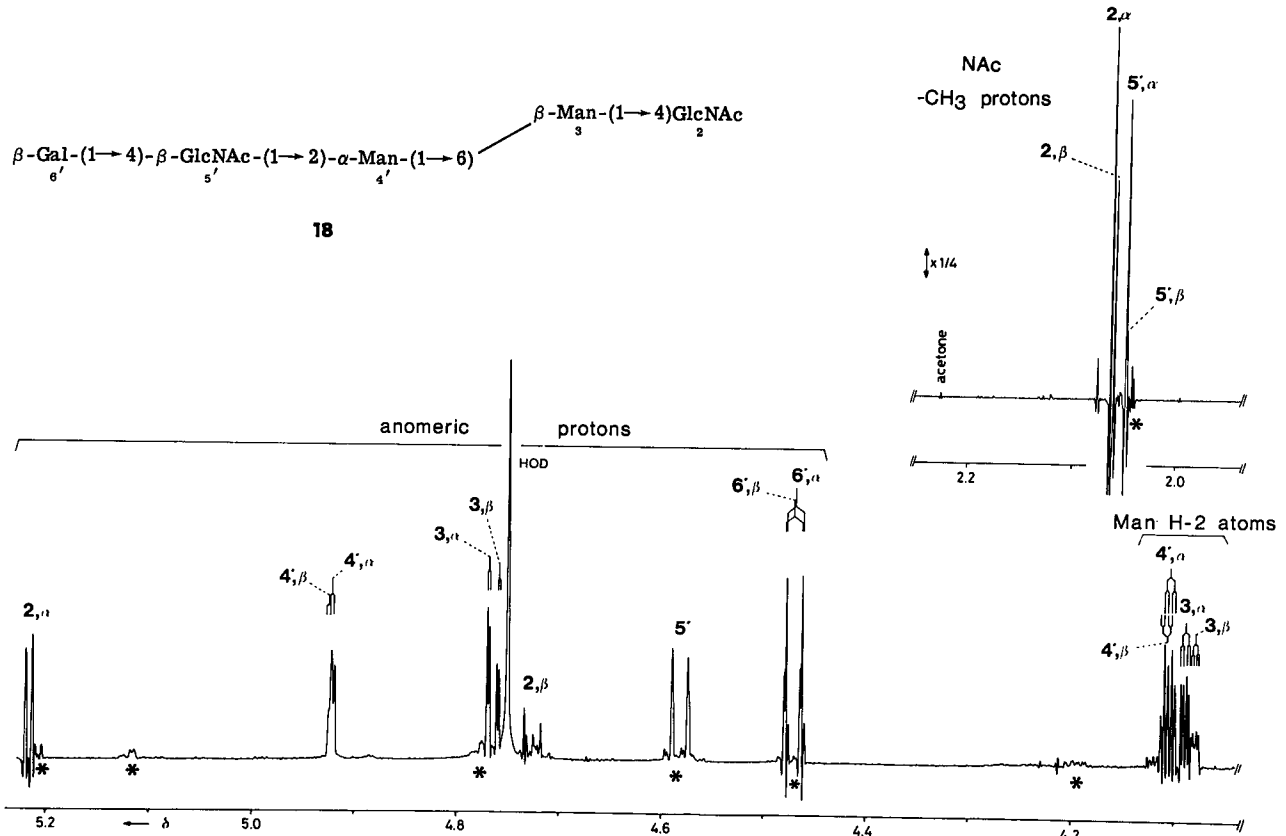


FIG. 18b.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **18**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **18**, occurring in this anomeric mixture of the ratio of 2 : 1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated. In addition to **18**, the sample contained a very small proportion of **17**, as can be inferred from the signals marked by asterisks.]

of GlcNAc-2, are specific for mono-α-(1→6) substitution of Man-3. Also, the chemical shifts for H-1 and H-2 of Man-3 can be used to determine whether Man-3 is mono-α-(1→3) or -α-(1→6) substituted by another Man residue. The values for H-1 and H-2 of Man-3 in the β anomer of **18** are in accord with those observed for the glyco-asparagines **3** and **4**, provided that they are corrected for the attachment of GlcNAc-5′ (for H-1, $\Delta\delta$ −0.009 p.p.m.; for H-2, $\Delta\delta \sim -0.005$ p.p.m., as may be deduced from comparison of the spectrum of compound **19** with that of **6**). The relatively large, $J_{1,2}$ value for Man-3 in compound **18** (1.1 Hz) gives rise to two doublets of doublets for H-2 of Man-3, resembling the resonance pattern of the H-2 atom of an α-linked Man residue in a reducing oligosaccharide (see Fig. 18a and b). The effects of anomerization on the chemical shifts of the signals of H-1 and H-2 of Man-3 are comparable to those observed for other oligosaccharides (for example, **6**, **7** and **10**; see Table IV).

The H-1 and H-2 signals of Man-4′, the H-1 doublet of Gal-6′, and the *N*-acetyl signal of GlcNAc-5′ are doubled. This is in line with the observation that the influence of anomerization is, in general, more pronounced in the signals of the lower- than of the upper-branch residues.

The chemical shifts of the structural-reporter groups of Gal-6′ and GlcNAc-5′ are in full accord with those observed for the corresponding protons in the diantennary oligosaccharide **7** (see Table IV). The chemical shift and the line width of the H-1 doublet of GlcNAc-5′ differ considerably from those observed for GlcNAc-5 in the spectrum of **17** (see Fig. 18a). The chemical shifts of the H-1 and H-2 atoms of Man-4′ reflect the incompleteness of the branching core (compare with **6**, **7**, **19**, and **20**; see Tables IV and VIII).

Compound **19** is a reducing oligosaccharide consisting of the mannotriose branching core bearing an *N*-acetyllactosamine unit in the upper branch. It has been isolated from the urine of a patient suffering from[64,65] Morquio syndrome type B, as a constituent of a mixture containing compounds **19** and **20** in the ratio of 1:3. The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture is shown in Fig. 19. The spectral parameters for compound **19**, which belong to the signals having the lower intensities in Fig. 19, are listed in Table VIII.

For the spectral interpretation, **19** is considered as being structure **17** extended with Man-4′, α-(1→6)-linked to Man-3. The attachment of Man-4′ leads to chemical shifts for H-1 of the α anomer of GlcNAc-2 and for the *N*-acetyl protons of 2 for both anomers of **19**, which are identical to those for the diantennary compound **7** (at 300 K, the β-anomeric, H-1 signal of GlcNAc-2 is partially hidden under the HOD line).

The major effect on Man-3 due to the introduction of Man-4′ is found in the chemical shift of its H-2 signal ($\Delta\delta$ 0.02 p.p.m.). Man-4′ causes an upfield shift ($\Delta\delta$ −0.005 p.p.m.) and a line broadening of the H-1 signal of Man-4. However, the structural-reporter-group signals of GlcNAc-5 and Gal-6 are unaffected.

The well resolved, H-1 doublets of the terminal Man-4′ are found at δ 4.920 (α) and 4.925 (β anomer of **19**). The latter value differs from that observed for H-1 of the terminal Man-4′ in glyco-asparagine **3**. This difference can be ascribed to the influence of the presence of Man-4. The H-2 signal of Man-4′ is hidden in the bulk of the sugar skeleton protons (δ <4.0), which is characteristic for a terminal α-(1→6)-linked Man group (see Table II).

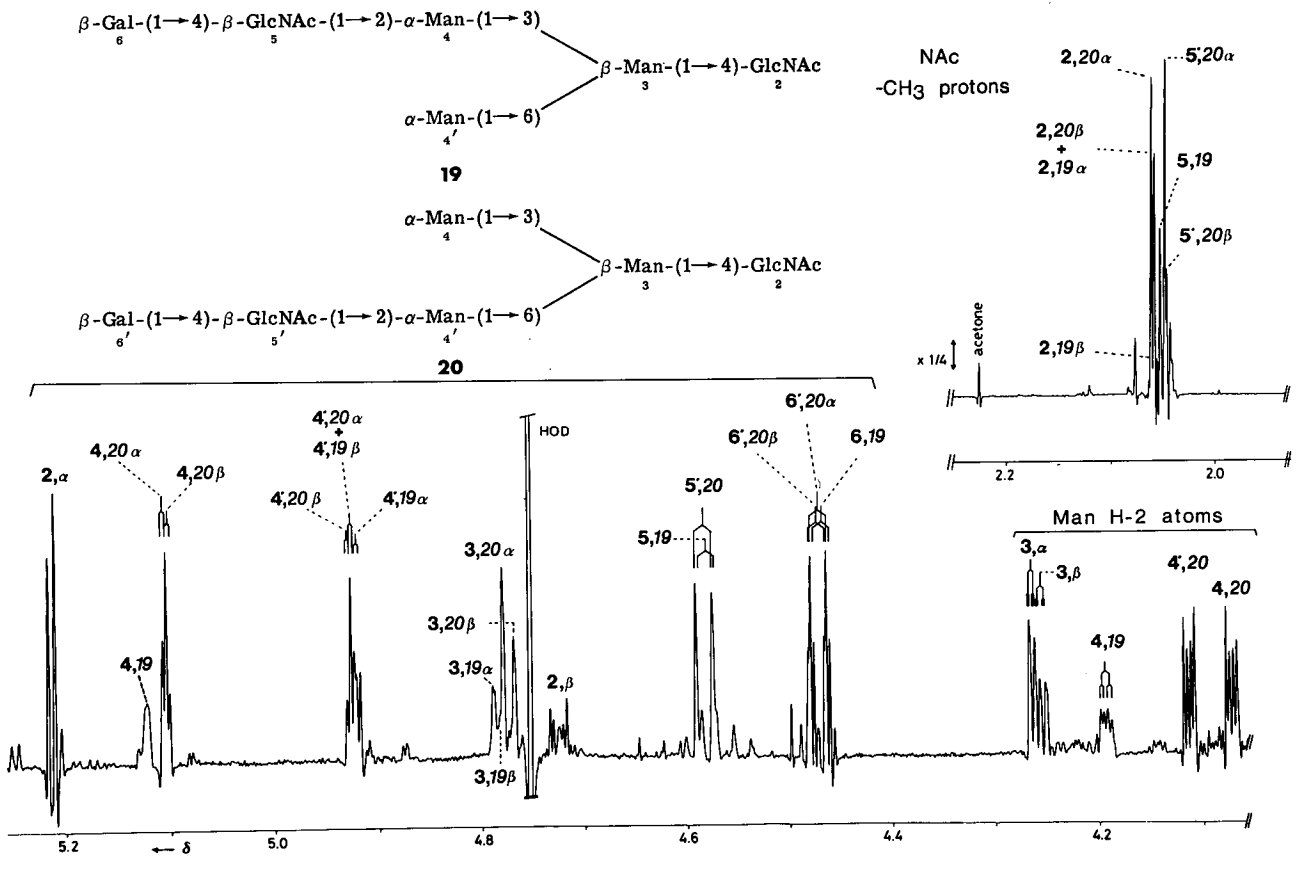
Comparison of the spectrum of **19** with that of the diantennary oligosaccharide **7** shows that the absence of GlcNAc-5′ and Gal-6′ has no influence on the structural-reporter-group signals of Gal-6, GlcNAc-5, and Man-4, but results in changes for those of Man-3 and -4′. In fact, the absence of GlcNAc-5′ is responsible for the alterations, which is evident from comparison with the data for oligosaccharide **6**.

Compound **20** is a reducing oligosaccharide consisting of the mannotriose branching core bearing an *N*-acetyllactosamine unit in the lower branch; **20** was obtained as the main component of a mixture also containing **19**. As already described, this mixture has been isolated from the urine of a patient suffering[64,65] from Morquio syndrome type B. The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture is depicted in Fig. 19; the relevant, spectral parameters for **20** are compiled in Table VIII.

For the interpretation of the signals having the higher intensities (see Fig. 19), belonging to **20**, this oligosaccharide is conceived as structure **18** extended with Man-4 that is α-(1→3)-linked to Man-3. Comparison of the spectral data with those for **18** demonstrates that the chemical shifts of H-1 and H-2 of Man-3 and -4′ have undergone alterations towards values that are typical for a complete branching-core (see Table VIII).

The (terminal) nonreducing position of Man-4 in **20** is revealed by its H-1 (δ 5.106 for the α, and δ 5.102 for the β anomer of the oligosaccharide) and H-2 chemical-shift values (δ 4.071), which are closely related to those for the corresponding residue in oligosaccharide **5** (see Table III). The slight, upfield shift of H-1 of Man-4 compared with **5** ($\Delta\delta \sim -0.007$ p.p.m.) is caused by the presence of Man-4′; the same effect is observed in the step from **17** to **19**. Furthermore, the narrow lines of the signals of Man-4 reflect its terminal, relatively mobile position in the carbohydrate chain.

Comparison of the spectral data for **20** with those for the dianten-


β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-Man-(1→6)
4'
19
α-Man-(1→3)
4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'
20
NAc
-CH3 protons
2,20α
5,20α
2,20β
+
2,19α
5,19
5',20β
acetone
x 1/4
2,19β
2.2
2.0
2,α
4,20α
4,20β
4',20α
+
4',19β
4',20β
4',19α
HOD
6',20α
6',20β
6,19
5',20
5,19
3,20α
3,20β
3,19α
3,19β
2,β
4,19
Man H-2 atoms
3,α
3,β
4',20
4,20
5.2
δ
5.0
4.8
4.6
4.4
4.2

nary oligosaccharides **6** and **7** reveals that the essential differences are located in the spectral features of Man-4. In particular, the effect of anomerization is lessened by the attachment of GlcNAc-5.

b. Extensions of Carbohydrate Chains of the *N*-Acetyllactosamine Type with Sialic Acid Groups (Compounds 21–41).—Many carbohydrate chains of the *N*-acetyllactosamine type, attached to, or liberated from, a glycoprotein peptide backbone, bear one or more sialic acid (NeuAc) groups (in a terminal position).[1–4] Symbols employed for compounds **21–41** are depicted in Chart 2.

Two well defined ways of terminating chains of the *N*-acetyllactosamine type by a NeuAc residue will be treated here, namely, (*i*) the α-(2→6) and (*ii*) the α-(2→3) attachment of NeuAc to a Gal residue that forms part of an *N*-acetyllactosamine unit. First, the spectral features of structures containing only α-(2→6)-linked NeuAc (compounds **21–34**) will be discussed. Then, the characteristics and influences of NeuAc in α-(2→3) linkage to Gal will be considered (compounds **35–38**). Finally, (*iii*) some structures (compounds **39–41**) possessing α-NeuAc(2→6)Gal, as well as αNeuAc(2→3)Gal, will be the subjects of discussion.

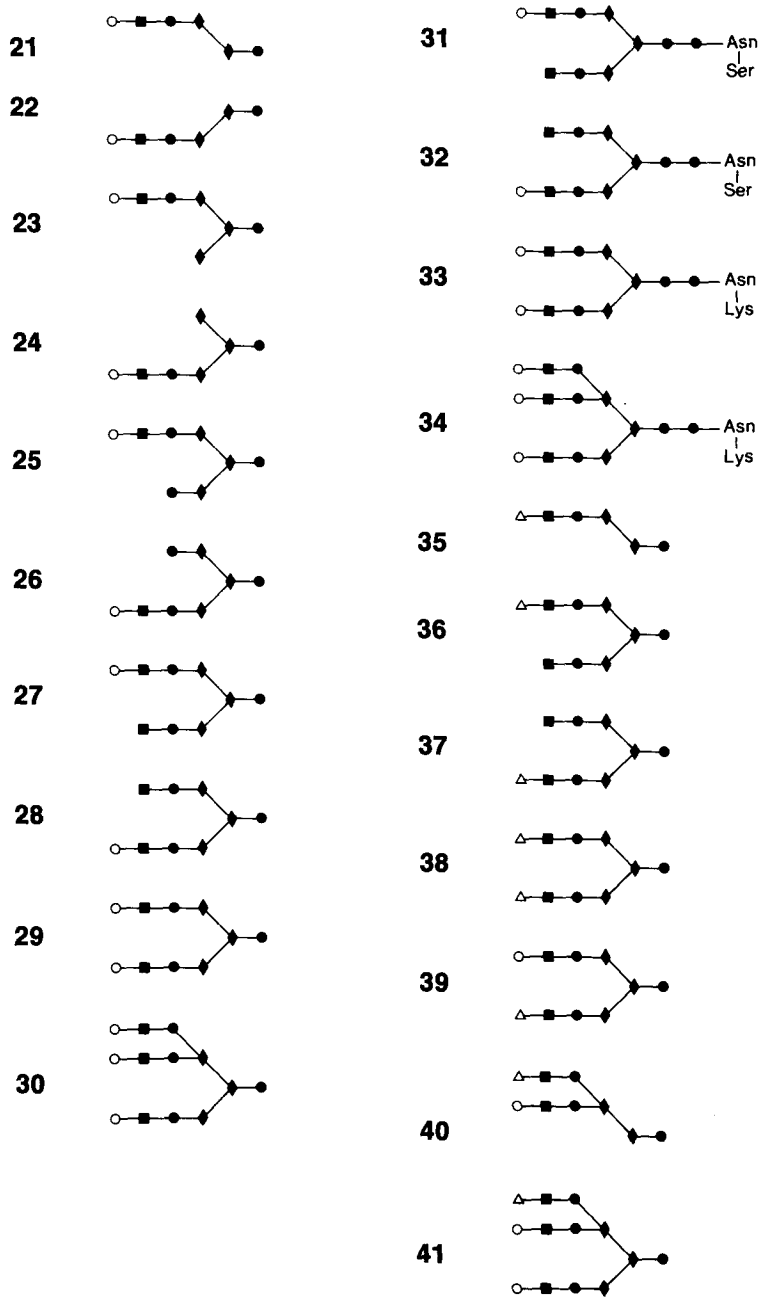
(i) NeuAc α-(2→6)-Linked to Gal.—Compound **21** is a mono-antennary, upper-branch, sialo oligosaccharide. It was available as the main component (90%) of a mixture of two oligosaccharides also containing **22** that has been isolated from the urine of a sialidosis patient.[52,72] The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture, recorded at pD 7 (as for all of these sialo compounds, except **60**) is shown in Fig. 20, and the n.m.r. parameters for **21** are listed in Table IX.

Comparison with the spectrum of **17** demonstrates that the attachment of a NeuAc residue α-(2→6)-linked to Gal gives some significant, additional signals in the spectrum, and causes some changes in chemical shifts of structural-reporter groups of other constituent monosac-

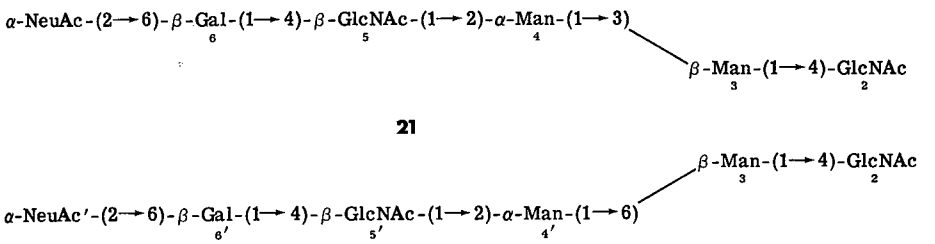
(72) L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, G. Strecker, J.-C. Michalski, B. Fournet, G. Spik, and J. Montreuil, *Eur. J. Biochem.*, 87 (1978) 323–329.

FIG. 19.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **19** and **20** in the Ratio of 1:3. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. (Anomeric-proton signal designated as 4′, **19**α means: H-1 of Man-4′ in the α anomer of compound **19**.) Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other part of the spectrum, as indicated.]

CHART 2. Symbols Employed[a] for Compounds 21–41



[a] For the key to the symbolic notation, see Chart 1.



21

22

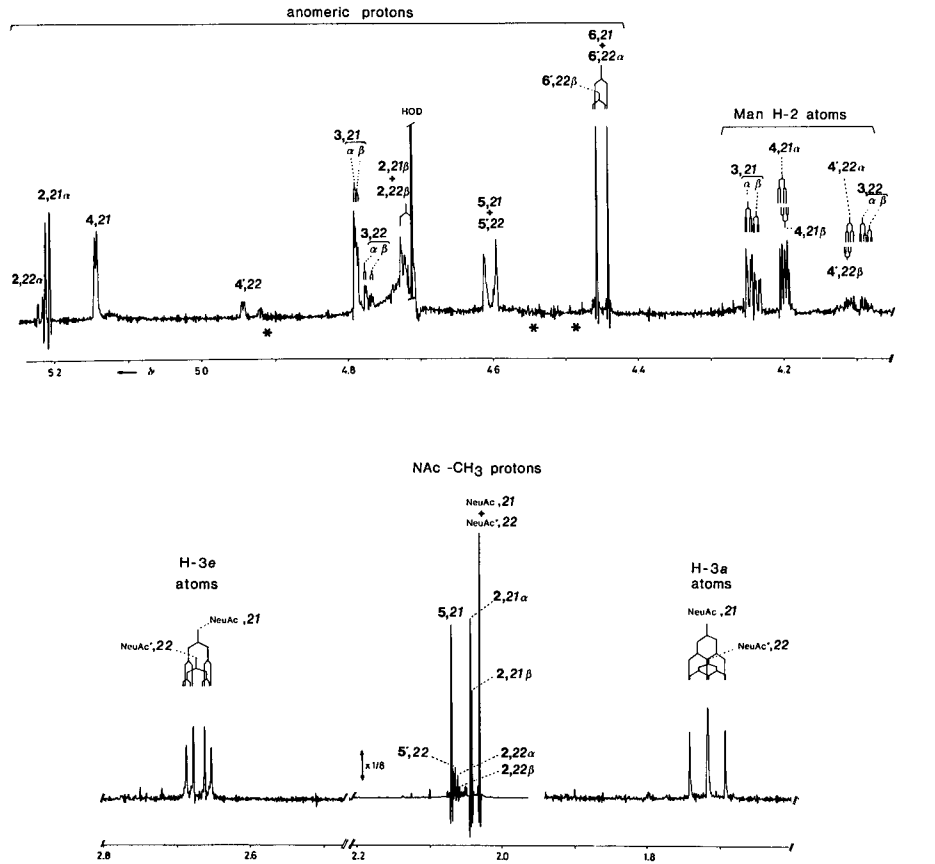
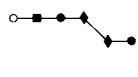
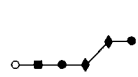


FIG. 20.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **21** and **22** in the Ratio of 9:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The sample was contaminated with a very small proportion of the asialo analog of **22**, that is, compound **18**, as can be inferred from the signals marked by asterisks (compare Fig. 18b).]

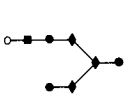
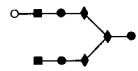
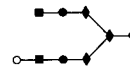
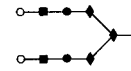
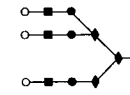
[1]H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Sialic Acid (Com

			Compound and			
			21	22	23	24
Reporter group	Residue	Anomer of oligosaccharide				
H-1 of	2	α	5.208	5.218	5.215	5.215
		β	~4.72	~4.72	~4.72	~4.72
	3	α	4.788	4.774	4.792	4.784
		β	4.785	4.765	4.784	4.775
	4	α	5.143	—	5.140	5.107
		β				5.103
	4′	α	—	4.942	4.921	4.949
		β				
	5	α,β	4.602	—	4.606	—
	5′	α,β	—	4.602	—	4.606
	6	α,β	4.446	—	4.446	—
	6′	α	—	4.446	—	4.446
		β		4.449		4.450
	7	α,β	—	—	—	—
	8	α,β	—	—	—	—
H-2 of	3	α	4.244	4.088	4.268	4.263
		β	4.233	4.078	4.258	4.253
	4	α	4.198	—	4.198	4.073
		β	4.196			
	4′	α	—	4.106	<4.0	4.119
		β		4.110		
H-3 of	4	α,β	<4.0	—	<4.0	<4.0
	4′	α,β	—	<4.0	<4.0	<4.0
H-3*a* of	[a]NeuAc	α,β	1.716	—	1.719	—
	[b]NeuAc′	α,β	—	1.714	—	1.718
	[c]NeuAc*	α,β	—	—	—	—
H-3*e* of	NeuAc	α,β	2.670	—	2.669	—
	NeuAc′	α,β	—	2.672	—	2.672
	NeuAc*	α,β	—	—	—	—
NAc of	2	α	2.043	2.061	2.059	2.063
		β	2.041	2.058	2.055	2.059
	5	α,β	2.069	—	2.070	—
	5′	α	—	2.064	—	2.070
		β				2.067
	7	α,β	—	—	—	—
	NeuAc	α,β	2.030	—	2.030	—
	NeuAc′	α,β	—	2.030	—	2.030
	NeuAc*	α,β	—	—	—	—

[a] NeuAc denotes NeuAc linked to Gal-6. [b] NeuAc′ denotes NeuAc linked to Gal-6′.

E IX

Oligosaccharides of the *N*-Acetyllactosamine Type Containing Only α-(2→6)-Linked, pounds 21–30)

schematic structure					
25	**26**	**27**	**28**	**29**	**30**
					
5.213	5.213	5.214	5.214	5.216	5.215
~4.72	~4.72	~4.72	~4.72	~4.72	~4.72
4.784	4.784	4.781	4.781	4.786	~4.77
4.776	4.776	4.771	4.771	4.777	~4.76
5.137	5.121	5.138	5.124	5.137	5.134
4.921	4.950	4.929 4.932	4.949	4.952	4.943
4.607	4.557	4.608	4.583	4.608	4.594
4.558	4.607	4.586	4.608	4.608	4.605
4.446	—	4.446	4.468	4.445	4.443
—	4.446	4.471	4.446	4.445	4.443
	4.450	4.473	4.450	4.449	4.447
—	—	—	—	—	4.573
—	—	—	—	—	4.441
4.264	4.259	4.260	4.265	4.266	4.228
4.253	4.248	4.249	4.254	4.255	4.217
4.198	4.190	4.197	4.193	4.199	4.221
4.109	4.115	4.114	4.118	4.121	4.119
<4.0	<4.0	<4.0	<4.0	<4.0	4.049
<4.0	<4.0	<4.0	<4.0	<4.0	<4.0
1.718	—	1.720	—	1.721	1.720
—	1.718	—	1.719	1.719	1.717
—	—	—	—	—	1.706
2.669	—	2.669	—	2.669	2.670
—	2.671	—	2.672	2.672	2.672
—	—	—	—	—	2.670
2.059	2.063	2.061	2.063	2.063	2.062
2.055	2.059	2.058	2.060	2.060	2.059
2.070	2.055	2.070	2.053	2.071	2.069
2.053	2.070	2.050	2.069	2.069	2.067
	2.066	2.048	2.066	2.066	2.065
—	—	—	—	—	2.102
2.030	—	2.031	—	2.031	2.030
—	2.030	—	2.031	2.031	2.030
—	—	—	—	—	2.028

[c] NeuAc* denotes NeuAc linked to Gal-8.

charides. Typical resonances for sialic acid, which do not coincide with the bulk, are those of H-3*a* (δ 1.716) and H-3*e* (δ 2.670), and of the *N*-acetyl protons (δ 2.030). The signals of the H-3 atoms have characteristic patterns: H-3*a* gives a triplet ($|J_{3a,4}| = |J_{3a,3e}|$), and H-3*e*, a doublet of doublets. The configuration of the linkage between NeuAc and Gal may be derived from the chemical shift of H-3*e* (in β-linkage δH-3*e* = 2.2–2.4 p.p.m.).[73–75] The chemical shifts of H-3*a* and H-3*e* are sensitive to changes in the pD of the oligosaccharide solution in D_2O. At pD 2, δH-3*a* ~1.79 and δH-3*e* ~2.66 (see Ref. 75).

The introduction of a sialic acid group at O-6 of Gal-6 gives rise to a shift decrement for H-1 of Gal-6 (Δδ −0.022 p.p.m.) and to shift increments for H-1 of GlcNAc-5 (Δδ 0.025 p.p.m.), the *N*-acetyl protons of GlcNAc-5 (Δδ 0.016 p.p.m.), H-1 of Man-4 (Δδ 0.016 p.p.m.), and H-2 of Man-4 (Δδ 0.005 p.p.m.), whereas the line widths of their signals, and the anomerization effects upon these, are unaffected, as compared to **17**.

These changes in chemical shifts, in combination with the chemical-shift values of the H-3 signals of NeuAc, are specific for the type and configuration of the linkage of NeuAc to Gal-6.

Compound **22** is a mono-antennary, lower-branch, sialo oligosaccharide, isolated in admixture with **21** from the urine of a patient with sialidosis. The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture is given in Fig. 20; the set of lower-intensity signals belongs to **22**. The n.m.r. parameters for **22** are summarized in Table IX. For the interpretation of the n.m.r. spectrum, comparison with the spectra of **18** and **21** is appropriate.

The chemical shifts of the H-3 signals of the NeuAc′ residue, that is, NeuAc α-(2→6)-linked to Gal-6′, in **22** differ slightly from those of NeuAc α-(2→6)-linked to Gal-6 in **21** (δH-3*e* 2.672 as against 2.670, and δH-3*a* 1.714 as against 1.716). The δ value of the *N*-acetyl protons of NeuAc′ in **22** exactly matches that for NeuAc in **21**.

The observed shift increments and decrements, due to the introduction of sialic acid, for structural-reporter groups of neighboring residues are of the same order of magnitude as for the step from **17** to **21**, but larger for H-1 of Man-4′ than for H-1 of Man-4. In this case, also

(73) J. Haverkamp, L. Dorland, J. F. G. Vliegenthart, J. Montreuil, and R. Schauer, *Abstr. Int. Symp. Carbohydr. Chem., 9th*, London, 1978, 281–282.

(73a) J. F. G. Vliegenthart, L. Dorland, H. van Halbeek, and J. Haverkamp, in R. Schauer (Ed.), *Sialic Acids—Chemistry, Metabolism and Function, Cell Biol. Monogr.*, Vol. 10, Springer, Wien, 1982, pp. 127–172.

(74) J. F. G. Vliegenthart, in L. Svennerholm, H. Dreyfus, and P. F. Urban (Eds.), *Structure and Function of Gangliosides, Adv. Exp. Med. Biol.*, 125 (1980) 77–91.

(75) J. Haverkamp, H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, R. Pfeil, and R. Schauer, *Eur. J. Biochem.*, 122 (1982) 305–311.

the resonance position of H-1 of Man-3 is slightly, but significantly, influenced by the extension of **18** with NeuAc: $\Delta\delta$ 0.005 p.p.m. for H-1 of Man-3 in the α and β anomer of **22**. The observation that the effects of NeuAc in the lower branch reach farther than in the upper branch supports the proposal that these two branches have different conformations (compare oligosaccharides **7**, **10–12**, and **17** as against **18**, and **19** as against **20**).

Concerning anomerization effects, it is noteworthy that differences in chemical shifts of corresponding protons in the α and β anomer of **22** remain unchanged in comparison to **18**, except for the H-1 signals of Man-3 and Man-4′. The $\Delta\delta_{\alpha-\beta}$ for these protons is (for Man-3) almost, or (for Man-4′) completely, nullified by attachment of the NeuAc group.

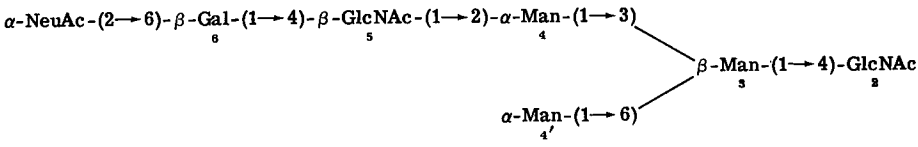
Compound **23** is the sialo analog of **19**, bearing NeuAc in α-(2→6) linkage to Gal-6. Compound **24** is the sialo analog of **20**, having NeuAc′ in α-(2→6) linkage to Gal-6′. The oligosaccharides **23** and **24** could be obtained as a mixture from the urine of a patient with sialidosis[72] and from new-born, human meconium.[76] The 500-MHz, ^{1}H-n.m.r. spectrum of such a mixture containing **23** and **24** in the ratio of 3:2 is presented in Fig. 21. The n.m.r. parameters for both compounds are summarized in Table IX.

The attachment of NeuAc to Gal-6 of **19** or to Gal-6′ of **20** gives rise to typical sets of H-3 and *N*-acetyl signals for NeuAc, and to alterations of chemical shifts of structural-reporter groups of neighboring residues, as described for **21** and **22**. In **23**, no effect is observed on the chemical shift of H-1 of Man-4′ in comparison to **19**. Similarly, the signals of H-1 and H-2 of Man-4 in **24** occupy the same positions as in the spectrum of **20**. In contrast to the situation for the asialo analog, the H-1 doublet of Man-4′ of **23** is not doubled due to anomerization; however, its line width is not affected.

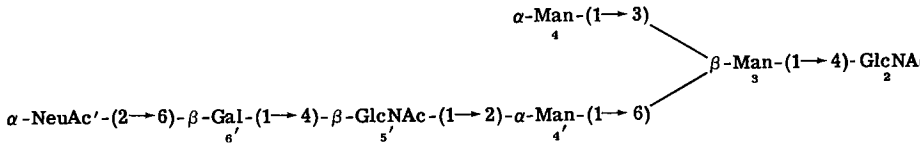
Compounds **25** and **26** are reducing oligosaccharides having incomplete, diantennary structures. They were isolated from new-born, human meconium[76] in a mixture containing **25** and **26** in the ratio of 3:1. The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture is shown in Fig. 22, and the n.m.r. data for both compounds are compiled in Table IX.

For the interpretation of their n.m.r. spectra, compounds **25** and **26** are considered to be extensions of **23** and **24** with a GlcNAc residue in β-(1→2) linkage to Man-4′ or -4, respectively. The H-1 atoms and the *N*-acetyl-group protons for the terminal GlcNAc residues 5′ (compound **25**) and 5 (compound **26**) resonate at positions similar to those

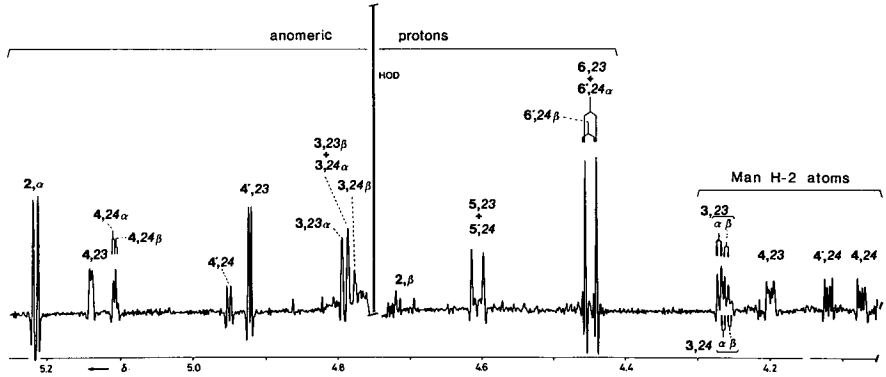
(76) M.-C. Herlant-Peers, J. Montreuil, G. Strecker, L. Dorland, H. van Halbeek, G. A. Veldink, and J. F. G. Vliegenthart, *Eur. J. Biochem.*, 117 (1981) 291–300.

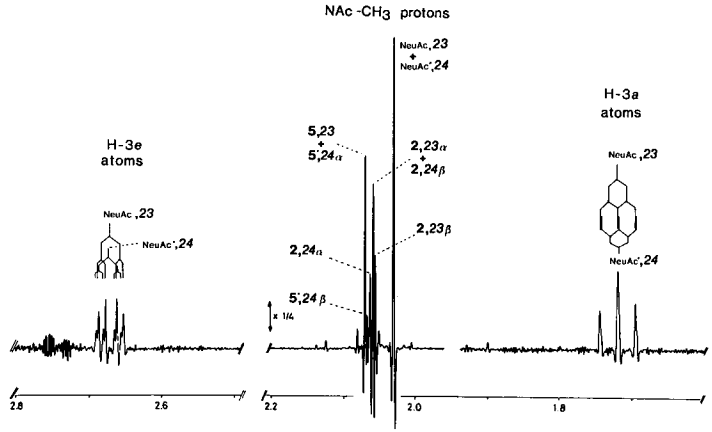

α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-Man-(1→6)
4'

23


α-Man-(1→3)
4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc'-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'

24


anomeric
protons
HOD
6,23 + 6',24α
6',24β
3,23β + 3,24α
3,24β
2,α
4,24α
4,24β
4,23
4',23
3,23α
4',24
5,23 + 5',24
2,β
Man H-2 atoms
3,23 α β
4,23
4',24
4,24
3,24 α β
5.2
δ
5.0
4.8
4.6
4.4
4.2


NAc-CH3 protons
NeuAc,23 + NeuAc',24
H-3e atoms
H-3a atoms
5,23 + 5',24α
2,23α + 2,24β
2,23β
2,24α
5',24β
x 1/4
NeuAc,23
NeuAc',24
2.8
2.6
2.2
2.0
1.8

for compound **6**. The attachment of GlcNAc-5′ to O-2 of Man-4′ has hardly any effect on the chemical shift of H-1 of Man-4′, but the H-2 signal shifts downfield. The observed chemical shifts are in accord with those of **6**.

The chemical shifts of H-1 and H-2 of Man-4 in **26** are practically identical to those for **6**; they are different from the values for **24**, owing to the presence of GlcNAc-5 (compare the positions of H-1 and H-2 of Man-4 in **20** and **6**).

The presence of GlcNAc-5′ is revealed by small shift-decrements of H-1 of Man-3 in both anomers of **25**, whereas GlcNAc-5 does not affect the chemical shift of H-1 of Man-3 in either the α or the β anomer of **26**.

Compounds **27** and **28** are diantennary, monosialo oligosaccharides that could be obtained in a mixture from the urine of a patient with sialidosis,[72] as well as from new-born, human meconium.[76] The 500-MHz, ^{1}H-n.m.r. spectrum of such a mixture containing **27** and **28** in the ratio of 11:9 is presented in Fig. 23. The n.m.r. data are listed in Table IX.

For the spectral interpretation, the oligosaccharides are conceived as extensions of **25** and **26**, with a Gal residue β-(1→4)-linked to GlcNAc-5′ or -5, respectively. As already described, this attachment almost exclusively influences the chemical shift of H-1 of the penultimate GlcNAc residue ($\Delta\delta$ 0.027 p.p.m.). The resonance positions of the H-1 signals of the terminal Gal residues are in complete accord with those of the diantennary, asialo oligosaccharide **7**.

The H-1 signal of Gal-6′ in **27** is clearly doubled, due to anomerization, an effect that is also observable for the residues Man-4′ (H-1) and GlcNAc-5′ (*N*-acetyl protons).

The location of NeuAc in a diantennary, monosialo structure may be directly inferred from the chemical-shift values of the H-1 signals of Man-4 and -4′. The chemical shifts of the *N*-acetyl signals of GlcNAc-5 and -5′ also reflect whether or not a certain branch is terminated with a NeuAc residue in α-(2→6) linkage to Gal. In principle, the chemical

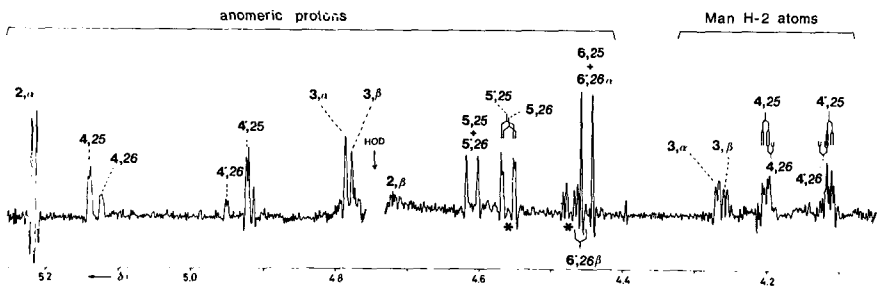
FIG. 21.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **23** and **24** in the Ratio of 3:2. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta$ = 2:1. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

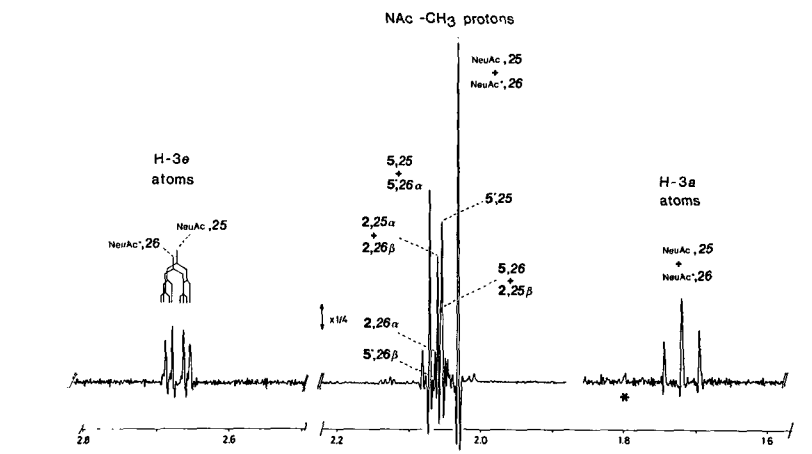
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-GlcNAc-(1→2)-α-Man-(1→6)
5′ 4′

25

β-GlcNAc-(1→2)-α-Man-(1→3)
5 4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc′-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′

26





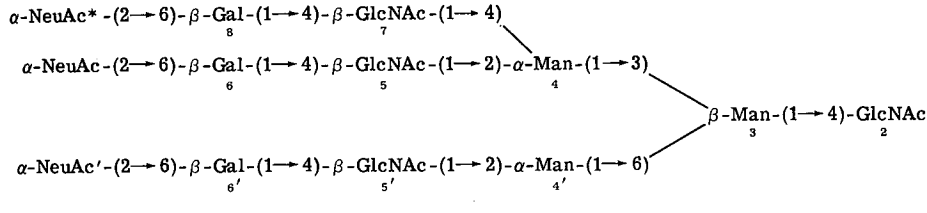
shifts of the H-1 signals of Gal-6 and -6′, and those of the H-3 signals of NeuAc itself, can also be utilized for the locating of NeuAc.

Compound **29** is a diantennary, sialo oligosaccharide that has been isolated from the urine of a sialidosis patient,[72] as well as from newborn, human meconium.[76] The 500-MHz, ^{1}H-n.m.r. spectrum of **29** is depicted in Fig. 24, and the n.m.r. data are compiled in Table IX.

The chemical shifts of the structural-reporter groups of the sialylated branches of compounds **27** and **28** are found unchanged in this spectrum. This shows that the shift effects introduced by the extension of a branch with sialic acid in α-(2→6) linkage are limited to the branch to which it is attached.

The pattern of the signals of the H-1 and H-2 resonances of Man residues 3, 4, and 4′ is, as a whole, typical of a diantennary, sialo structure wherein the chemical shifts of the H-1 signals of Man-4 and -4′ indicate that each branch is terminated with NeuAc in α-(2→6) linkage to Gal.

The NeuAc residues are not completely equivalent, as is evident from the differences in the chemical shifts of their H-3*e* as well as their H-3*a* signals. It should be noted that this doubling of signals is not caused by anomerization, as is clear from the 1 : 1 intensity ratio of the signals. The difference in chemical shift between the H-1 signals of Gal-6 and -6′ is slightly smaller than for the asialo analog **7**.



30

Compound **30** is a triantennary, sialo oligosaccharide terminated with NeuAc α-(2→6)-linked to Gal in the three branches. It has been

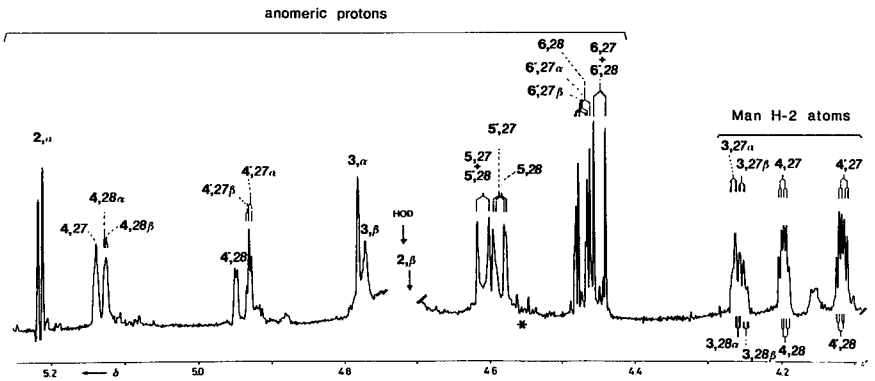
FIG. 22.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **25** and **26** in the Ratio of 3 : 1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow. The sample was contaminated with very small proportions of **27** and **28**, and, to a very small extent, with an oligosaccharide similar to **25**, **26**, **27**, or **28**, containing a NeuAc group in α-(2→3), instead of α-(2→6), linkage to Gal. This can be inferred from the signals marked by asterisks.]

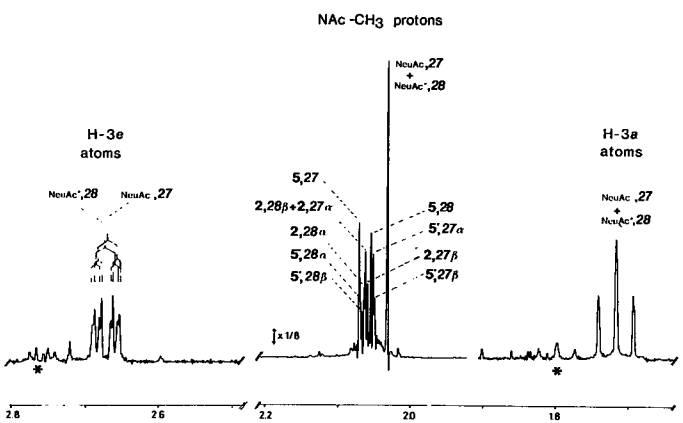
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′

27

β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc′-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′

28



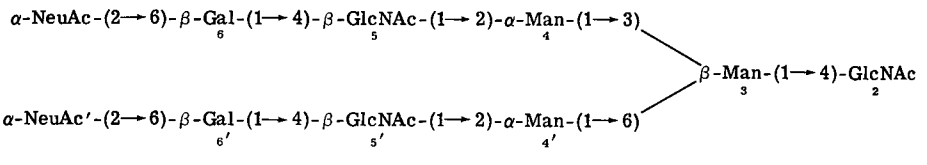


obtained from the urine of a patient suffering from sialidosis, as a minor constituent of a mixture of two oligosaccharides containing **41** as the major part.[13] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 32. The n.m.r. parameters for **30** are compiled in Table IX.

For the interpretation of the signals of lower intensity in the spectrum, **30** is considered to be an extension of **10** with three α-(2→6)-linked NeuAc residues. The introduction of NeuAc groups α-(2→6)-linked to Gal-6 and -6′ influences the δ values of the structural-reporter groups of the upper-branch residues 6, 5, and 4, and the lower-branch residues 6′, 5′, and 4′ in a way similar to that already described (see Tables IV, VIII, and IX). The attachment of NeuAc*, that is, NeuAc α-(2→6)-linked to Gal-8, is revealed in a shift decrement for H-1 of Gal-8 ($\Delta\delta$ −0.022 p.p.m.) with respect to **10**, and an increment for H-1 of GlcNAc-7 ($\Delta\delta$ 0.027 p.p.m.) (compare Table XIII). The shift of the *N*-acetyl singlet of GlcNAc-7 from δ 2.078 for **10** to δ 2.102 for **30** ($\Delta\delta$ 0.024 p.p.m.) is highly characteristic for the elongation of **10** with NeuAc* α-(2→6)-linked to Gal-8. No additional effects of this NeuAc* upon δH-1 and δH-2 of the more remote Man-4 and Man-3 are observed.

The most striking feature of the 500-MHz, ^{1}H-n.m.r. spectrum of compound **30** is the occurrence of three well-resolved, H-3*a* triplets of NeuAc at δ 1.720, 1.717, and 1.706. The first is assigned to NeuAc α-(2→6)-linked to Gal-6, and the second, to NeuAc′ linked to Gal-6′, in agreement with **29**. Consequently, the third signal, at δ 1.706, must be ascribed to H-3*a* of NeuAc* in the additional (third) branch of **30**. Based on the relative intensities of the H-3*e* doublets of doublets, the H-3*e* signal of this NeuAc* group apparently coincides with the doublet of doublets at δ 2.670 stemming from NeuAc linked to Gal-6, as H-3*e* of the lower-branch NeuAc′ resonates at δ 2.672. Finally, it is remarkable that the NeuAc* group linked to Gal-8 possesses *N*-acetyl

FIG. 23.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **27** and **28** in the Ratio of 11:9. [The bold numbers in the spectrum refer to the corresponding residues in the structures, the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta$ = 2:1. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of GlcNAc-2 for the β anomers of **27** and **28**, have been omitted from the spectrum; their positions are indicated by arrows. The sample was contaminated with a small proportion of **36** or **37**, or both, containing α-(2→3)-linked NeuAc, as can be inferred from the signals marked by asterisks (compare Fig. 28).]



29

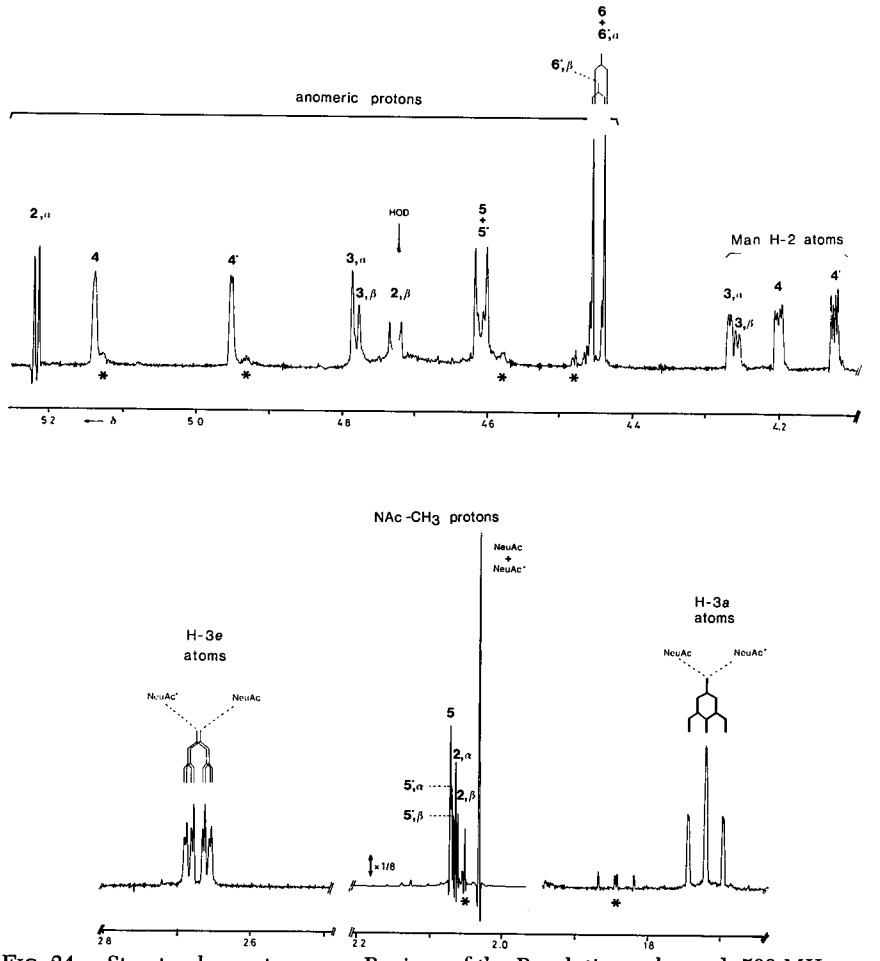


FIG. 24.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **29**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **29**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow. The sample was con-

protons in a chemical environment different from that of the corresponding protons in the other two NeuAc groups (δ 2.028 as against 2.030).

It must be emphasized that the extension of a triantennary structure with a NeuAc group α-(2→6)-linked to Gal-8 can be unambiguously established on the basis of both the unique set of chemical shifts for its H-3*a*, H-3*e*, and *N*-acetyl protons, as well as its specific effects upon the chemical shifts of the structural-reporter groups of GlcNAc-7 and Gal-8 (compare Table XIII).

Compounds **31** and **32** are diantennary, monosialo glycopeptides that may be obtained from numerous sialo-glycoproteins, for example, from rabbit serotransferrin.[77] Compounds **31** and **32** can occur in different ratios in glycopeptide mixtures thus prepared. It is unclear whether these monosialo compounds reflect partial degradation, or incomplete biosynthesis; the origin of this microheterogeneity is still under investigation. It should be noted that the preparation *in vitro* of the Asn-Lys analog of the monosialo glycopeptide **31**, using **8** as the substrate, by means of a β-D-galactoside α-(2→6)-sialyltransferase (EC 2.4.99.1) may be efficiently monitored by high-resolution, ^{1}H-n.m.r. spectroscopy.[78] The spectral parameters for **31** and **32** are listed in Table X.

With regard to the spectral interpretation, it may be mentioned that **31** and **32** are the glycopeptide analogs of oligosaccharides **27** and **28**, respectively. The presence of α-(2→6)-linked NeuAc in one of the branches can be derived from the chemical shifts of the anomeric protons of Man-4 and -4′, or, alternatively, from those of the *N*-acetyl protons of GlcNAc-5 and -5′, as already described.

The chemical shifts of H-1 and of the *N*-acetyl protons of GlcNAc-1 deviate from those for **8** (see Table V) and **33** (see Table X), illustrating the effect of the peptide moiety (Asn-Ser as against Asn-Lys) on the chemical shifts of the structural-reporter groups of this residue.

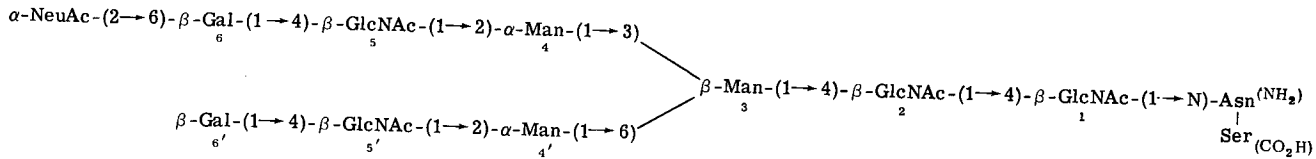
Compound **33** is a diantennary, sialo glycopeptide that can be obtained from various glycoproteins[1–4]; for example, from human sero-

(77) D. Leger, V. Tordera, G. Spik, L. Dorland, J. Haverkamp, and J. F. G. Vliegenthart, *FEBS Lett.*, 89 (1978) 149–152.

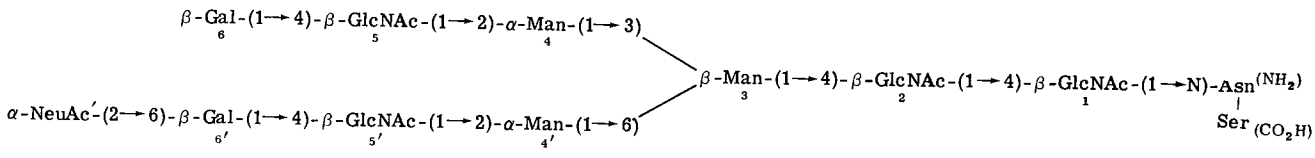
(78) D. H. van den Eijnden, D. H. Joziasse, L. Dorland, H. van Halbeek, J. F. G. Vliegenthart, and K. Schmid, *Biochem. Biophys. Res. Commun.* 92 (1980) 839–845.

---

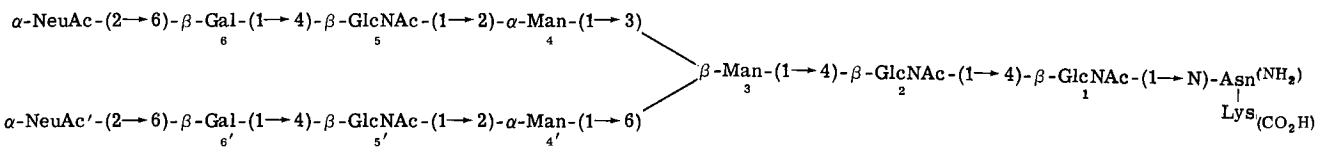
taminated with small proportions of the monosialo oligosaccharides **27** and **28** (or even with the asialo compound **7**), as well as the corresponding proportion of free NeuAc. This is evident from the signals marked by asterisks. Special attention should be paid to the chemical shifts of the *N*-acetyl signal (δ 2.050) and the H-3*a* signal (δ 1.838), and to the typical doublet of doublets for the latter proton, of free NeuAc (β anomer).]



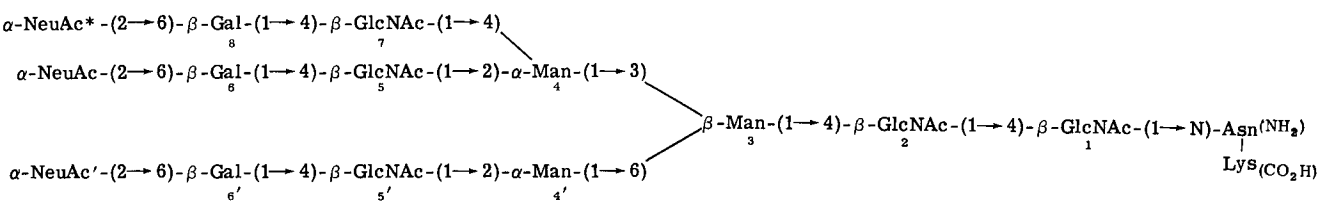
31



32



33



34

transferrin,[72] and from α_2HS glycoprotein.[79] Structural investigations employing 360-MHz, ^{1}H-n.m.r. spectroscopy have been conducted on such a glycopeptide isolated from rat-liver, plasma membrane.[80] The 500-MHz, ^{1}H-n.m.r. spectrum of **33** from serotransferrin is presented in Fig. 25, and its n.m.r. parameters are compiled in Table X.

For the interpretation of the spectrum, the compound may be conceived of as a glycopeptide analog of oligosaccharide **29**, which allows a straightforward assignment of all signals of structural-reporter groups. The resonances of the H-3*a* atoms of the two NeuAc groups coincide, whereas those of the H-3*e* atoms are well separated. The assignment of the latter signals was achieved on the basis of the n.m.r. spectrum of the enzymically prepared,[78] diantennary, monosialo glycopeptide (Asn-Lys) having the NeuAc group α-(2→6)-linked to Gal-6 (compare **31**). Interestingly, the *N*-acetyl signals of the two NeuAc groups in **33** are observed separately (δ 2.029 and 2.028), in contrast to the corresponding signals of oligosaccharides **29** and **30**.

Compound **34** is a triantennary, sialo glycopeptide terminated with NeuAc α-(2→6)-linked to Gal in all three branches; it could be obtained by exhaustive, *in vitro* sialylation of compound **9** with β-D-galactoside α-(2→6)-sialyltransferase.[13,78] The Gln-Asn analog of compound **34**, in admixture with compound **53** and a glycopeptide analog of **41**, has been isolated from human-plasma ceruloplasmin (see compound **53**). The 500-MHz, ^{1}H-n.m.r. spectrum of **34** is given in Fig. 26, and its pertinent spectral parameters are listed in Table X.

Compared to the spectrum of **33**, an additional set of NeuAc H-3 signals is present, at δ 1.706 (H-3*a*) and δ 2.670 (H-3*e*), which is specific for NeuAc* in α-(2→6) linkage to Gal-8. This set has also been observed in the spectrum of the oligosaccharide analog of **34**, namely, **30**. The chemical shift of the *N*-acetyl singlet of GlcNAc-7 (δ 2.101), in combination with its relatively large line-width, is also characteristic for an O-6 substitution of Gal-8 by a NeuAc group (compare **30**).[13] All other spectral features of **34** are in agreement with those described for **30**.

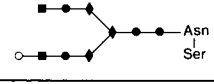
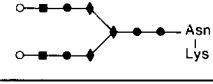
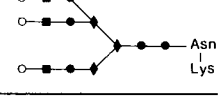
The difference in chemical shift between the H-1 signals of Gal-6 and Gal-8 is significantly lessened, as compared to that for the asialo analog **9** ($\Delta\delta$ 0.001 as against 0.006 p.p.m.). This may be attributable to a small, secondary, downfield-shift effect, induced on δH-1 of Gal-8, by the attachment of NeuAc to Gal-6 in α-(2→6) linkage. Furthermore, it

(79) M. Endo, K. Hoare, K. Schmid, H. van Halbeek, L. Dorland, and J. F. G. Vliegenthart, *J. Biol. Chem.*, **258** (1983) in press; K. Schmid, M. Endo, H. van Halbeek, L. Dorland, and J. F. G. Vliegenthart, *Fed. Proc.*, 40 (1981) 1598.

(80) H. Debray, B. Fournet, J. Montreuil, L. Dorland, and J. F. G. Vliegenthart, *Eur. J. Biochem.*, 115 (1981) 559–563.

TABLE X

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glycopeptides of the *N*-Acetyllactosamine Type Containing Only α-(2→6)-Linked Sialic Acid (Compounds 31–34)

		Compound and schematic structure			
		31[a]	32[a]	33	34
					
Reporter group	**Residue**				
H-1 of	1	5.073	5.073	5.088	5.060
	2	4.620	4.620	4.616	4.615
	3	4.769	4.769	4.773	~4.76
	4	5.138	5.120	5.133	5.133
	4′	4.928	4.944	4.949	4.936
	5	4.599	4.577	4.603	4.594
	5′	4.577	4.599	4.603	4.602
	6	4.445	4.470	4.442	4.440
	6′	4.472	4.445	4.447	4.448

	7	—	—	—	4.571
	8	—	—	—	4.439
H-2 of	3	4.251	4.251	4.254	4.220
	4	4.192	4.192	4.195	4.225
	4′	4.112	4.112	4.116	4.116
H-3 of	4	<4.0	<4.0	<4.0	4.050
	4′	<4.0	<4.0	<4.0	<4.0
H-3*a* of	NeuAc	1.716	—	1.716	1.717
	NeuAc′	—	1.716	1.716	1.717
	NeuAc*	—	—	—	1.706
H-3*e* of	NeuAc	2.670	—	2.666	2.670
	NeuAc′	—	2.670	2.672	2.674
	NeuAc*	—	—	—	2.670
NAc of	1	2.005	2.005	2.002	2.008
	2	2.078	2.078	2.081	2.082
	5	2.069	2.050	2.067	2.069
	5′	2.046	2.067	2.063	2.065
	7	—	—	—	2.101
	NeuAc	2.030	—	2.029[b]	2.030
	NeuAc′	—	2.030	2.028[b]	2.030
	NeuAc*	—	—	—	2.030

[a] Chemical shifts measured at 360 MHz, T = 298 K. [b] Assignments may have to be interchanged.

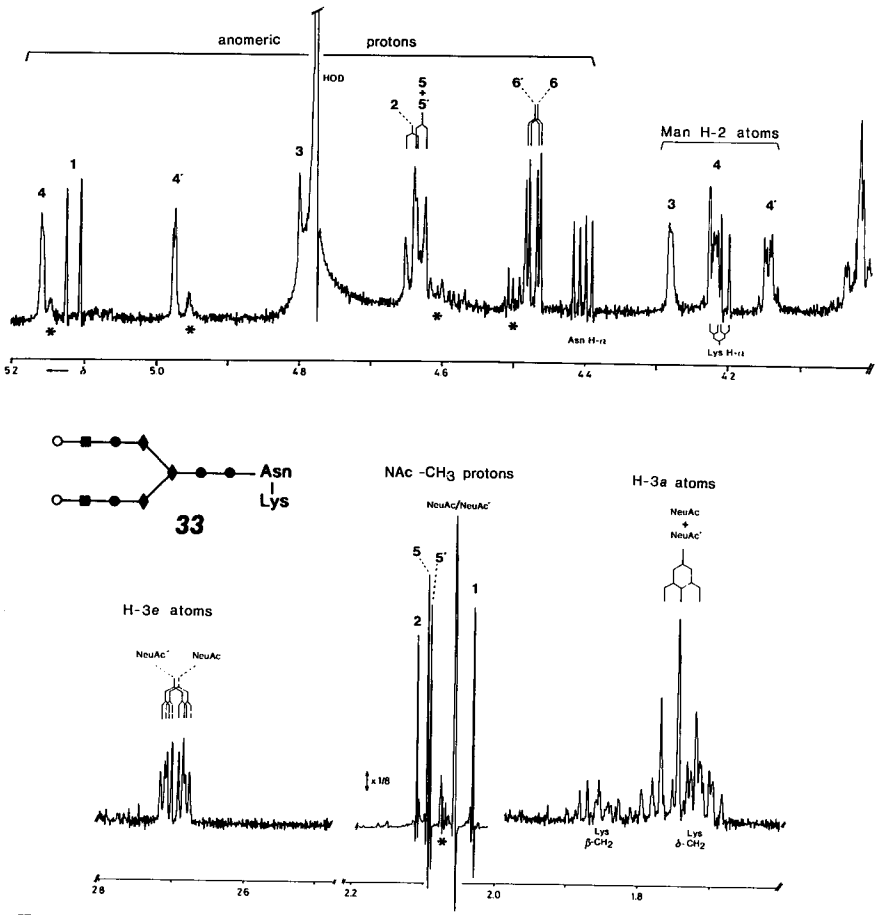


FIG. 25.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **33**. [The bold numbers in the spectrum refer to the corresponding residues in the structure (see page 284). The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The sample contained small proportions of the monosialo analogs of **33**, that is, **31** and **32** (or even of the asialo glycopeptide **8**), as can be inferred from the signals marked by asterisks.]

is remarkable that the H-1 signal of Man-4′ undergoes an upfield shift upon extension of the diantennary, sialo glycopeptide **33** with the third branch. The origin of this effect is not yet clear; the third branch may exert a direct, through-space action on Man-4′, or it may, more probably [compare the partly sialylated triantennary glycopeptide, lacking NeuAc′ in comparison to **34**: δH-1 of Man-4′ = 4.927 (Ref.

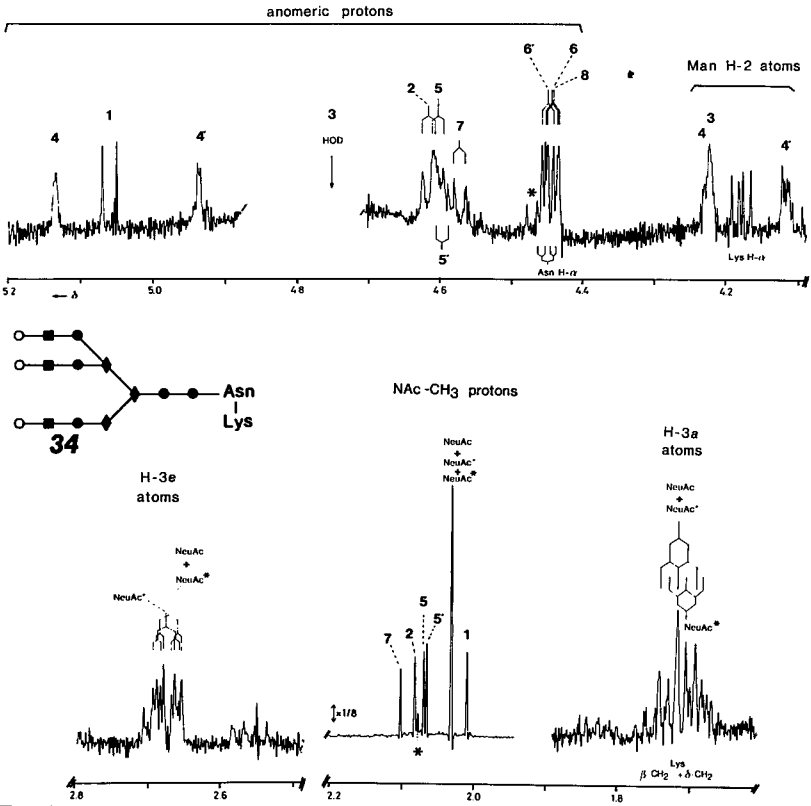
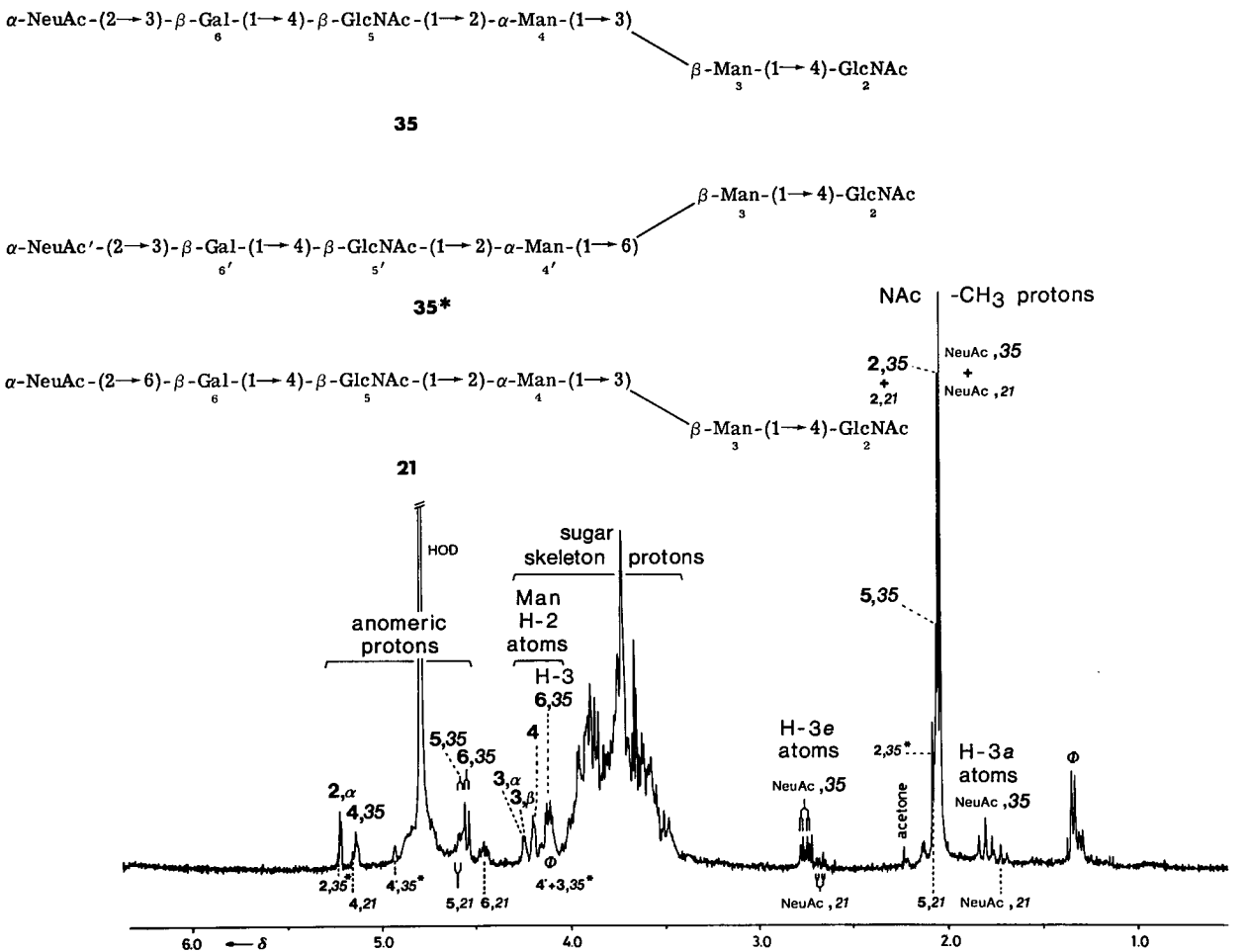


FIG. 26.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **34**. [The bold numbers in the spectrum refer to the corresponding residues in the structure (see page 284). The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from the spectrum; their position is indicated by an arrow. From the signals marked by asterisks, especially from the low-intensity *N*-acetyl-proton singlet of GlcNAc-7 at δ 2.077, it can be inferred that the sample contained a small proportion of the disialo analog of **34** that lacks the NeuAc* residue.]

78)], lessen the influence of NeuAc′ attached to Gal-6′ on this Man residue.

(ii) NeuAc α-(2→3)-Linked to Gal.—Compound **35** is a mono-antennary, upper-branch, sialo oligosaccharide. It was available as the main component (75%) of a mixture of at least three isomeric oligosaccharides also containing the lower-branch analog of **35**, namely, αNeuAc(2→3)βGal(1→4)βGlcNAc(1→2)αMan(1→6)βMan(1→4)Glc-


α-NeuAc-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Man-(1→4)-GlcNAc
35
α-NeuAc'-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-GlcNAc
35*
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Man-(1→4)-GlcNAc
21
NAc -CH3 protons
2,35 + 2,21
NeuAc ,35 + NeuAc ,21
HOD
anomeric protons
sugar skeleton protons
Man H-2 atoms
H-3 6,35
5,35
4
6,35
3,α
3,β
2,α
4,35
H-3e atoms
NeuAc ,35
2,35*
acetone
H-3a atoms
NeuAc ,35
Φ
2,35*
4,21
4',35*
5,21
6,21
4'+3,35*
NeuAc ,21
5,21
NeuAc ,21
6.0 δ 5.0 4.0 3.0 2.0 1.0

NAc (7%) (compound **35***),[80a] and oligosaccharide **21**, differing from **35** in the type of linkage between NeuAc and Gal (12%). Small proportions of the corresponding asialo analogs (compound **17** or **18**, or both) were also detected in the sample. This mixture was isolated from the urine of a patient with sialidosis.[72] The 360-MHz, ^{1}H-n.m.r. spectrum of the mixture is given in Fig. 27; the n.m.r. parameters for **35** are listed in Table XI.

Comparison with the spectrum of **17** demonstrates that the attachment of a NeuAc group in α-(2→3) linkage to Gal affords three NeuAc structural-reporter-group signals. The chemical shifts of H-3*a* (δ 1.799) and H-3*e* (δ 2.758) are both clearly discernible from the corresponding ones for NeuAc in α-(2→6) linkage to Gal (**21**), making these parameters excellently suited for determination of the type (and configuration) of the NeuAc→Gal linkage, but the *N*-acetyl protons of NeuAc in α-(2→3) linkage resonate at the same position as those of NeuAc α-(2→6)-linked to Gal (compare Tables XIII and XIV).

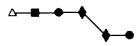
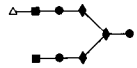
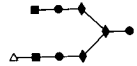
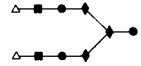
As compared to the spectrum of the asialo analog **17**, the introduction of NeuAc at O-3 of Gal-6 gives rise to a limited number of effects on the chemical shifts of structural-reporter groups of neighboring sugar residues, namely, a shift increment for H-1 of Gal-6 (Δδ 0.076

(80a) The occurrence of the lower-branch, sialo oligosaccharide **35*** in this complex mixture is based upon the presence of an additional H-1 doublet at δ 5.215 that belongs to the α anomer of GlcNAc-2 in a lower-branch oligosaccharide (compare **17** and **18**; see Table VIII) in combination with the H-1 signal at δ 4.925 derived from Man-4′. The *N*-acetyl signal at δ 2.060, which is ascribed to the α anomer of GlcNAc-2 in lower-branch oligosaccharides (compare **18**), and the relatively intense signal at δ ~4.11, partly stemming from the nearly coinciding H-2 resonances of Man-3 (δ ~4.08) and -4′ (δ ~4.10) in such lower-branch compounds (compare **3**, **18**, and **22**; see Tables II, VIII, and IX), support this conclusion. However, the resolution of the 360-MHz, ^{1}H-n.m.r. spectrum available (see Fig. 27) is rather poor; therefore, it is impossible to deduce the detailed n.m.r. data for this lower-branch oligosaccharide with desirable accuracy. The presence of the Gal H-1 doublet(s) at δ ~4.46, next to the signal at δ ~4.44, suggests the occurrence of compound **17** or **18**, or both, in the sample (compare with Table VIII).

FIG. 27.—Resolution-enhanced, Overall, 360-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **35**, **35***, and **21** in the Ratios of 75:7:12. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The signals marked by ϕ (doublet at δ 1.32; quartet at δ 4.11) originate from a frequently occurring nonprotein, noncarbohydrate contaminant of unknown structure. In addition, the sample contained small proportions of the asialo analogs of **35** and **35***, that is, **17** and **18** (see also, footnote 80a).]

TABLE XI

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Oligosaccharides of the *N*-Acetyllactosamine Type Containing Only α-(2→3)-Linked Sialic Acid (Compounds 35–38)

Reporter group	Residue	Anomer of oligosaccharide	Compound and schematic structure			
			35[a] 	36 	37 	38 
H-1 of	2	α	5.206	5.214	5.214	5.213
		β	~4.72	~4.72	~4.72	~4.72
	3	α	~4.79	~4.77	~4.77	4.775
		β	~4.78	~4.76	~4.76	4.763
	4	α	5.122	5.123	5.123	5.122
		β		5.119	5.119	5.120
	4′	α	—	4.929	4.929	4.923
		β		4.932	4.932	4.928
	5	α,β	4.579	4.579	4.587	4.578
	5′	α,β	—	4.587	4.579	4.578

	6	α,β	4.544	4.546	4.469	4.544
	6′	α	—	4.472	4.548	4.546
		β		4.474	4.550	4.547
H-2 of	3	α	4.245	4.260	4.260	4.256
		β	4.234	4.249	4.249	4.244
	4	α,β	4.197	4.195	4.195	4.192
	4′	α,β	—	4.117	4.117	4.117
H-3 of	6	α,β	4.115	4.115	<4.0	4.113
	6′	α,β	—	<4.0	4.117	4.115
H-3*a* of	NeuAc	α,β	1.799	1.796	—	1.797
	NeuAc′	α,β	—	—	1.799	1.799
H-3*e* of	NeuAc	α,β	2.758	2.757	—	2.758
	NeuAc′	α,β	—	—	2.757	2.758
NAc of	2	α	2.043	2.060	2.060	2.061
		β	2.041	2.057	2.057	2.058
	5	α,β	2.049	2.049	2.052	2.049
	5′	α	—	2.049	2.045	2.046
		β		2.047	2.042	2.043
	NeuAc	α,β	2.029	2.031	—	2.032
	NeuAc′	α,β	—	—	2.031	2.032

[a] Chemical shifts measured at 360 MHz, T = 298 K.

α-NeuAc-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′

36

β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc′-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6′ 5′ 4′

37

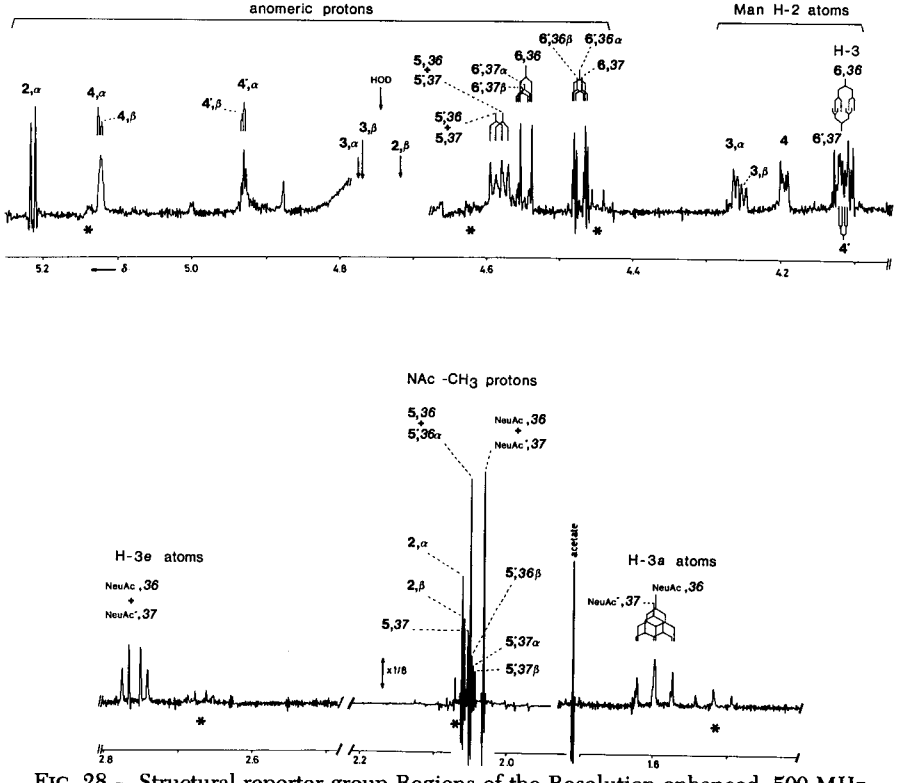


FIG. 28.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **36** and **37** in the Ratio of 3:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the

p.p.m.) and a small shift decrement for the *N*-acetyl protons of GlcNAc-5 ($\Delta\delta$ −0.004 p.p.m.). Furthermore, the H-3 signal of Gal-6 undergoes a relatively large, downfield shift ($\Delta\delta$ ~0.45 p.p.m., compare compound **59**, Table XXIII), thereby emerging from the bulk of skeleton protons, and becoming a structural-reporter group for the αNeuAc(2→3)βGal(1→·) sequence.

In summary, the type of linkage of NeuAc to Gal, namely, α-(2→3) or α-(2→6), can be distinguished on the basis of (*i*) the set of chemical shifts for the H-3 signals of NeuAc itself, and (*ii*) the typical effects of attachment of NeuAc on the chemical shifts of structural-reporter groups of neighboring residues (see Tables XIII and XIV).

It should be noted that the anomerization effect upon the structural-reporter group signals of **35** is only observable for those of GlcNAc-2 and Man-3, probably due to the lower resolving power of the 360-MHz with respect to the 500-MHz spectrometer. Apparently, the effect of anomerization on the various structural-reporter groups is similar to that for **17**, as will later be considered in more detail (for example, **36** and **37**).

Compounds **36** and **37** are diantennary, monosialo oligosaccharides, each containing an α-(2→3)-linked NeuAc group. They were obtained in a mixture (together with a small amount of **27**) from the urine of a patient with sialidosis.[52,72] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture, containing **36** and **37** in the ratio of 3:1, is presented in Fig. 28. The n.m.r. data for both oligosaccharides are given in Table XI.

For the spectral interpretation, **36** and **37** are conceived of as extensions of the diantennary, asialo oligosaccharide **7**. The spectral parameters of the structural-reporter groups of the core residues, that is, GlcNAc-2, and Man-3, -4, and -4′, as well as doubling, due to anomerization, of the structural-reporter-group signals of the constituent monosaccharides (including those of GlcNAc-5′ and Gal-6′) are not influenced by the attachment of a NeuAc group in α-(2→3) linkage to one of the Gal residues. The unaltered differences in the influence of anomerization on the chemical shifts of upper- and lower-branch *N*-acetyllactosamine structural-reporter groups turn out to be valuable for spectral assignments.

For locating an α-(2→3)-linked NeuAc in one of the branches of a diantennary, monosialo structure, the combination of three effects is

---

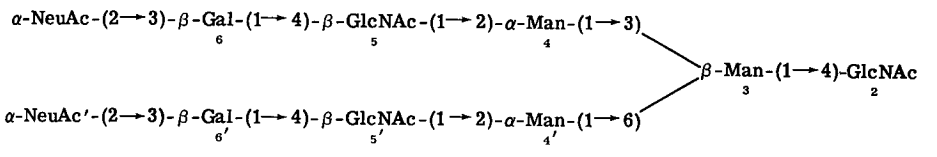
spectrum, as indicated. The HOD resonance, as well as the H-1 signals of Man-3 for both anomers, and that of GlcNAc-2 for the β anomer, of **36** and **37**, have been omitted from the spectrum; their positions are indicated by arrows. The sample was contaminated with a small proportion of **27**, containing α-(2→6)-linked NeuAc, as can be inferred from the signals marked by asterisks.]

specific. (*i*) As already described, the H-1 signal of Gal undergoes a downfield shift ($\Delta\delta$ 0.077 p.p.m.) upon substitution by NeuAc in α-(2→3) linkage. Therefore, the relatively intense, single doublet at δ 4.546 is ascribed to H-1 of Gal-6 in **36**; the set of doublets at δ 4.472 and 4.474, occurring in the anomeric ratio, belongs to H-1 of Gal-6′ in the α and β anomer of **36**, respectively. The latter values of chemical shift are in accord with those for **7** (see Table IV). For the minor component **37**, the H-1 resonance of the terminal Gal-6 is found at δ 4.469 (compare **7**), whereas the H-1 signal of Gal-6′, still doubled due to anomerization, is shifted towards δ 4.548 (α anomer) and 4.550 (β anomer of **37**). This shows that the attachment of NeuAc in α-(2→3) linkage does not affect the influences of anomerization. (*ii*) From the relatively complex pattern at $4.10 < \delta < 4.13$ p.p.m., it may be inferred that, besides the H-2 resonance of Man-4′, two separate Gal H-3 signals occur in this spectral region, at δ 4.115 and 4.117, which are assigned to Gal-6 in **36** and to Gal-6′ in **37**, respectively, on the basis of their relative intensities, namely, 3:1. (*iii*) The *N*-acetyl singlet of GlcNAc-5 in **36** is found at δ 2.049, showing a small upfield shift ($\Delta\delta$ −0.003 p.p.m.) as compared to the asialo upper branch of **7**. The *N*-acetyl signal of GlcNAc-5′ in **36** is doubled due to anomerization; the pair of *N*-acetyl singlets is observed at δ 2.049 (coinciding with the signal of GlcNAc-5, see earlier) and δ 2.047 for the α and β anomer of **36**, respectively (compare **7**). For compound **37**, the *N*-acetyl signal of GlcNAc-5 in the asialo upper branch is found at δ 2.052, whereas both singlets of GlcNAc-5′ are shifted upfield, to δ 2.045 and 2.042, on attachment of NeuAc′ in α-(2→3) linkage to Gal-6′.

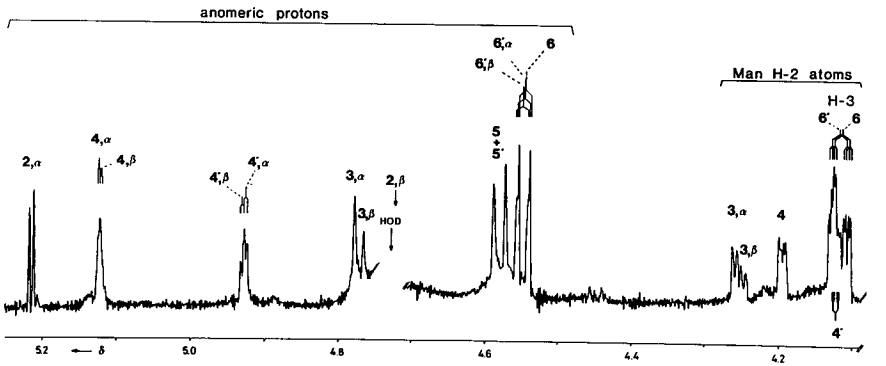
The effects of attachment of NeuAc in α-(2→3) linkage to Gal are restricted to the structural-reporter groups of the *N*-acetyllactosamine unit to which Gal belongs; as the chemical shifts of corresponding protons of these units in the asialo upper and lower branch of a diantennary structure differ only slightly (compare compounds **7**, **8**, and **15**; see Tables IV, V, and VII), it is advisable to utilize the combination of the three effects just mentioned, in order to establish unambiguously the upper- or lower-branch position of NeuAc in α-(2→3) linkage to Gal.

Also, the H-1 signal of GlcNAc in the *N*-acetyllactosamine moiety is influenced by sialylation of Gal at O-3 ($\Delta\delta$ −0.003 p.p.m., and the line width of the resonance is increased), but the signals of the H-1 atoms of GlcNAc-5 in **36** and GlcNAc-5′ in **37** remain indistinguishable from each other (see Table XI).

Essentially, the resonance positions of the structural-reporter-group signals of NeuAc in **36** and **37** are identical to those described for **35**. However, there is a small difference between the chemical shifts of



38



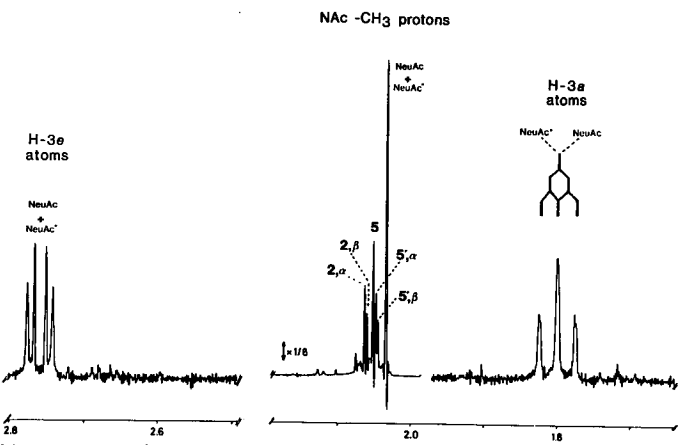


FIG. 29.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **38**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **38**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of GlcNAc-2 for the β anomer of **38**, have been omitted from the spectrum; their positions are indicated by arrows.]

the H-3*a* signals of the upper- (δ 1.796) and lower-branch NeuAc (δ 1.799); the relative intensities of the latter signals corroborate that the molar ratio of the mixture of **36** and **37** is 3 : 1.

Compound **38** is a diantennary, sialo oligosaccharide that has been isolated from the urine of a patient with sialidosis.[72] The 500-MHz, ^{1}H-n.m.r. spectrum of **38** is depicted in Fig. 29, and its relevant n.m.r. parameters are included in Table XI.

The chemical shifts of the structural-reporter groups of the core monosaccharides, that is, GlcNAc-2, and Man-3, -4, and -4′, and of those of the sialylated *N*-acetyllactosamine branches of compounds **36** (5-6) and **37** (5′-6′), are found unaltered in the spectrum of **38**. The attachment of NeuAc in α-(2→3) linkage to Gal-6 or Gal-6′, or both, does not influence the anomerization effect upon any structural-reporter group of the oligosaccharide, neither with respect to its occurrence nor to its magnitude, as compared to the asialo analog **7** (see Table IV).

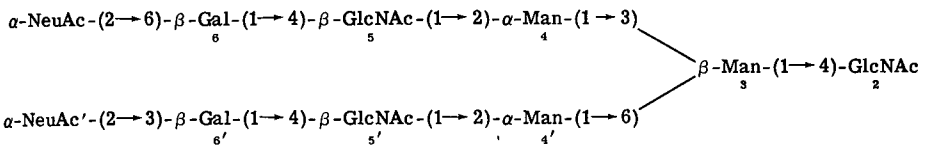
The NeuAc groups in **38** are not completely equivalent, as is evident from the differences in the chemical shifts of their H-3*a* signals (see Table XI).

(iii) Structures Possessing α-(2→6)- and α-(2→3)-Linked NeuAc Groups.—Compound **39** is a diantennary, sialo oligosaccharide having its two branches terminated with NeuAc, but in a mixed type of linkage; it has been isolated from the urine of a patient suffering from sialidosis.[52,72] The 500-MHz, ^{1}H-n.m.r. spectrum of **39** is presented in Fig. 30, and its spectral data are compiled in Table XII.

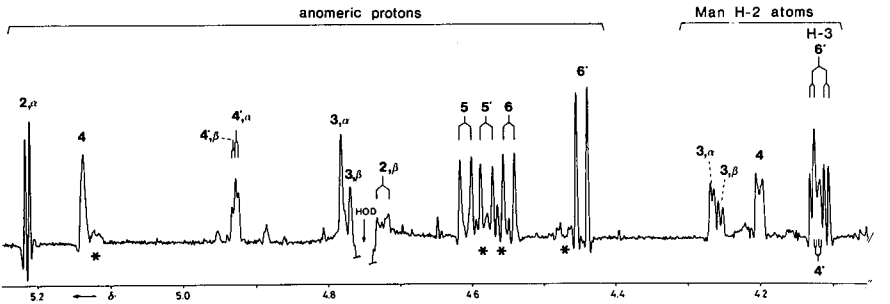
The spectral features of the upper-branch residues Man-4, GlcNAc-5, Gal-6, and NeuAc are essentially identical to those of other oligosaccharides possessing an α-(2→6)-sialylated, upper branch (namely, **21**, **23**, **25**, **27**, and **29**; see Table IX). The n.m.r. parameters of the structural-reporter groups of the lower-branch residues Man-4′, GlcNAc-5′, Gal-6′, and NeuAc′ very closely resemble those of the corresponding residues in **37** and **38**, also having NeuAc′ in α-(2→3) linkage to Gal-6′ (see Table XI).

The spectrum of **39** (see Fig. 30) clearly illustrates that the sets of chemical shifts for the NeuAc H-3 signals are characteristic for the α-(2→3) or α-(2→6) type of linkage to the Gal residue of an *N*-acetyllactosamine moiety. These values of chemical shift are independent of elongation of the other branch in a diantennary structure by NeuAc in any type of linkage.

Compound **40** is a double-branched oligosaccharide containing a disubstituted Man-4, but lacking the lower branch up to and including Man-4′, as compared to di- or tri-antennary structures; it is an example



39



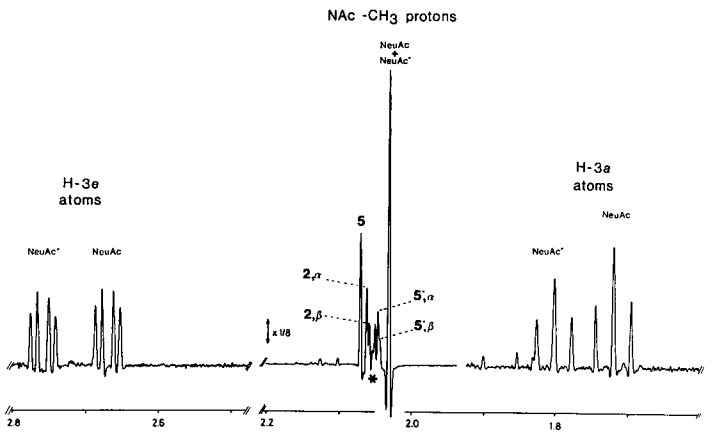
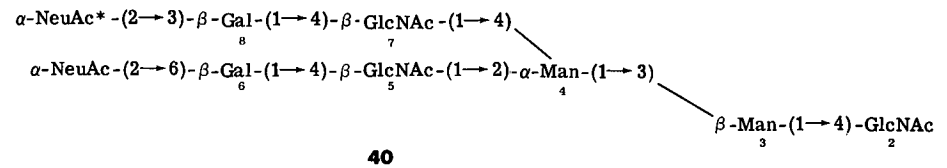


FIG. 30.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **39**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **39**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow. The sample was contaminated with small proportions of the monosialo analogs of **39** lacking NeuAc or NeuAc', or even the asialo analog (oligosaccharide **7**). This can be inferred from the signals marked by asterisks.]

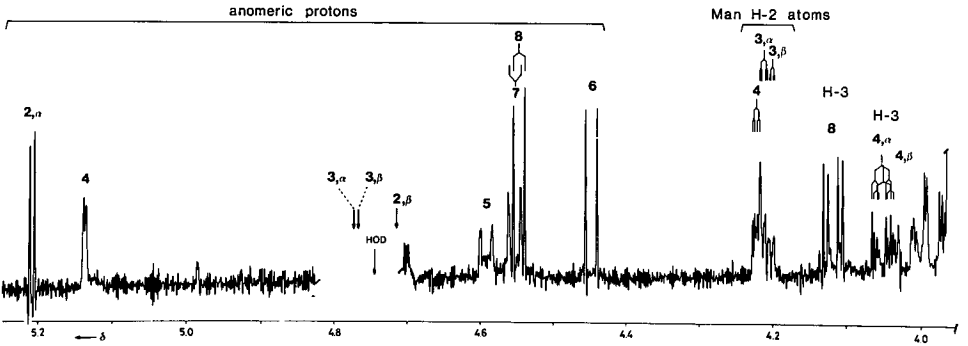
TABLE XII

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Oligosaccharides of the *N*-Acetyllactosamine Type Containing Both α-(2→3)- and α-(2→6)-Linked Sialic Acid (Compounds 39–41)

			Compound and schematic structure		
			39	40	41
Reporter group	**Residue**	**Anomer of oligosaccharide**			
H-1 of	2	α	5.213	5.207	5.215
		β	~4.72	~4.72	~4.72
	3	α	4.780	~4.78	~4.77
		β	4.767	~4.77	~4.76
	4	α,β	5.136	5.135	5.134
	4′	α	4.924	—	4.943
		β	4.929		
	5	α,β	4.606	4.590	4.594
	5′	α,β	4.578	—	4.605
	6	α,β	4.444	4.445	4.443
	6′	α	4.547	—	4.443
		β			4.447
	7	α,β	—	4.552	4.551
	8	α,β	—	4.546	4.545
H-2 of	3	α	4.263	4.211	4.228
		β	4.251	4.201	4.217
	4	α,β	4.197	4.219	4.221
	4′	α,β	4.119	—	4.119
H-3 of	4	α	<4.0	4.053	4.049
		β		4.049	
	4′	α,β	<4.0	—	<4.0
	6	α,β	<4.0	<4.0	<4.0
	6′	α,β	4.115	—	<4.0
	8	α,β	—	4.116	4.116
H-3*a* of	αNeuAc(2→6)	α,β	1.719	1.719	1.720
	αNeuAc′(2→6)	α,β	—	—	1.717
	αNeuAc(2→3)	α,β	1.800	1.801	1.802
H-3*e* of	αNeuAc(2→6)	α,β	2.668	2.669	2.670
	αNeuAc′(2→6)	α,β	—	—	2.672
	αNeuAc(2→3)	α,β	2.756	2.757	2.757
NAc of	2	α	2.061	2.042	2.062
		β	2.057	2.040	2.059
	5	α,β	2.069	2.068	2.067
	5′	α	2.045	—	2.067
		β	2.044		2.065
	7	α,β	—	2.073	2.074
	αNeuAc(2→6)	α,β	2.030	2.030	2.030
	αNeuAc′(2→6)	α,β	—	—	2.030
	αNeuAc(2→3)	α,β	2.030	2.030	2.030



40



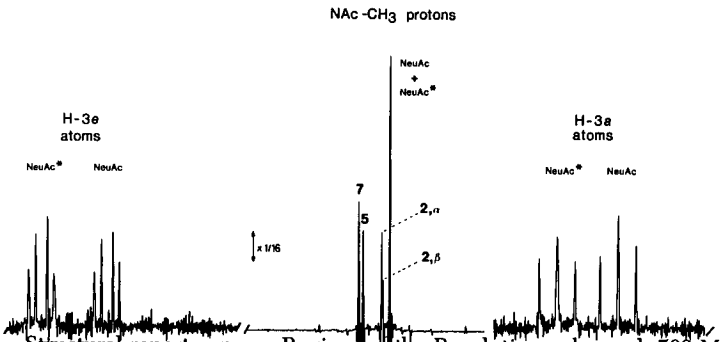


FIG. 31.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **40**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **40**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signals of Man-3 for both anomers and that of GlcNAc-2 for the β anomer of **40**, have been omitted from the spectrum; their positions are indicated by arrows.]

of an oligosaccharide possessing two NeuAc residues in different types of linkage, and it has been obtained from the urine of a patient with sialidosis.[72] The 500-MHz, ^{1}H-n.m.r. spectrum of **40** is given in Fig. 31, and its n.m.r. parameters are summarized in Table XII.

Mono-α-(1→3) substitution of Man-3 is evident from the presence of

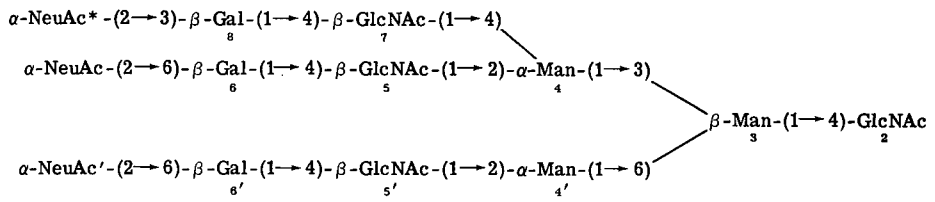
the Man-4 H-1 doublet at δ 5.135, in combination with the absence of a Man-4′ H-1 signal in the spectral region $4.85 < \delta < 4.95$ p.p.m. This type of substitution of Man-3 is corroborated by the chemical shifts of H-1 of GlcNAc-2 in the α anomer of **40**, and of the *N*-acetyl singlets of this monosaccharide at the reducing end (compare **5**, **17**, **21**, and **35**; see Tables III, VIII, IX, and XI). The chemical shifts of the H-2 signals of Man-3 and -4 indicate disubstitution of Man-4 at O-2 and O-4, as in tri- and tetra-antennary structures (see Table VI and compounds **9**, **10**, **12**, **13**, **16**, **30**, and **34**).

The α-(2→6)-linked NeuAc group (δH-3*a* 1.719; δH-3*e* 2.669) is attached to Gal-6, as may be inferred from the chemical shifts of H-1 of Man-4 (δ 5.135), of the *N*-acetyl signal of GlcNAc-5 (δ 2.068), and of the single H-1 signal of Gal-6 (δ 4.445). Furthermore, the anomerization effect upon the chemical shift of H-1 of Man-4 is strongly lessened ($|\Delta\delta_{\alpha-\beta}| < 0.001$ p.p.m.) by the attachment of this NeuAc (compare the step from **10** to **30** or to **41**).

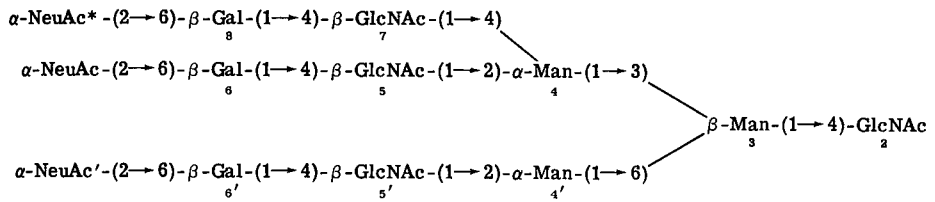
The *N*-acetyllactosamine unit 7-8 is extended with an α-(2→3)-linked NeuAc group (δH-3*a* 1.801; δH-3*e* 2.757). In accordance with the effects of introduction of NeuAc α-(2→3)-linked to an *N*-acetyllactosamine unit (see **35–39**, and Table XIV), the H-1 doublet of Gal-8 is found at δ 4.546; the H-3 signal of the latter residue is observed at δ 4.116, and the *N*-acetyl singlet of GlcNAc-7 is shifted slightly upfield, to δ 2.073, compared to the triantennary, asialo oligosaccharide **10** (see Table IV). Finally, H-1 of GlcNAc-7 shows a small downfield shift (Δδ 0.004 p.p.m.) with respect to the asialo 7–8 branch (for example, in **10**). It should be noted that the assignment of the doublets at δ 4.552 and 4.546 to the H-1 atoms of GlcNAc-7 and Gal-8, respectively, is facilitated by the considerable difference in the line width of the signals.

Compound **41** is a triantennary, sialo oligosaccharide possessing α-(2→6)-linked NeuAc groups in the terminal position of both β-(1→2)-linked *N*-acetyllactosamine branches, and an α-(2→3)-linked NeuAc as the terminating group of the β-(1→4)-linked part. Compound **41** was available as the main component (92%) of a mixture of two oligo-

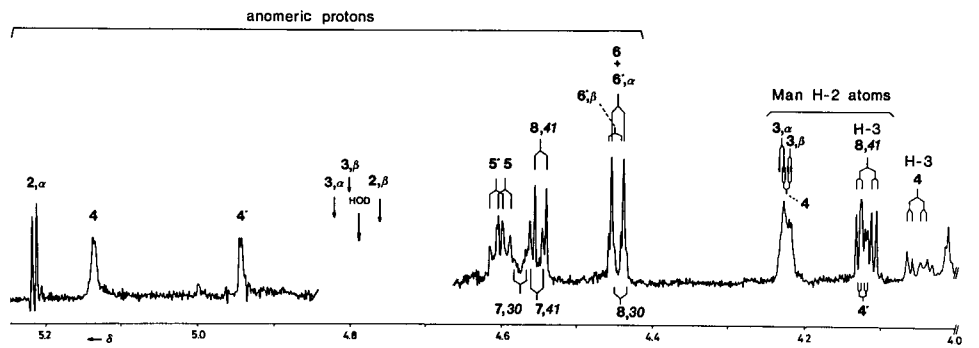
FIG. 32.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **41** and **30** in the Ratio of 23:2. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture, each of which occurs as an anomeric mixture in the ratio of $\alpha:\beta = 2:1$. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signals of Man-3 for both anomers and that of GlcNAc-2 for the β anomer of **41**, and also of **30**, have been omitted from the spectrum; their positions are indicated by arrows.]

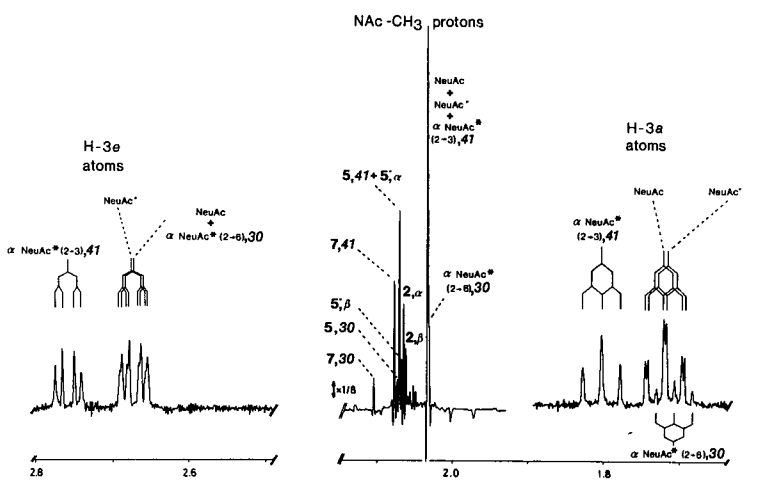

α-NeuAc*-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→4)
8 7
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc'-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'

41


α-NeuAc*-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→4)
8 7
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
6 5 4
β-Man-(1→4)-GlcNAc
3 2
α-NeuAc'-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
6' 5' 4'

30


anomeric protons
Man H-2 atoms
2,α
4
4'
3,β
3,α
2,β
HOD
5'
5
8,41
6
6',α
6',β
3,α
3,β
4
H-3
8,41
H-3
4
7,30
7,41
8,30
4'
5.2
δ
5.0
4.8
4.6
4.4
4.2
4.0


NAc-CH3 protons
H-3e
atoms
NeuAc'
NeuAc
+
α NeuAc* (2-6),30
α NeuAc* (2-3),41
NeuAc
+
NeuAc'
+
α NeuAc*
(2-3),41
5,41+5',α
7,41
2,α
5',β
5,30
2,β
7,30
α NeuAc*
(2-6),30
H-3a
atoms
NeuAc
NeuAc'
α NeuAc*
(2-3),41
α NeuAc* (2-6),30
2.8
2.6
2.0
1.8

saccharides also containing **30**. This mixture was isolated from the urine of a patient suffering from sialidosis.[13,52,72] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is given in Fig. 32, and the n.m.r. parameters for **41** are listed in Table XII.

The triantennary type of branching of the trimannosyl part of the oligosaccharide was established on the basis of the chemical shifts of the H-1 and H-2 signals of Man-3, -4, and -4′, taking into consideration the shift effects upon these due to extension of both the 5-6 and the 5′-6′ branch with NeuAc in α-(2→6) linkage (compare Tables VI and XIII; see also, compound **30**).

TABLE XIII

Recognition of α-(2→6)-Linked, Terminal NeuAc in the Sequence αNeuAc(2→6)βGal(1→4)βGlcNAc(1→*x*)αMan(1→*y*)βMan(1→·)

^{1}H Chemical Shifts of Structural-reporter Groups of α-(2→6)-Linked NeuAc (at pD ~7)

Reporter group	δ ±s.d.
H-3*a* (*x* = 2; *y* = 3)	1.718 ±0.002
H-3*a* (*x* = 2; *y* = 6)	1.717 ±0.001
H-3*a* (*x* = 4; *y* = 3)	1.706 ±0.002
H-3*e* (*x* = 2; *y* = 3)	2.669 ±0.001
H-3*e* (*x* = 2; *y* = 6)	2.672 ±0.001
H-3*e* (*x* = 4; *y* = 3)	2.670 ±0.002
NAc (*x* = 2; *y* = 3 or 6)	2.030 ±0.001
NAc (*x* = 4; *y* = 3)	2.029 ±0.002

Influence of α-(2→6)-Linked NeuAc on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[a]

	Δδ (p.p.m.) ±s.d. (p.p.m.)
H-1 of Gal (*x* = 2 or 4; *y* = 3 or 6)	−0.024 ±0.001
H-1 of GlcNAc (*x* = 2 or 4; *y* = 3 or 6)	+0.024 ±0.001
NAc of GlcNAc (*x* = 2; *y* = 3 or 6)	+0.019 ±0.001
NAc of GlcNAc (*x* = 4; *y* = 3)	+0.025 ±0.002
H-1 of αMan (*x* = 2; *y* = 3)	+0.014 ±0.002
H-1 of αMan (*x* = 2; *y* = 6)	+0.020 ±0.003[b]
H-1 of αMan (*x* = 4; *y* = 3)	~0.0
H-1 of βMan (*x* = 2; *y* = 6)	+0.008 ±0.003
H-2 of αMan (*x* = 2; *y* = 3)	+0.004 ±0.002
H-2 of αMan (*x* = 2; *y* = 6)	+0.005 ±0.002
H-2 of αMan (*x* = 4; *y* = 3)	~0.0
H-2 of βMan (*x* = 2; *y* = 6)	+0.007 ±0.002

[a] Values of Δδ are mean values ± standard deviation (s.d.), at pD ~7 and T = 300 K, calculated from asialo → (2→6)-sialo steps (see Tables IV, V, VIII, IX, X, XII, XV, and XVIII). [b] Except for Man-4′ in triantennary compounds having NeuAc* in α-(2→6) linkage to Gal-8 (see, for example, compound **34** as compared to **9**: Δδ 0.012 p.p.m.).

The spectral features of the α-(2→6)-sialylated *N*-acetyllactosamine branches are identical to those of the corresponding parts of **29** and **30**, whereas the spectral characteristics of the third branch, bearing a NeuAc in α-(2→3) linkage to Gal-8, are identical to those of the same branch in **40** (see Table XII). The H-3*a* signals, and also the H-3*e* signals, of the α-(2→6)-linked NeuAc groups are well separated; furthermore, they are clearly distinguishable from those of NeuAc* α-(2→3)-linked to Gal-8.

The spectral characteristics of carbohydrate chains terminating in *N*-acetyllactosamine residues bearing NeuAc linked to Gal are summarized, for the α-(2→6) and α-(2→3) type of linkage, respectively, in Tables XIII and XIV.

c. Extensions of Carbohydrate Chains of the *N*-Acetyllactosamine Type with Fucose Groups (Compounds 42–54).—Symbols employed for compounds **42–54** are depicted in Chart 3.

In many carbohydrate chains of the *N*-acetyllactosamine type that

TABLE XIV

Recognition of α-(2→3)-Linked, Terminal NeuAc in the Sequence αNeuAc(2→3)βGal(1→4)βGlcNAc(1→*x*)αMan(1→*y*)βMan(1→•)

^{1}H Chemical Shifts of Structural-reporter Groups of α-(2→3)-Linked NeuAc (at pD ~7)

Reporter group	δ ±s.d.
H-3*a* (*x* = 2; *y* = 3)	1.798 ±0.002
H-3*a* (*x* = 2; *y* = 6)	1.800 ±0.001
H-3*a* (*x* = 4; *y* = 3)	1.801 ±0.001
H-3*e* (*x* = 2; *y* = 3)	2.758 ±0.001
H-3*e* (*x* = 2; *y* = 6)	2.757 ±0.001
H-3*e* (*x* = 4; *y* = 3)	2.757 ±0.001
NAc (*x* = 2 or 4; *y* = 3 or 6)	2.030 ±0.001

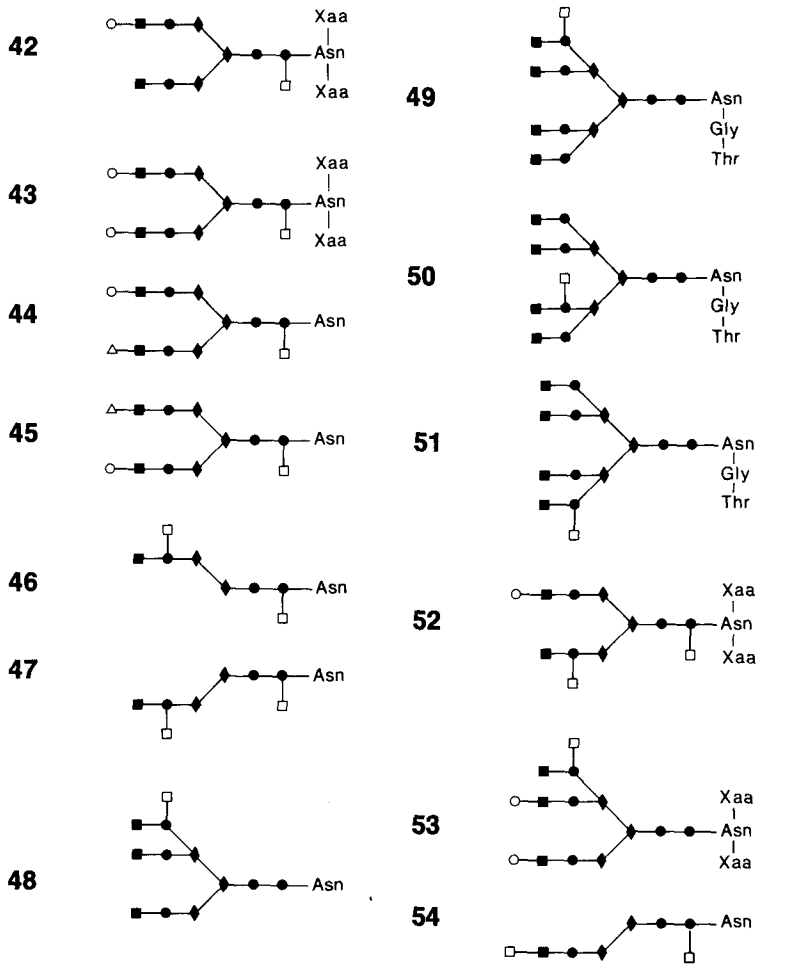
Influence of α-(2→3)-Linked NeuAc on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[a]

	Δδ (p.p.m.) ±s.d. (p.p.m.)
H-1 of Gal (*x* = 2; *y* = 3 or 6)	+0.076 ±0.001
H-1 of Gal (*x* = 4; *y* = 3)	+0.083 ±0.002
H-3 of Gal (*x* = 2 or 4; *y* = 3 or 6)	+0.453 ±0.003
H-1 of GlcNAc (*x* = 2; *y* = 3 or 6)	−0.004 ±0.002
H-1 of GlcNAc (*x* = 4; *y* = 3)	+0.005 ±0.002
NAc of GlcNAc (*x* = 2 or 4; *y* = 3 or 6)	−0.003 ±0.001

[a] Values of Δδ are mean values ±s.d., at pD ~7 and T = 300 K, calculated from asialo → (2→3)-sialo steps (see Tables IV, V, VIII, XI, XII, and XV).

CHART 3

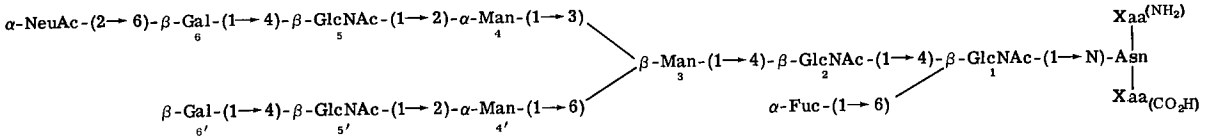
Symbols Employed[a] for Compounds 42–54



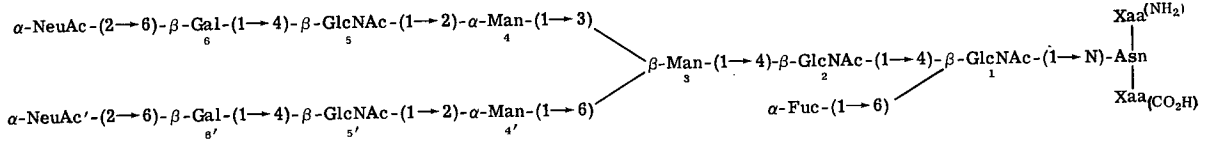
[a] For the key to the symbolic notation, see Chart 1.

are potentially *N*-glycosylically linked to a peptide backbone, one or more Fuc groups occur as terminal monosaccharide(s).[1–4]

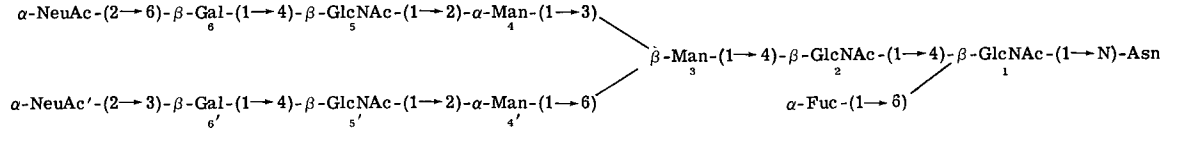
(*i*) When a Fuc group is linked to the trimannosyl-*N*,*N*′-diacetylchitobiose core, reliable (n.m.r.) evidence exists[13,55] only for an α-(1→6) linkage to GlcNAc-1. Compounds **2** and **4**, containing this structural unit, have already been discussed; several others will be treated in this Section.


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)
Xaa(NH2)
Xaa(CO2H)

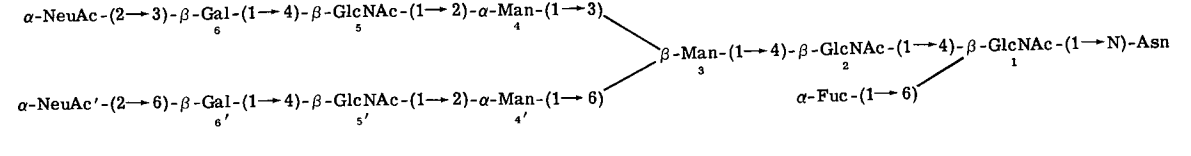
42


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-NeuAc′-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)
Xaa(NH2)
Xaa(CO2H)

43


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-NeuAc′-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)

44


α-NeuAc-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-NeuAc′-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)

45

(*ii*) Another, well defined type of Fuc linkage is the α-(1→3) attachment to a peripheral, GlcNAc residue forming part of an asialo *N*-acetyllactosamine branch. Compounds **46** to, and including, **53** contain the structural unit βGal(1→4)[αFuc(1→3)]βGlcNAc(1→•).

(*iii*) Finally, Fuc may be present as the terminal group in *N*-acetyllactosamine-type structures, in α-(1→2) linkage to a Gal residue. An example of a compound containing the αFuc(1→2)βGal(1→4)-βGlcNAc(1→•) moiety is the glyco-asparagine **54**.

(i) **Fuc α-(1→6)-Linked to GlcNAc-1.**—Compound **42** is a diantennary, monosialo glycopeptide bearing an additional Fuc in α-(1→6) linkage to GlcNAc-1. Compound **42** was obtained from human lactotransferrin[81] in a mixture of glycopeptides consisting of **42** and **52** in the ratio of 3:1. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is given in Fig. 37, and the n.m.r. data for **42** are compiled in Table XV.

For the spectral interpretation, the carbohydrate chain of **42** is regarded as an extension of that of the diantennary, monosialo glycopeptide **31**, with a Fuc residue α-(1→6)-linked to GlcNAc-1. The chemical shifts of the Fuc structural-reporter groups, namely, H-1 (δ 4.877), H-5 (δ 4.125), and CH_3 (δ ~ 1.208), are in perfect agreement with those reported for **4** (see Table II). For the protons of the CH_3 group, two doublets are observed, at δ 1.206 and 1.210, probably due to the heterogeneity of the peptide moiety. The effects on the chemical shifts of structural-reporter groups of neighboring residues, due to the attachment of Fuc in α-(1→6) linkage to GlcNAc-1, are restricted to H-1 (Δδ 0.062 p.p.m.) and the *N*-acetyl protons of GlcNAc-2 (Δδ 0.016 p.p.m., in comparison to **31**). All other n.m.r. parameters of **31** are found, essentially unaltered, in the spectrum of **42**. The value of the chemical shift of H-1 of Man-4 shows that **42** contains NeuAc in α-(2→6) linkage to Gal-6 (see Tables VI and XIII).

Compound **43** is a diantennary, sialo glycopeptide containing a terminal α-(2→6)-linked NeuAc group in each branch, and a Fuc group α-(1→6)-linked to GlcNAc-1. Compound **43** was derived from human lactotransferrin.[81] It could also be identified as a minor constituent of a complex mixture of glycopeptides that additionally contained compounds **34** and **53** and a glycopeptide analog of **41**. This mixture was derived from human-plasma ceruloplasmin; its 500-MHz, ^{1}H-n.m.r. spectrum is given in Fig. 38. The relevant n.m.r. data for **43** are included in Table XV.

As compared to the spectrum of the afuco analog **33**, the spectrum of **43** reveals three additional, structural-reporter-group signals, namely,

(81) G. Spik, G. Strecker, B. Fournet, S. Bouquelet, J. Montreuil, L. Dorland, H. van Halbeek, and J. F. G. Vliegenthart, *Eur. J. Biochem.*, 121 (1982) 413–419.

the Fuc H-1, H-5, and CH_3 resonances, in conjunction with downfield-shift effects upon the H-1 and *N*-acetyl proton signals of GlcNAc-2. The step from **33** to **43** is completely analogous to that from **31** to **42**. The set of chemical shifts of these five structural-reporter groups is typical for the presence of a Fuc group α-(1→6)-linked to GlcNAc-1 of the *N*, *N'*-diacetylchitobiose core-region of carbohydrate chains that are *N*-glycosylically linked.

Compounds **44** and **45** are diantennary glyco-asparagines termi-

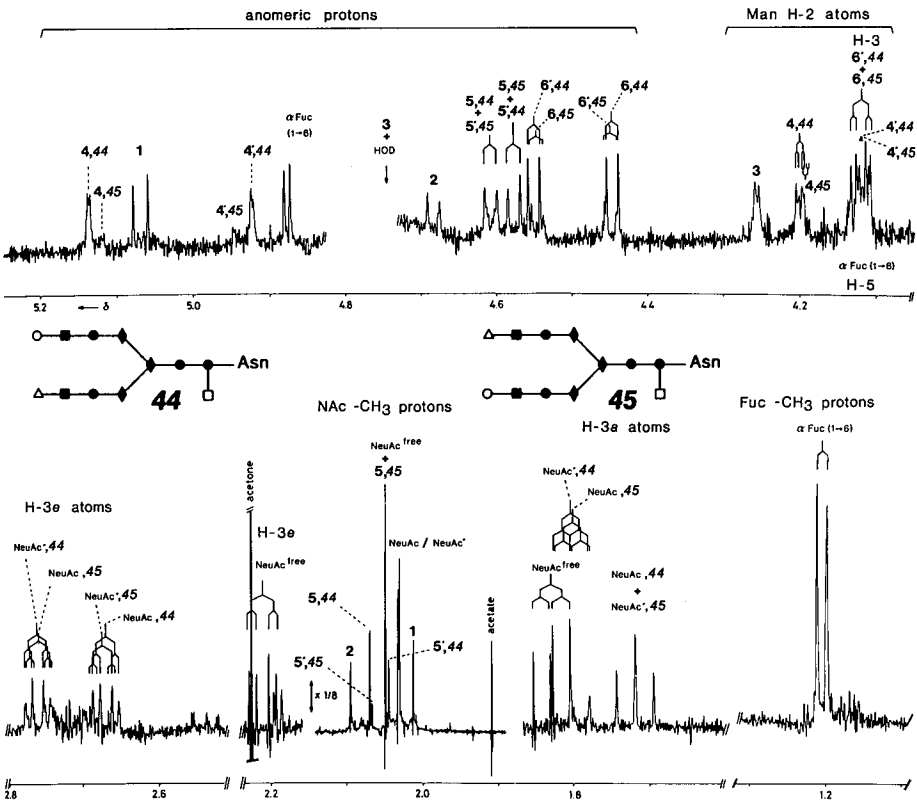
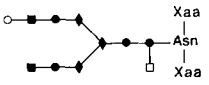
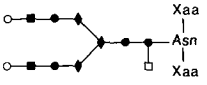
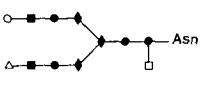
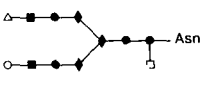


FIG. 33.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **44** and **45** in the Ratio of 5:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures (see p. 307), and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the two components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from the spectrum; their position is indicated by an arrow. The sample was contaminated with free NeuAc, of which only the β anomer is detectable in this spectrum (δH-3*a* 1.829, δH-3*e* 2.210, and δNAc 2.050).]

TABLE XV

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glycopeptides of the *N*-Acetyllactosamine Type, Containing Fuc in α-(1→6) Linkage to GlcNAc-1 (Compounds 42–45)

		Compound and schematic structure			
		42 	**43** 	**44** 	**45** 
Reporter group	**Residue**				
H-1 of	1	5.102[a]	5.045	5.068	5.068
	2	4.682[a]	4.684	4.682	4.682
	3	4.771	~4.77	~4.77	~4.77
	4	5.134	5.131	5.136	5.118
	4′	4.927	4.946	4.922	4.940
	5	4.605	4.607	4.606	4.575
	5′	4.579	4.607	4.575	4.606
	6	4.445	4.444	4.444	4.545
	6′	4.472	4.446	4.549	4.447
H-2 of	3	4.256	4.257	4.253	4.253
	4	4.191	4.200	4.197	4.190
	4′	4.112	4.112	~4.11[b]	~4.11[b]

H-3 of	6	<4.0	<4.0	<4.0	~4.11[b]
	6′	<4.0	<4.0	~4.11[b]	<4.0
H-3*a* of	αNeuAc(2→6)	1.720	1.717[c]	1.717	1.717
	αNeuAc(2→3)	—	—	1.802	1.800
H-3*e* of	αNeuAc(2→6)	2.667	2.670	2.669	—
	αNeuAc′(2→6)	—	2.673	—	2.671
	αNeuAc(2→3)	—	—	2.760	2.758
H-1 of	αFuc(1→6)	4.877	4.873	4.876	4.876
H-5 of	αFuc(1→6)	4.125	~4.12	~4.12[b]	~4.12[b]
CH_3 of	αFuc(1→6)	1.206[a]	1.200	1.202	1.202
		1.210[a]			
NAc of	1	2.006[a]	2.020	2.012	2.012
		2.016[a]			
	2	2.094	2.094	2.094	2.094
	5	2.069	2.065	2.069	2.048
	5′	2.048	2.065	2.044	2.066
	αNeuAc(2→6)	2.030	2.030[d]	2.030[e]	2.030[e]
	αNeuAc(2→3)	—	—	2.032[e]	2.032[e]

[a] Signals stemming from the main component(s) with respect to the heterogeneity of the peptide moiety. [b] Values could not be determined more accurately (±0.01 ppm), due to partial overlapping of the H-5 signal of Fuc, the H-2 signals of Man-4′, and the H-3 signals of Gal-6 in the subspectrum of **45** and -6′ in that of **44**. [c] Signal of two protons. [d] Signal of two methyl groups. [e] Assignments may have to be interchanged.

nated by differently linked NeuAc groups and bearing Fuc in α-(1→6) linkage to GlcNAc-1. These compounds were obtained in a mixture, containing **44** and **45** in the ratio of 5:1, from horse-pancreatic ribonuclease.[66,74,82] The 500-MHz, ^{1}H-n.m.r. spectrum of the mixture is presented in Fig. 33; the n.m.r. parameters for **44** and **45** are summarized in Table XV.

For the spectral interpretation, the glyco-asparagines are conceived of as extensions of the corresponding oligosaccharides ending in GlcNAc-2 (for example, **39**), with an αFuc(1→6)βGlcNAc(1→N)Asn moiety. The presence of GlcNAc-1 *N*-glycosylically linked to Asn is evident from the H-1 doublet at δ 5.068, and the singlet of its *N*-acetyl protons at δ 2.012. The H-1 and *N*-acetyl signals of GlcNAc-2 are found at the same positions as described for compound **4** (see Table II). In combination with the chemical shifts of the Fuc structural-reporter groups (δH-1 4.876, δH-5 ~4.12, and δCH$_3$ 1.202), the chemical shifts of the GlcNAc-2 reporter groups are characteristic for the α-(1→6) type of linkage of Fuc to GlcNAc-1.

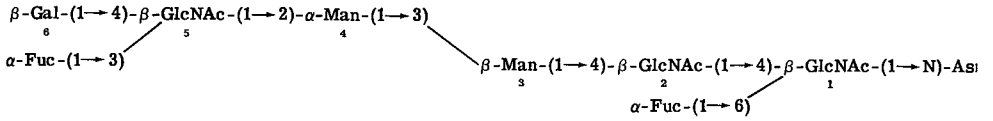
The total pattern of the H-1 and H-2 signals of Man-3, -4, and -4′ is indicative of the diantennary type of structure (compare Table VI and Fig. 14a).

The spectral parameters of the upper- and lower-branch residues of **44**, from Man-4/-4′ up to, and including, the NeuAc residues, are identical with those for the β anomer of oligosaccharide **39** (see Table XII).

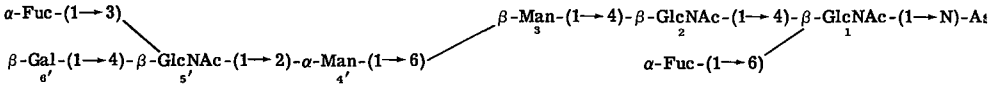
The presence in the mixture of the isomeric, minor compound **45**, containing NeuAc in α-(2→3) linkage in the upper branch, and NeuAc′ in α-(2→6) linkage in the lower branch, is proved by the occurrence of relatively low Gal H-1 doublets at δ 4.545 (Gal-6) and δ 4.447 (Gal-6′), and of *N*-acetyl signals of GlcNAc-5 and -5′ at δ 2.048

(82) B. L. Schut, L. Dorland, J. Haverkamp, J. F. G. Vliegenthart, and B. Fournet, *Biochem. Biophys. Res. Commun.*, 82 (1978) 1223–1228.

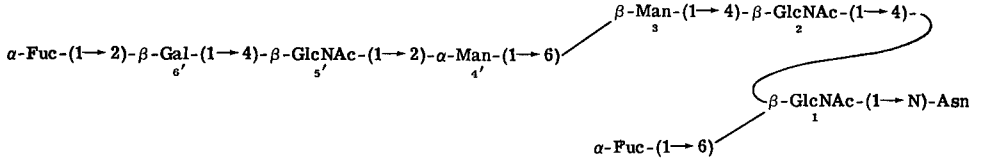
FIG. 34.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **46**, **47**, and **54** in the Ratios of 1:6:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow. In addition, the sample contained one (or more) structurally related, difuco glyco-asparagine(s). The presence of a mono-antennary, lower-branch analog of **47** that contains a repeating *N*-acetyllactosamine moiety, namely, βGal(1→4)[αFuc(1→3)]βGlcNAc(1→3)βGal(1→4)GlcNAc, β-(1→2)-linked to Man-4′, may be suggested, mainly on the basis of the signals marked by asterisks (see Ref. 13).]


β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-Fuc-(1→3)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)

46


α-Fuc-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)

47


α-Fuc-(1→2)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-
β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)

54

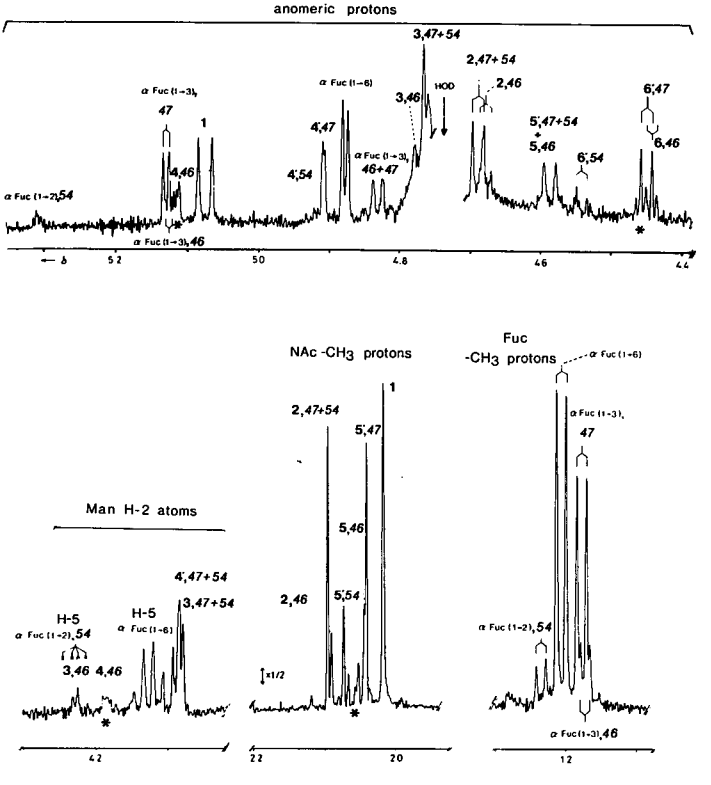
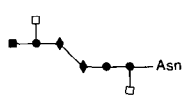
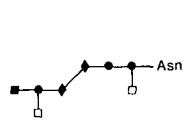
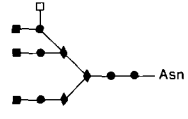
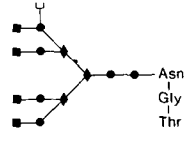
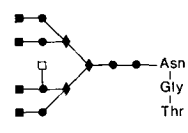
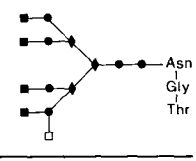

anomeric protons
NAc -CH3 protons
Fuc
-CH3 protons
Man H-2 atoms
H-5
H-5
δ

TABLE XVI

[1]H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glycopeptides of the *N*-Acetyllactosamine Type Containing α-(1→3)-Linked Fucose (Compounds 46–51)

Reporter group	Residue	Compound and schematic structure					
		46 	47 	48 	49 	50 	51 
H-1 of	1	5.073	5.073	5.071	5.067	5.067	5.067
	2	4.677	4.687	4.616	4.611	4.611	4.611
	3	4.775	4.764	4.756	4.758	4.758	4.758
	4	5.110	—	5.116	5.124	5.124	5.124
	4′	—	4.906	4.925	4.866	4.866	4.866
	5	4.585	—	4.568	4.570	4.570	4.570
	5′	—	4.585	4.580	4.594	~4.61[a]	4.594
	6	4.442	—	4.464	4.462	4.468	4.468
	6′	—	4.449	4.474	4.470	4.462	4.470

	7	—	—	4.557	4.558	4.545	4.545
	7′	—	—	—	4.550	4.550	4.563[a]
	8	—	—	4.447	4.447	4.462	4.462
	8′	—	—	—	4.479	4.479	4.473
H-2 of	3	4.228	4.081	4.212	4.209	4.209	4.209
	4	4.187	—	4.219	4.222	4.222	4.222
	4′	—	4.087	4.112	4.092	4.092	4.092
H-3 of	4	<4.0	—	4.038	4.052	4.052	4.052
	4′	—	<4.0	<4.0	<4.0	<4.0	<4.0
H-1 of	αFuc(1→3)	5.126	5.127	5.112	5.112	5.124	5.112
H-5 of	αFuc(1→3)	4.830	4.830	4.835	4.832	4.832	4.832
CH_3 of	αFuc(1→3)	1.172	1.177	1.176	1.176	1.178	1.177
H-1 of	αFuc(1→6)	4.876	4.876	—	—	—	—
H-5 of	αFuc(1→6)	4.126	4.126	—	—	—	—
CH_3 of	αFuc(1→6)	1.206	1.206	—	—	—	—
NAc of	1	2.016	2.016	2.013	2.006	2.006	2.006
	2	2.090	2.095	2.078	2.076	2.076	2.076
	5	2.043	—	2.049	2.052	2.052	2.052
	5′	—	2.040	2.045	2.040	2.033	2.040
	7	—	—	2.067	2.067	2.077	2.077
	7′	—	—	—	2.038	2.038	2.029

[a] Tentative assignments, because of the low amount of compounds **50** and **51** in the mixture with **49**, in combination with the relative broadness of the signals.

and 2.066, respectively. In comparison to the diantennary, asialo glycopeptide **8**, these signals show shift influences that are characteristic for the α-(2→3) type of NeuAc linkage to Gal in the upper branch, and the α-(2→6) linkage type in the lower branch (see Tables XIII and XIV). In accordance with this interpretation, the H-1 signals of Man-4 and -4′ for **45** are observed at δ 5.118 and 4.940, respectively. Their relative intensities correspond with those of the Gal H-1 doublets of **45**. It should be mentioned that the chemical shifts of the H-1 signals of Gal-6 and -6′ in the spectra of **44** and **45** lend independent support for the assignment of the Gal H-1 doublets in the asialo afuco glycopeptide analog **8** (compare **31**; see Ref. 66).

The coincidence of the H-3*a* signals of the α-(2→6)-linked NeuAc groups for both compounds is in line with the observation for **33**. The assignments of the well separated H-3*a* signals of the α-(2→3)-linked NeuAc groups, and also those of the H-3*e* signals of the α-(2→6)-linked NeuAc groups, are based on their relative intensities; they are in accord with those for **38** and **33**, respectively. Finally, the H-3*e* signals of the α-(2→3)-linked NeuAc groups are observed separately, in contrast to the situation for **38**. The assignment given in Table XV is based on the relative intensities of the signals.

(ii) Fuc α-(1→3)-Linked to a Peripheral, GlcNAc Residue.—Compound **46** is a mono-antennary glyco-asparagine bearing two differently linked Fuc groups. Compound **47** is an isomer of **46**, deviating in the type of α-glycosidic linkage between the Man residues. Both compounds occur, together with **54**, in a mixture (ratios of **46**:**47**:**54** = 1:6:1) isolated from the urine of a patient suffering from fucosidosis.[13,52,56] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is given in Fig. 34. The n.m.r. data for **46** and **47** are presented in Table XVI.

The occurrence of the βMan(1→4)βGlcNAc(1→4)[αFuc(1→6)]-βGlcNAc(1→N)Asn moiety in compounds **46** and **47** may be inferred from comparison of the n.m.r. data for **46** and **47** with those for **4** and **42**–**45** (see Tables II and XV). The presence of Fuc α-(1→6)-linked to GlcNAc-1 is obvious from the chemical-shift values of its H-1, H-5, and CH_3 signals, and from its $J_{1,2}$ value (3.6 Hz), as well as from the effects on the chemical shifts of H-1 and the *N*-acetyl protons of GlcNAc-2, as compared to an afuco structure (compare the steps from **3** to **4**, from **31** to **42**, and from **33** to **43**). The chemical shifts of the structural-reporter groups of Fuc in α-(1→6) linkage to GlcNAc-1, together with the influences of its introduction upon the chemical-shift values of reporter groups of neighboring residues, are summarized in Table XVII.

The resonance position of H-2 of Man-3 (δ 4.228) indicates that this residue in **46** is substituted only at O-3 by another Man residue (com-

TABLE XVII

Recognition of α-(1→6)-Linked, Terminal Fuc in the Sequence βGlcNAc(1→4)[αFuc(1→6)]βGlcNAc(1→N)Asn

^{1}H Chemical Shifts of Structural-reporter Groups of α-(1→6)-Linked Fuc[a]

Reporter group	δ ±s.d.
H-1	4.876 ±0.001
H-5	4.125 ±0.003
CH_3	1.207 ±0.003[b]

Influences of α-(1→6)-Linked Fuc on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[c]

	Δδ (p.p.m.) ± s.d. (p.p.m.)
H-1 of GlcNAc-1	+0.01[d]
NAc of GlcNAc-1	+0.008[d]
H-1 of GlcNAc-2	+0.068 ±0.004
NAc of GlcNAc-2	+0.016 ±0.002

[a] Mean values ±s.d. at T = 300 K; see Tables I, II, XV, XVI, and XVIII (the mean value for δH-1 was derived from Tables II, XV, XVI, and XVIII). [b] The value of δCH_3 depends slightly on the composition of the peptide moiety (see, for example, compounds **42** and **52**). [c] Values of Δδ are mean values ±s.d., calculated for afuco → (1→6)-fuco steps (see Tables I, II, V, X, XV, XVI, and XVIII); comparison may be made with the values of Δδ for H-1 and NAc of H GlcNAc-1, and for H-1 of Gal-a, for a similar step in Table XXIII. [d] These effects are barely traceable, as they are negligible compared to the influences of changes in the pD of the solution, or in the composition of the peptide part.

pare **17** and **21**; see Tables VIII and IX). Small differences between the chemical shifts of the structural-reporter groups of GlcNAc-2, and H-1 of Man-3, of **46** as compared to **4** (and to **47**!), also reflect mono-α-(1→3) rather than -α-(1→6) substitution of Man-3 in **46** (see later).

Concerning the peripheral part of the molecule, **46** may be regarded as an extension of the (β anomer of the) mono-antennary, upper-branch, asialo oligosaccharide **17**, with Fuc α-(1→3)-linked to GlcNAc-5. This additional Fuc manifests itself by its structural-reporter-group signals at positions clearly distinguishable from those of the α-(1→6)-linked Fuc already mentioned: δH-1 5.126, δH-5 4.830, and δCH_3 1.172. The $J_{1,2}$ value of the (1→3)-linked Fuc (4.1 Hz) indicates an α linkage. The rather low-field, resonance position of H-5 of Fuc is indicative of a location of this residue in vicinal position to Gal, both being substituents[67,83] of GlcNAc.

(83) R. U. Lemieux, K. Bock, L. T. J. Delbaere, S. Koto, and V. S. Rao, *Can. J. Chem.*, 58 (1980) 631–653; R. U. Lemieux, D. R. Bundle, and D. A. Baker, *J. Am. Chem. Soc.*, 97 (1975) 4076–4083.

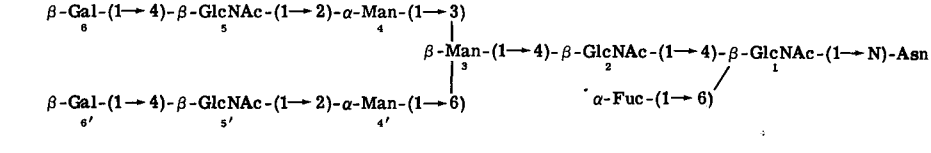
Furthermore, the introduction of this type of Fuc causes some shift effects on structural-reporter groups of neighboring residues. Most specific is the upfield shift of the *N*-acetyl protons of GlcNAc-5 ($\Delta\delta$ −0.01 p.p.m.). Also, the H-1 doublet of Gal-6 is shifted upfield ($\Delta\delta$ −0.026 p.p.m.) towards δ 4.442. Moreover, the chemical shifts of H-1 of Man-4 ($\Delta\delta$ −0.017 p.p.m.) and H-1 of GlcNAc-5 ($\Delta\delta$ +0.008 p.p.m.), both compared to **17**β, are significantly influenced. The introduction of Fuc α-(1→3)-linked to GlcNAc-5 does not cause strong changes in line widths of structural-reporter-group signals of neighboring residues.

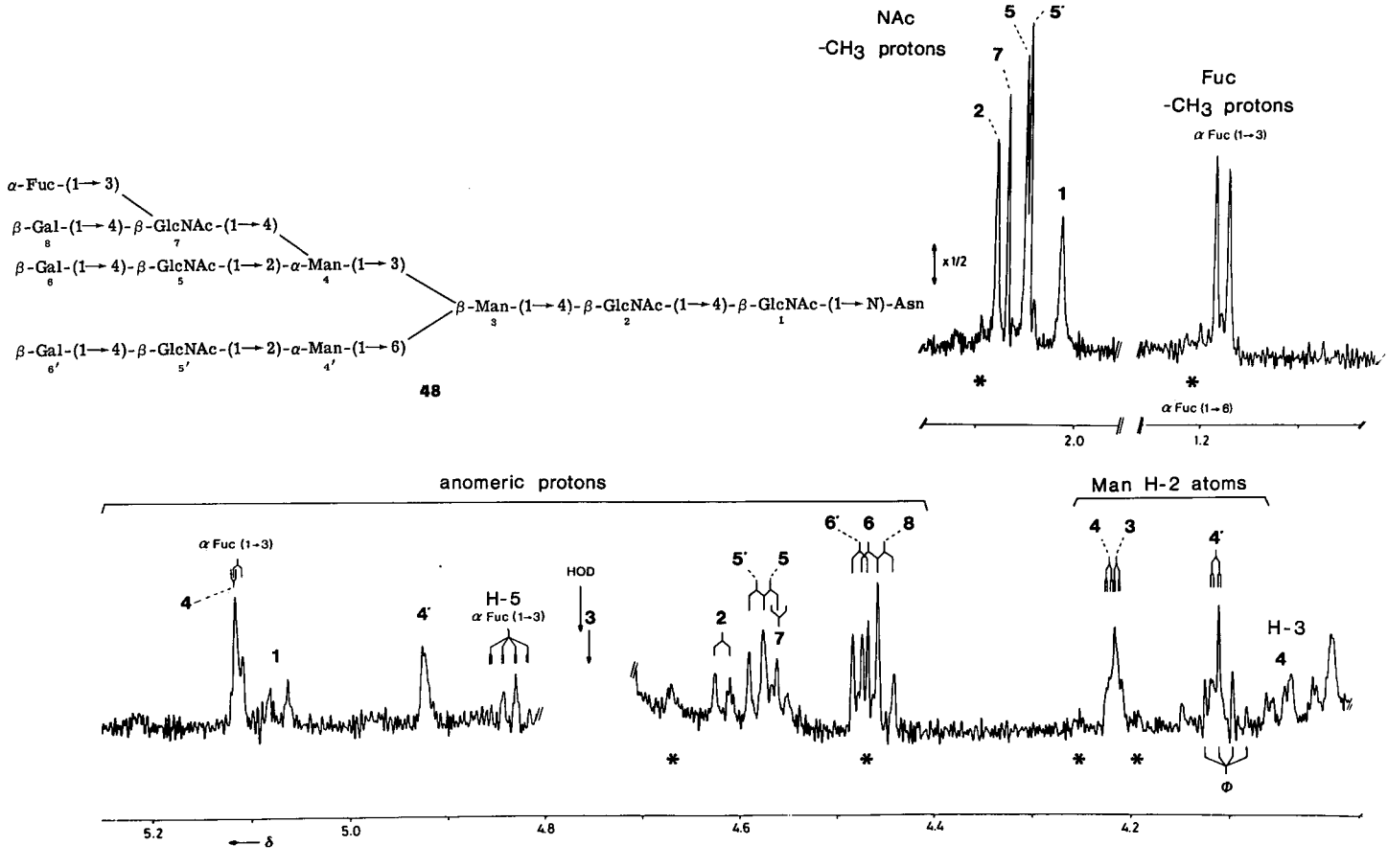
In **47**, the main component of the mixture, Man-3 is mono-α-(1→6)-substituted by Man-4′. This is obvious from the chemical-shift values of H-1 and H-2 of Man-3 (compare compounds **3**, **4**, **18**, **22**, **61**, and **62**; see Tables II, VIII, IX, and XXIV). The set of chemical shifts for H-1 and the *N*-acetyl protons of GlcNAc-2 [despite the alterations caused by Fuc in α-(1→6) linkage to GlcNAc-1] is also indicative of the α-(1→6) type of mono-substitution of Man-3 (compare **46**), as has already been pointed out for mono-antennary oligosaccharides (for example, **17**, **18**, **21**, **22**, and **35**).

The presence of Fuc in α-(1→3) linkage to GlcNAc-5′ is revealed in the resonance positions of H-1, H-5, and its CH_3 protons (δ 5.127, 4.830, and 1.177, respectively), as well as in its $J_{1,2}$ value (4.1 Hz). It should be noted that the chemical shifts of the CH_3-group protons of **46** and **47** differ clearly.

On comparing the n.m.r. data for **47** with those for the β anomer of **18** (see Table VIII), some characteristic shift increments and decrements of reporter groups may be ascribed to the introduction of Fuc α-(1→3)-linked to GlcNAc-5′. Although the chemical shift of H-1 of 5′

FIG. 35.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **48**. [The bold numbers in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from the spectrum; their positions are indicated by arrows. The signal marked by φ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure. The sample was contaminated with a small proportion of a diantennary, asialo glycopeptide, containing Fuc in α-(1→6) linkage to GlcNAc-1, that is depicted as follows. This can be inferred from the signals marked by asterisks (compare Ref. 84).]




α-Fuc-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→4)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
48
NAc
-CH3 protons
Fuc
-CH3 protons
α Fuc (1→3)
α Fuc (1→6)
x1/2
anomeric protons
Man H-2 atoms
H-5
H-3
HOD
δ

remains essentially unaltered, the *N*-acetyl signal of this residue shows a highly significant, shift decrement upon attachment of Fuc: $\Delta\delta$ −0.007 p.p.m. Furthermore, H-1 of Gal-6′ undergoes a shift decrement from δ 4.472 to 4.449. Finally, the chemical shifts of H-1, as well as of H-2 of Man-4′, in **47** are different from those in **18**β, which may be due to the presence of Fuc.

Most of the shift increments and decrements mentioned are similar for the α-(1→3)-linked Fuc in **46** and **47**. Therefore, they are characteristic for attachment of Fuc in α-(1→3) linkage to GlcNAc of a βGal(1→4)βGlcNAc(1→2)αMan(1→•) moiety, and the chemical shift of the CH_3 group of Fuc might be a helpful parameter for its localization in a certain branch.

Compound **48** is a triantennary glycopeptide containing a terminal Fuc group α-(1→3)-linked to GlcNAc-7. This compound was derived from asialo α_1-acid glycoprotein[67–69] (peptide moiety: Asn-Lys) and from asialo ceruloplasmin[84] (peptide moiety: Asn). The 500-MHz, ^{1}H-n.m.r. spectrum of the glycopeptide obtained from the latter source is shown in Fig. 35; its spectral parameters are listed in Table XVI.

The type of branching of the trimannosyl-*N*,*N*′-diacetylchitobiose core is evident from the set of chemical shifts of the H-1 and H-2 signals of Man-3, -4, and -4′ (compare Table VI and Fig. 14a). For the spectral interpretation, **48** is considered to be an extension of its afuco analog, **9**. The presence of a Fuc group α-(1→3)-linked to GlcNAc in the structural unit βGal(1→4)[αFuc(1→3)]βGlcNAc(1→•) may be inferred from the chemical shifts and the $J_{1,2}$ value (3.6 Hz) of the Fuc structural-reporter-group resonances, which are essentially the same as those for **46** and **47** (see Table XVI). However, the chemical shift of H-1 of Fuc α-(1→3)-linked to GlcNAc-7 is significantly different from that for Fuc α-(1→3)-linked to GlcNAc-5 (**46**) or -5′ (**47**).

Among the shift effects brought about by the introduction of Fuc into **9**, the shift decrement of the *N*-acetyl protons of GlcNAc-7 ($\Delta\delta$ −0.008 p.p.m.) is the most typical; the resonance positions of all other *N*-acetyl signals remain unchanged in comparison to those of **9**. Furthermore, H-1 of Gal-8 undergoes a significant, upfield shift, from δ 4.462 for **9** to δ 4.447 for **48**. The H-1 doublet of GlcNAc-7 itself is shifted downfield ($\Delta\delta$ 0.012 p.p.m.), and is somewhat broadened. Also, the chemical shift of H-1 of Gal-6 is affected by the attachment of Fuc in α-(1→3) linkage to GlcNAc-7 ($\Delta\delta$ −0.004 p.p.m.). This secondary effect probably originates from close spatial proximity of Fuc linked to GlcNAc-7 and the anomeric proton of Gal-6. For the characterization

(84) M. Endo, K. Suzuki, K. Schmid, B. Fournet, Y. Karamanos, J. Montreuil, L. Dorland, H. van Halbeek, and J. F. G. Vliegenthart, *J. Biol. Chem.*, 257 (1982) 8755–8760.

of the location of Fuc in this triantennary structure, the combination of these effects, together with the chemical shift of H-1 of Fuc itself, is decisive.

Compounds **49, 50,** and **51** are isomeric, tetra-antennary, monofuco glycopeptides differing in the type of glycosidic linkage of the peripheral GlcNAc to which Fuc is attached. These glycopeptides were prepared in a mixture, containing **49, 50,** and **51** in the ratios of 12:2:3, from asialo α_1-acid glycoprotein.[13,67,85] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 36; the pertinent n.m.r. parameters for the three constituents are summarized in Table XVI.

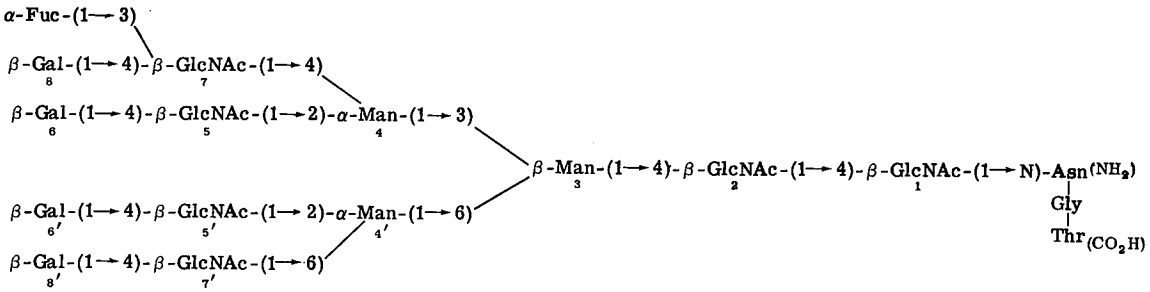
The tetra-antennary type of branching of the trimannosyl part of the pentasaccharide core in the three glycan chains may be deduced from the set of chemical-shift values for the H-1 and H-2 signals of Man-3, -4, and -4′ (see Fig. 14a, Table VI, and the afuco analog of the compounds, namely, **13**).

In the main component (**49**) of the mixture, the terminal Fuc group is α-(1→3)-linked (δH-1 5.112, δH-5 4.832, and δCH_3 1.176; $J_{1,2}$ 3.8 Hz) to GlcNAc-7. The location of this Fuc group is primarily evident from the shift decrement of the *N*-acetyl protons of GlcNAc-7 from δ 2.077 to 2.067, which is in accord with that described for the step from **9** to **48** (see also, Table XIX). The decrease of the intensity of the *N*-acetyl signal at δ 2.077 due to attachment of Fuc to GlcNAc-7 (compare Fig. 36 with Figs. 13 and 14a) gives independent proof of the assignment of the *N*-acetyl signals at δ 2.079 and 2.078, given before for the tetra-antennary, afuco compound **13** (see Table V). Owing to the introduction of Fuc α-(1→3)-linked to GlcNAc-7, H-1 of Gal-8 ($\Delta\delta$ −0.015 p.p.m.), H-1 of GlcNAc-7 ($\Delta\delta$ 0.013 p.p.m.), and H-1 of Gal-6 ($\Delta\delta$ − 0.006 p.p.m.) show changes in their chemical shifts, in comparison to **13**, that are very similar to those observed for the corresponding protons going from **9** to **48**. Again, the H-1 signal of GlcNAc-7 is significantly broadened.

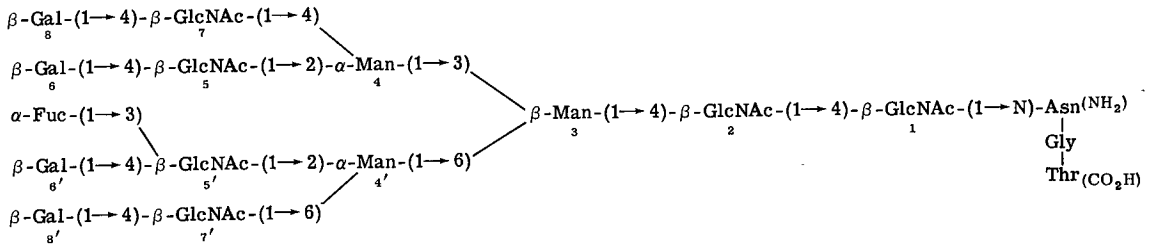
Compound **49** was earlier[67] judged to be homogeneous (by 360-MHz, ^{1}H-n.m.r. spectroscopy). However, the presence of two minor components, namely, **50** and **51** in the actual mixture can be clearly seen from the *N*-acetyl region of the 500-MHz spectrum of this sample; this conclusion is supported by the occurrence of three partially overlapping Fuc CH_3 signals, at $1.17 < \delta < 1.18$ p.p.m., and of more than one Fuc H-1 signal at $5.11 < \delta < 5.12$ p.p.m.

The locations of the Fuc groups in **50** and **51** may be inferred from the *N*-acetyl region of the spectrum. In this region, nine singlets are

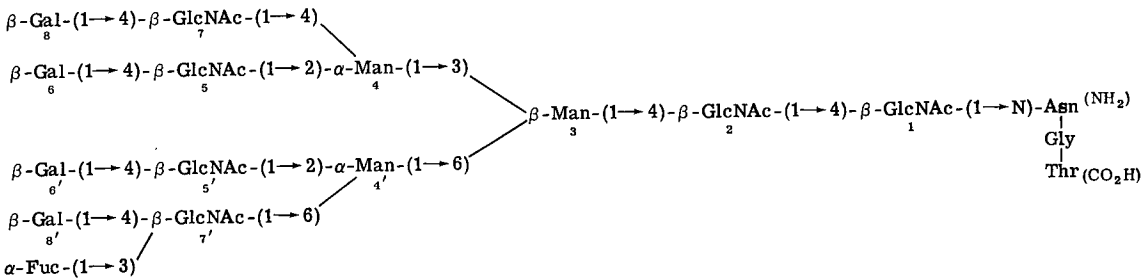
(85) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, J. Montreuil, B. Fournet, and K. Schmid, *J. Biol. Chem.*, 256 (1981) 5588–5590.



49



50



51

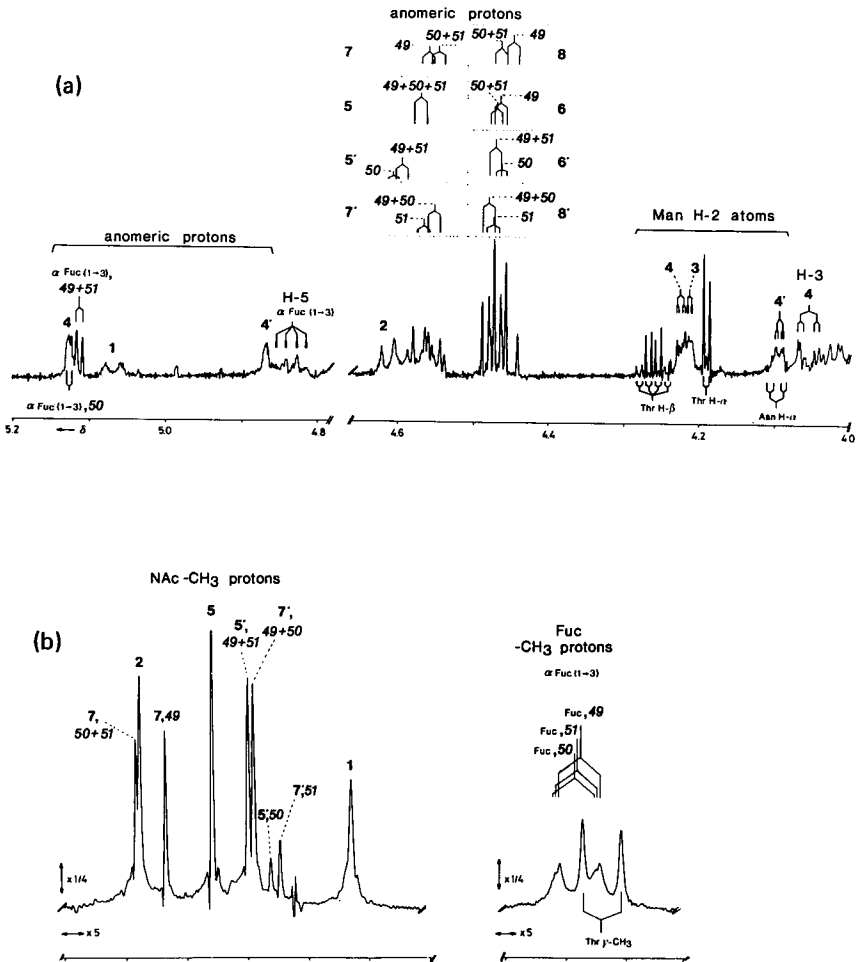


FIG. 36.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ¹H-N.m.r. Spectrum of a Mixture Containing Compounds **49**, **50**, and **51** in the Ratios of 12:2:3. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale, as well as the chemical-shift scale, of the *N*-acetyl-proton region and the Fuc CH_3-proton region (b), differ from those of the other parts (a) of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from the spectrum.]

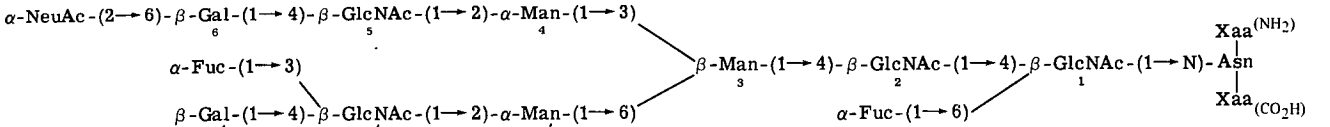
observed. The signals at δ 2.076, 2.052, and 2.006, which are of equal intensity, are assigned to GlcNAc-2, -5, and -1, respectively, in accordance with **13** (see Table V). The remaining six signals may be divided into three pairs. The signals of each pair are separated by $\Delta\delta \sim 0.01$ p.p.m. The signal pair at δ 2.077 and 2.067, having an intensity ratio of ~2:3, stems from GlcNAc-7 (see earlier). The former signal corresponds to GlcNAc-7 without Fuc (**50** and **51**), the latter to GlcNAc-7 with Fuc (**49**). Similarly, GlcNAc-5′ gives rise to a pair of signals, at δ 2.040 and 2.033, without (**49** and **51**) and with (**50**) Fuc, respectively. Also, GlcNAc-7′ gives rise to a pair of singlets, at δ 2.038 and 2.029, without (**49** and **50**) and with (**51**) Fuc, respectively. (This may be compared with the assignment of the *N*-acetyl signals of GlcNAc-5′ and -7′ for oligosaccharides **11** and **12**; see Table IV and Fig. 14b.) Therefore, **50** has its Fuc α-(1→3)-linked to GlcNAc-5′, and **51** to GlcNAc-7′. The interpretation of the *N*-acetyl region is corroborated by the evaluation of the Gal H-1 region of the spectrum (see Fig. 36 and Table XVI). For instance, the H-1 doublet of Gal-8′ for **51** is found at δ 4.473, showing an upfield shift as compared to **13**. It should be noted that the *N*-acetyl signals at δ 2.041 and 2.042 in the spectrum of **13** are suggested as belonging to GlcNAc-7′ and -5′, respectively, based on the foregoing interpretation.

The structural analysis of this glycopeptide sample (from asialo α_1-acid glycoprotein) by 500-MHz, ^{1}H-n.m.r. spectroscopy, revealing a new type of microheterogeneity, gives a nice impression of the potency of this technique in detecting and identifying compounds in mixtures, even if they are present in low proportions.

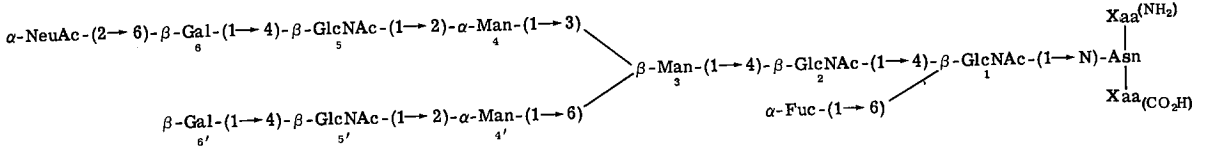
Compound **52** is a diantennary glycopeptide bearing NeuAc in α-(2→6) linkage to Gal-6, and also two differently linked Fuc groups, namely, α-(1→6)-linked to GlcNAc-1, and α-(1→3)-linked to GlcNAc-5′. Compound **52** has been obtained from human lactotransferrin,[81] as well as from[86] human-milk secretory IgA, as the minor constituent of a 3:1 mixture of glycopeptides, also containing **42**. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 37; the n.m.r. data for **52** are compiled in Table XVIII.

For the spectral interpretation, **52** is conceived of as an extension of **42** with a Fuc group α-(1→3)-linked to GlcNAc-5′. The spectral parameters of the structural-reporter groups of the fucosylated core pentasaccharide of **42** are found unaltered in **52**; the same holds for those of the sialylated, upper branch. It may be mentioned that the apparent heterogeneity of the peptide moiety of **52** (and **42**) clearly affects the

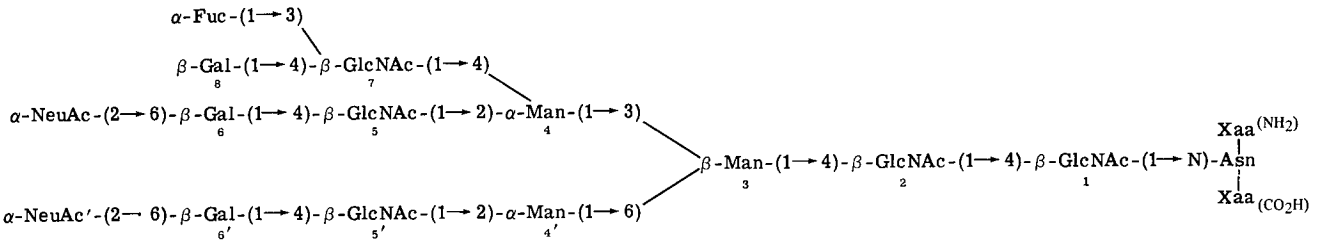
(86) A. Pierce-Crétel, M. Pamblanco, G. Strecker, J. Montreuil, G. Spik, L. Dorland, H. van Halbeek, and J. F. G. Vliegenthart, *Eur. J. Biochem.*, 125 (1982) 383–388.


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-Fuc-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)
Xaa(NH2)
Xaa(CO2H)

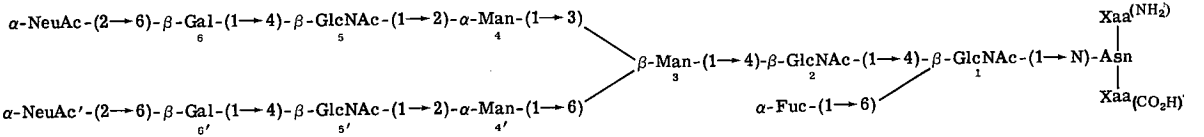
52


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)
Xaa(NH2)
Xaa(CO2H)

42


α-Fuc-(1→3)
β-Gal-(1→4)-β-GlcNAc-(1→4)
α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-NeuAc'-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
Xaa(NH2)
Xaa(CO2H)

53


α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→3)
α-NeuAc'-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→2)-α-Man-(1→6)
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
α-Fuc-(1→6)
Xaa(NH2)
Xaa(CO2H)

43

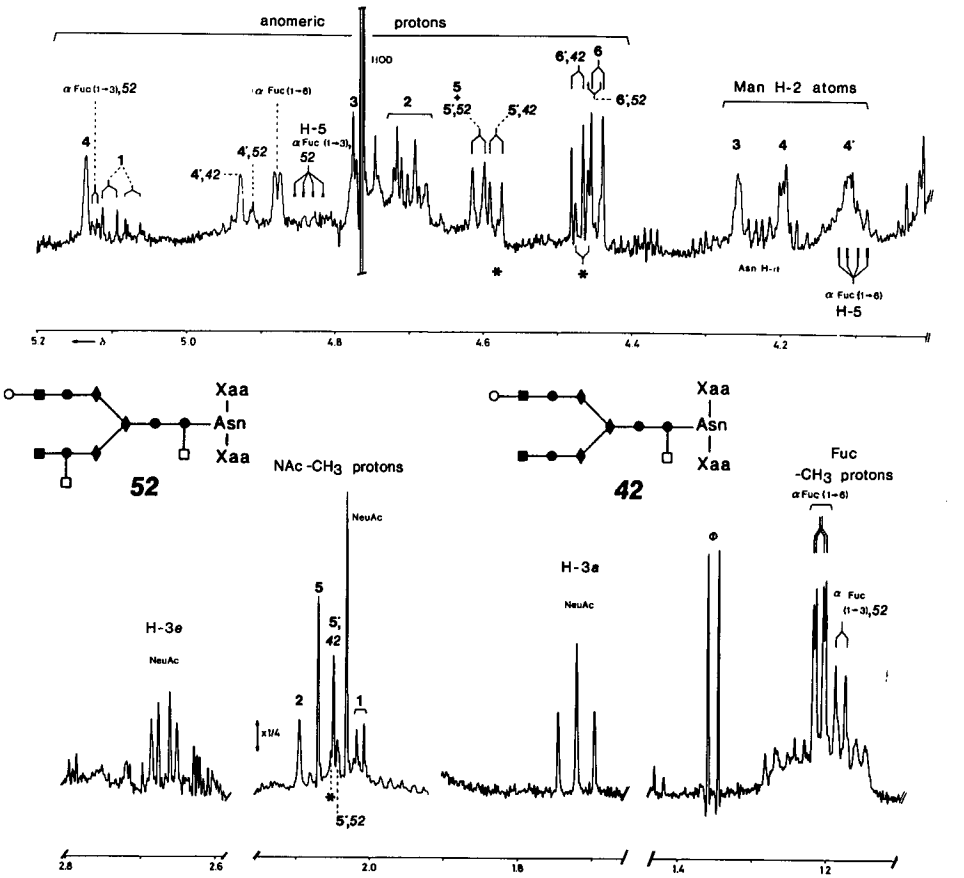


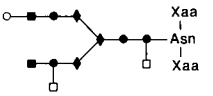
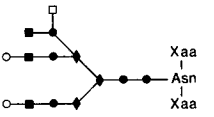
FIG. 37.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **42** and **52** in the Ratio of 3:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the two components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The signal marked by ϕ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure. For solvent-peak suppression, a water-eliminating, Fourier-transform (w.e.F.t.) pulse-sequence was used. The sample was contaminated with a very small proportion of an asialo analog of **42** or **52**, or both, as can be inferred from the signals marked by asterisks.]

chemical shifts of the structural-reporter groups of GlcNAc-1 and -2, as well as that of the CH_3 protons of Fuc α-(1→6)-linked to GlcNAc-1.

The chemical shifts of the structural-reporter groups of the α-(1→3)-linked Fuc group are in accord with those described for other compounds containing Fuc α-(1→3)-linked to GlcNAc-5′ (**47** and **50**, see

TABLE XVIII

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glycopeptides Containing Both NeuAc and Fuc as Terminating Groups (Compounds 52 and 53)

		Compound and schematic structure	
		52	**53**
			
Reporter group	**Residue**		
H-1 of	1	5.102[a]	5.036
	2	4.682[a]	4.615
	3	4.771	~4.76
	4	5.134	5.131
	4′	4.910	4.934
	5	4.605	4.59[b]
	5′	4.605	4.60[b]
	6	4.445	4.444
	6′	4.447	4.446
	7	—	4.557
	8	—	4.446
H-2 of	3	4.256	4.218
	4	4.191	4.218
	4′	4.112	4.112
H-3 of	4	<4.0	4.043
	4′	<4.0	<4.0
H-3*a* of	NeuAc	1.720	1.717
	NeuAc′	—	1.717
H-3*e* of	NeuAc	2.667	2.670
	NeuAc′	—	2.673
H-1 of	αFuc(1→3)	5.123	5.114
H-5 of	αFuc(1→3)	4.829	4.82[b]
CH_3 of	αFuc(1→3)	1.179	1.176
H-1 of	αFuc(1→6)	4.877	—
H-5 of	αFuc(1→6)	4.125	—
CH_3 of	αFuc(1→6)	1.206[a] 1.210[a]	—
NAc of	1	2.006[a] 2.016[a]	2.004
	2	2.094	2.079
	5	2.069	2.065
	5′	2.042	2.065
	7	—	2.068
	NeuAc	2.030	2.030
	NeuAc′	—	2.030

[a] Signals stemming from the main component(s) with respect to the heterogeneity of the peptide moiety. [b] Values could not be determined more accurately (±0.01 p.p.m.), due to the complexity of the mixture, of which compound **53** is one of the minor components (see also, legend to Fig. 38).

Table XVI). The location of this Fuc group may be established on the basis of the shift effects upon structural-reporter groups of neighboring residues: the *N*-acetyl protons of GlcNAc-5′ undergo a characteristic, upfield shift (from δ 2.048 for **42** to 2.042 for **52**); H-1 of Gal-6′ also shifts upfield (from δ 4.472 for **42** to 4.447 for **52**), and, last but not least, H-1 of Man-4′ is shifted significantly upfield from δ 4.927 in the asialo, lower branch of **42** to δ 4.910 in the asialo, fucosylated, lower branch of **52** (compare **47**).

Compound **53** is a triantennary glycopeptide that contains two extending units, namely, NeuAc groups α-(2→6)-linked to Gal-6 and -6′, and Fuc α-(1→3)-linked to GlcNAc-7. This compound has been derived from human-plasma ceruloplasmin[84] as a minor component of a mixture also containing compound **43**, and another two triantennary glycopeptides, both terminated with a NeuAc group in all three branches, that is, a glycopeptide analog of oligosaccharide **41**, and the Gln-Asn analog of compound **34** (compare Ref. 87). The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 38; the pertinent n.m.r. parameters for **53** are listed in Table XVIII.

The triantennary type of structure of **53** may be derived from comparison of the chemical shifts of the H-1 and H-2 signals of Man-3, -4, and -4′ with those of **9** and **48** (see Table VI). The structural-reporter-group signals for the upper (4-5-6-NeuAc) and lower (4′-5′-6′-NeuAc′) branch residues are in accord with those reported for compound **34** (see Table X).

The occurrence of an α-(1→3)-linked Fuc group attached to a peripheral GlcNAc residue is obvious from the set of chemical shifts of its structural-reporter groups (δH-1 5.114, δH-5 ~4.82, and δCH_3 1.176; δH-5 could not be established more accurately, because of the complexity of the mixture of which **53** forms a part, and of its partial coincidence with the broad, HOD signal). The chemical shifts of the *N*-acetyl signal of GlcNAc-7 (δ 2.068) and the H-1 of Gal-8 (δ 4.446) prove the attachment of Fuc α-(1→3) to GlcNAc-7. The shift decrements for the latter reporter-groups, compared to **9**, are in full accord with those observed for the asialo analog of **53**, namely, **48** (see Table XVI).

It should be noted that attachment of Fuc in α-(1→3) linkage to GlcNAc-7 also influences the chemical shift of H-1 of the sialylated Gal-6 (compare **34**; see Table X). Therefore, it may be concluded that shift effects on structural-reporter groups of neighboring residues, induced by attachment of Fuc α-(1→3) to a peripheral GlcNAc (sum-

(87) K. Yamashita, C. J. Liang, S. Funakoshi, and A. Kobata, *J. Biol. Chem.*, 256 (1981) 1283–1289.

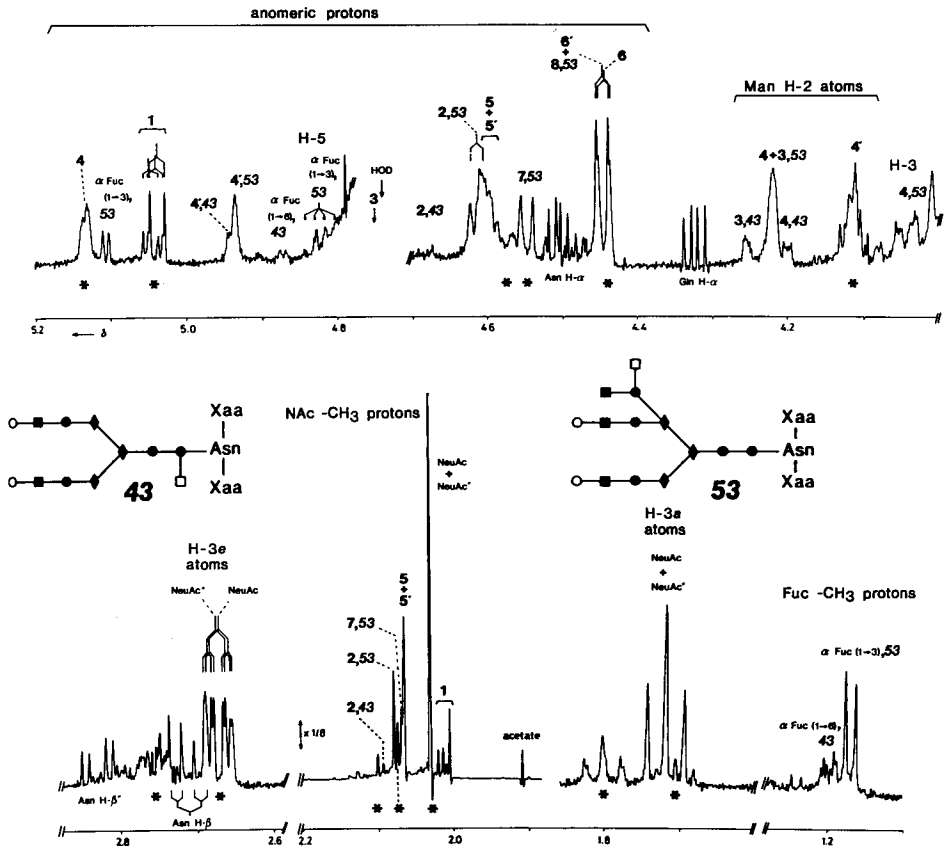


FIG. 38.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ¹H-N.m.r. Spectrum of a Mixture Containing Compounds **53** and **43** in the Ratio of 1:1. [The bold numbers in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in these two components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance, as well as the H-1 signal of Man-3, have been omitted from the spectrum; their positions are indicated by arrows. In addition to **53** and **43**, the sample contains another two triantennary glycopeptides, both terminated with NeuAc in all three branches, that is, the analogs of **53** containing NeuAc* in α-(2→3) and α-(2→6) linkage to Gal-8, respectively, instead of the Fuc group. In addition, the afuco analog of **43** is present in this mixture to a small extent. This can be inferred from the signals marked by asterisks (compare Ref. 84).]

marized in Table XIX), and those caused by extension of an *N*-acetyllactosamine unit in another branch with NeuAc in α-(2→3) or α-(2→6) linkage (see Tables XIV and XIII, respectively), are independent of each other, and are additive.

TABLE XIX

Recognition of α-(1→3)-Linked, Terminal Fuc in the Sequence βGal(1→4)[αFuc(1→3)]βGlcNAc(1→*x*)αMan(1→*y*)βMan(1→•)

^{1}H Chemical Shifts of Structural-reporter Groups of α-(1→3)-Linked Fuc

Reporter group	δ ±s.d.
H-1 (x = 2; y = 3)	5.126 ±0.002
H-1 (x = 2; y = 6)	5.124 ±0.002
H-1 (x = 4; y = 3)	5.113 ±0.002
H-1 (x = 6; y = 6)	5.112 ±0.002
H-5 (x = 2, 4, or 6; y = 3 or 6)	4.835 ±0.015[a]
CH_3 (x = 2; y = 3)	1.172 ±0.002
CH_3 (x = 2; y = 6)	1.178 ±0.002
CH_3 (x = 4; y = 3)	1.176 ±0.002
CH_3 (x = 6; y = 6)	1.177 ±0.002

Influence of α-(1→3)-Linked Fuc on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[b]

	Δδ (p.p.m.) ±s.d. (p.p.m.)
H-1 of GlcNAc (x = 2, 4, or 6; y = 3 or 6)	+0.011 ±0.003
NAc of GlcNAc (x = 2, 4, or 6; y = 3 or 6)	−0.009 ±0.002
H-1 of Gal (x = 2 or 4; y = 3 or 6)	−0.025 ±0.002
H-1 of Man (x = 2; y = 3 or 6)	−0.017 ±0.002
H-1 of Man (x = 4; y = 3)	−0.004 ±0.002

[a] δH-5 is extremely sensitive to changes in the temperature of the sample. [b] Values of Δδ are mean values ±s.d. at T = 300 K, calculated from afuco → (1→3)-fuco steps (see Tables V, VIII, X, XVI, and XVIII).

So far, we have never observed, by n.m.r. spectroscopy, oligosaccharides or glycopeptides bearing Fuc and NeuAc both linked to the same *N*-acetyllactosamine branch; this is in agreement with the biosynthetic principle of mutual exclusion existing in transferring (1→3)-linked α-fucosyl and (2→6)-linked α-sialyl groups to the same *N*-acetyllactosamine branch, as formulated by Hill and coworkers.[88]

(iii) Fuc α-(1→2)-Linked to βGal(1→4)GlcNAc.—Compound **54** is a mono-antennary glyco-asparagine bearing two terminating Fuc groups, namely, α-(1→6)-linked to GlcNAc-1, and α-(1→2)-linked to Gal-6′. This compound occurred, together with **46** and **47**, in the ratios of 1:1:6, respectively, in a mixture that could be isolated from the urine of a patient with fucosidosis.[13,56] The 500-MHz, ^{1}H-n.m.r. spec-

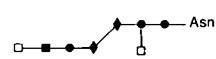
(88) J. C. Paulson, J. P. Prieels, L. R. Glasgow, and R. L. Hill, *J. Biol. Chem.*, 253 (1978) 5617–5624; T. A. Beyer, J. E. Sadler, J. I. Rearick, J. C. Paulson, and R. L. Hill, *Adv. Enzymol.*, 52 (1981) 23–175.

trum of this mixture is depicted in Fig. 34. The n.m.r. data for **54** are compiled in Table XX.

Mono-α-(1→6) substitution of Man-3 in **54** is evident from the relative intensities of several structural-reporter-group signals (in the spectrum of the mixture) that are markers for the type of mono-substitution of Man-3; it may most readily be inferred from the *N*-acetyl singlets of GlcNAc-2 at δ 2.095 (lower branch) and 2.090 (upper branch), occurring in the ratio of 7:1. As **47**, the main component of the mixture, is a lower-branch structure, whereas the other minor component, **46**, is an upper-branch glyco-asparagine (see earlier), it

TABLE XX

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for a Glyco-asparagine (Compound 54) of the *N*-Acetyllactosamine Type Containing α-(1→2)-Linked Fucose

		Compound and schematic structure
		54 
Reporter group	**Residue**	
H-1 of	1	5.073
	2	4.687
	3	4.764
	4	—
	4′	4.918
	5	—
	5′	~4.58[a]
	6	—
	6′	4.541
H-2 of	3	4.081
	4	—
	4′	4.087
H-1 of	αFuc(1→2)	5.309
H-5 of	αFuc(1→2)	~4.22[a]
CH_3 of	αFuc(1→2)	1.234
H-1 of	αFuc(1→6)	4.876
H-5 of	αFuc(1→6)	4.126
CH_3 of	αFuc(1→6)	1.206
NAc of	1	2.016
	2	2.095
	5	—
	5′	2.072

[a] Values could not be determined more accurately (±0.01 p.p.m.), due to the small amount of compound **54** present in the mixture with **46** and **47**; moreover, the H-1 doublet of GlcNAc-5′, as well as the H-5 signal of Fuc α-(1→2)-linked to Gal-6′, are relatively broad-lined.

was concluded that **54** contains the lower branch. The chemical shifts of the structural-reporter groups of the βMan(1→4)βGlcNAc(1→4)-[αFuc(1→6)]βGlcNAc(1→N)Asn moiety in **54** are identical to those for **47** (see Table XVI).

Concerning the peripheral part of the structure, **54** is considered to be an extension of (the β anomer of) the mono-antennary, lower-branch oligosaccharide **18**, with Fuc α-(1→2)-linked to Gal-6′. This additional Fuc group is characterized by the following set of structural-reporter-group signals: δH-1 5.309, δH-5 ~4.22, and δCH_3 1.234. The $J_{1,2}$ value (~4 Hz) of the Fuc group is indicative of an α-glycosidic linkage. The introduction of this Fuc group causes some shift effects on structural-reporter groups of neighboring residues. The H-1 doublet of Gal-6′ is shifted downfield to δ 4.541 (Δδ +0.069 p.p.m.); also, the *N*-acetyl signal of GlcNAc-5′ undergoes a downfield shift (Δδ +0.025 p.p.m.), but a possible influence on the chemical shift of H-1 of 5′ cannot be traced, due to the complexity of the mixture.

The chemical shifts of the structural-reporter groups of Fuc in α-(1→2) linkage to Gal of an *N*-acetyllactosamine branch, together with the influences of its introduction upon the chemical-shift values of reporter groups of neighboring residues, are summarized in Table XXI.

TABLE XXI

Recognition of α-(1→2)-Linked, Terminal Fuc in the Sequence αFuc(1→2)βGal(1→4)βGlcNAc(1→·)

^{1}H Chemical Shifts of Structural-reporter Groups of α-(1→2)-Linked Fuc[a]

Reporter group	δ ±s.d.
H-1	5.311 ±0.002
H-5	4.224 ±0.002
CH_3	1.237 ±0.003

Influences of α-(1→2)-Linked Fuc on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[b]

	Δδ (p.p.m.) ± s.d. (p.p.m.)
H-1 of Gal	+0.070 ±0.002
H-1 of GlcNAc	−0.008 ±0.004
NAc of GlcNAc	+0.025 ±0.003[c]

[a] Mean values ±s.d. at T = 300 K; see Tables XX and XXIII. [b] Values of Δδ are mean values ±s.d., calculated for afuco → (1→2)-fuco steps (see Tables VIII, XX, XXII, and XXIII). [c] The value of Δδ holds for a peripheral GlcNAc residue; for GlcNAc-1, see compound **57** as against **55**.

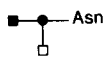
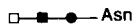
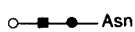
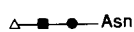
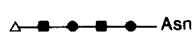
d. Carbohydrate Chains of the *N*-Acetyllactosamine Type Possessing the βGal(1→4)βGlcNAc(1→N)Asn Moiety (Compounds 55–60).—Symbols employed for compounds **55–60** are depicted in Chart 4.

Compounds **55** through **60** are glyco-asparagines of the *N*-acetyllactosamine type lacking the trimannosyl-*N*,*N*′-diacetylchitobiose core. These compounds occur, in addition to large amounts of βGlcNAc(1→N)Asn (compound **1**) and small amounts of αMan(1→6)-βMan(1→4)βGlcNAc(1→4)βGlcNAc(1→N)Asn (compound **3**), in the urine and other body fluids of patients suffering from aspartylglucosaminuria. It has been suggested[51,52,89] that βGlcNAc(1→N)Asn can serve as an acceptor compound for various glycosyltransferases, leading to the synthesis of neutral (**55–57**) or acidic (**58–60**) glyco-asparagines having in common the βGal(1→4)βGlcNAc(1→N)Asn moiety.

From the structural point of view, this series of relatively simple compounds offers the possibility of verifying and refining the n.m.r. parameters of the various units of extension of carbohydrate chains of the *N*-acetyllactosamine type, namely, Fuc and NeuAc, both in different types of linkage (compare, Tables XIII, XIV, XVII, and XXI). Moreover the parameters of an additional *N*-acetyllactosamine unit in β-(1→3) linkage to Gal of the original unit can be established (see compound **60**); such a structural unit is known to occur in various oligosac-

CHART 4

Symbols Employed[a] for Compounds 55–60

55	
56	
57	
58	
59	
60	

[a] For the key to the symbolic notation, see Chart 1.

(89) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, G. Strecker, J.-C. Michalski, J. Montreuil, W. Pollitt, and W. E. Hull, *Eur. J. Biochem.*, (1983) in press.

charides isolated from the urine of patients with GM_1-gangliosidosis,[1,64a,90] and in glycoproteins.[1-4,91-92a]

Compound **55** is a linear glyco-asparagine isolated from the urine of a patient with aspartylglucosaminuria,[52,89] as the main component of a mixture also containing **57**. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is depicted in Fig. 39. The chemical shifts of the structural-reporter groups of **55** are included in Table XXIII.

The attachment of Gal-a in β-(1→4) linkage to GlcNAc-1 of compound **1** is revealed in a significant downfield shift of the H-1 signal of this GlcNAc residue ($\Delta\delta \sim 0.03$ p.p.m.) in comparison to **1** (see Table I), as would be expected from the extension of **6** to **7** (see Table IV). The resonance positions of the structural-reporter groups of GlcNAc-1, together with its relatively large $J_{1,2}$ value (9.8 Hz), are characteristic for the β-*N*-glycosylic linkage to Asn (compare **1**). The chemical shift of the H-1 signal of Gal-a (δ 4.481) reflects the (terminal) nonre-

(90) G. Strecker, J.-C. Michalski, B. Fournet, J. Montreuil, H. van Halbeek, L. Dorland, and J. F. G. Vliegenthart, unpublished results.
(91) E. Li, R. Gibson, and S. Kornfeld, *Arch. Biochem. Biophys.*, 199 (1980) 393–399.
(92) H. Yoshima, S. Takasaki, and A. Kobata, *J. Biol. Chem.*, 255 (1980) 10,793–10,804.
(92a) T. Krusius, J. Finne, and H. Rauvala, *Eur. J. Biochem.*, 92 (1978) 289–300.

TABLE XXII

^{1}H Chemical Shifts and Coupling Constants for βGal(1→4)βGlcNAc(1→N)Asn (Compound 55, ■—●—Asn)

Residue	Proton	Chemical shift (p.p.m.)	Coupling constant (Hz)	
GlcNAc-1	H-1	5.101	$J_{1,2}$	9.8
	H-2	3.880	$J_{2,3}$	10.0
	H-3	3.778	$J_{3,4}$	8.75
	H-4	3.754	$J_{4,5}$	2.05
	H-5	3.678	$J_{5,6a}$	2.25
	H-6a	3.941	$J_{5,6b}$	4.6
	H-6b	3.837	$J_{6a,6b}$	−12.3
	NAc	2.018		
Gal-a	H-1	4.481	$J_{1,2}$	7.9
	H-2	3.549	$J_{2,3}$	9.9
	H-3	3.667	$J_{3,4}$	3.4
	H-4	3.927	$J_{4,5}$	0.9
	H-5	3.708	$J_{5,6a}$	8.7
	H-6a	3.785	$J_{5,6b}$	3.5
	H-6b	3.742	$J_{6a,6b}$	−11.6
Asn	H-α	3.985	$J_{\alpha,\beta}$	6.7
	H-β	2.875	$J_{\alpha,\beta'}$	4.5
	H-β′	2.942	$J_{\beta,\beta'}$	−17.2

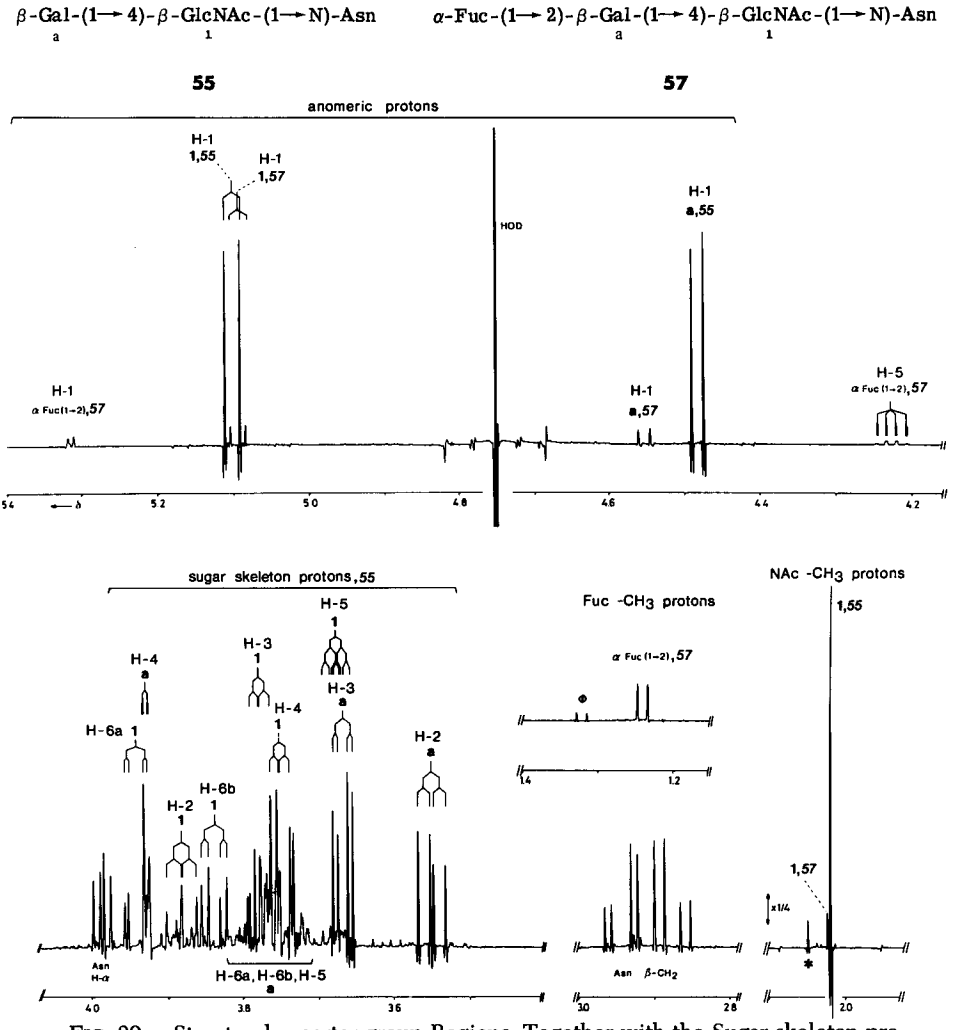


FIG. 39.—Structural-reporter-group Regions, Together with the Sugar-skeleton-proton Region, of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **55** and **57** in the Ratio of 9:1. [The bold number and letter in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The signal marked by ϕ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure. The assignments indicated in the bulk region of the spectrum belong to the main component (**55**) of the mixture. They were proved by double-resonance techniques, namely, by pseudo-INDOR and by spin-tickling experiments (see Ref. 89). Surprisingly, the sample was contaminated with a small proportion of free NeuAc, as can be inferred, among others, from the *N*-acetyl signal at δ 2.048, marked by an asterisk.]

ducing position of this group in an *N*-acetyllactosamine unit; the $J_{1,2}$ value (7.9 Hz) indicates a β-linkage.

For the relatively simple compound **55**, complete interpretation of all resonances in the 500-MHz, ^{1}H-n.m.r. spectrum was achieved.[89] The interpretation was confirmed, and refined, by computer simulation of the spectrum. The detailed, n.m.r.-spectral data for compound **55** are summarized in Table XXII. By means of a modified Karplus equation,[93] the $^4C_1$(D) conformation of the GlcNAc residue, as well as of the Gal group, was deduced.[89]

Compound **56** is a branched glyco-asparagine possessing GlcNAc-1 as the disubstituted residue; it has been isolated from the urine of a patient suffering from aspartylglucosaminuria.[52,89] The 360-MHz, ^{1}H-n.m.r. spectrum of **56** is presented in Fig. 40. The chemical shifts of the structural-reporter groups of **56** are compiled in Table XXIII.

For the spectral interpretation, compound **56** may be conceived of either as an extension of **55** with a Fuc group α-(1→6)-linked to GlcNAc-1, or as an extension of **2** with Gal-a in β-(1→4) linkage to

(93) C. A. G. Haasnoot, F. A. A. M. de Leeuw, and C. Altona, *Bull. Soc. Chim. Belg.*, 89 (1980) 125–131; *Tetrahedron*, 36 (1980) 2783–2792.

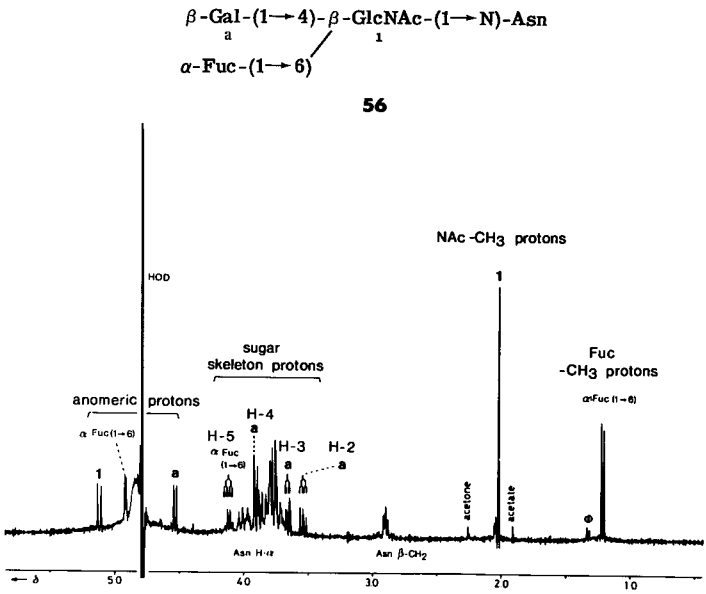


FIG. 40.—Resolution-enhanced, Overall, 360-MHz, ^{1}H-N.m.r. Spectrum of Compound **56**. [The bold number and letter in the spectrum refer to the corresponding residues in the structure. The signal marked by φ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure.]

TABLE XXIII

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glyco-asparagines Possessing the βGal(1→4)βGlcNAc(1→N)Asn Moiety (Compounds 55–60)

		Compound and schematic structure					
		55 ■—● Asn	56[a] ■—●(□) Asn	57 □—■—● Asn	58 ○—■—● Asn	59 △—■—● Asn	60 △—■—●—■—● Asn
Reporter group	Residue						
H-1 of	GlcNAc-1	5.101	5.117	5.093	5.133	5.096	5.094
	Gal-a	4.481	4.538	4.552	4.454	4.557	4.468
	GlcNAc-b	—	—	—	—	—	4.699
	Gal-c	—	—	—	—	—	4.554
H-3*a* of	αNeuAc(2→6)	—	—	—	1.709	—	—
	αNeuAc(2→3)	—	—	—	—	1.796	1.799
H-3*e* of	αNeuAc(2→6)	—	—	—	2.677	—	—
	αNeuAc(2→3)	—	—	—	—	2.759	2.756
H-3 of	Gal-a	3.667	3.665	~3.8[b]	<3.9	4.113	<4.0
	Gal-c	—	—	—	—	—	4.121
H-4 of	Gal-a	3.927	3.923	3.927	3.929	3.959	4.154
	Gal-c	—	—	—	—	—	3.959
H-1 of	αFuc(1→6)	—	4.918	—	—	—	—
H-5 of	αFuc(1→6)	—	4.123	—	—	—	—
CH_3 of	αFuc(1→6)	—	1.212	—	—	—	—
H-1 of	αFuc(1→2)	—	—	5.312	—	—	—
H-5 of	αFuc(1→2)	—	—	4.224	—	—	—
CH_3 of	αFuc(1→2)	—	—	1.240	—	—	—
NAc of	GlcNAc-1	2.018	2.024	2.022	2.042	2.015	2.014
	GlcNAc-b	—	—	—	—	—	2.032
	αNeuAc(2→6)	—	—	—	2.028	—	—
	αNeuAc(2→3)	—	—	—	—	2.031	2.030

[a] Measured at 360 MHz; T = 295 K. [b] The value could not be determined more accurately (±0.05 p.p.m.), due to the small amount of compound **57** present in the mixture with **55**.

GlcNAc-1. The set of chemical shifts of the Fuc H-1 and H-5 and the $-CH_3$ protons is indicative of this type of linkage to GlcNAc-1 (compare Table XVII). Besides for H-1 of GlcNAc-1, the presence of Gal-a, instead of GlcNAc-2, β-(1→4)-linked to GlcNAc-1 causes a significant alteration in the chemical shift of H-1 of Fuc, as compared to the mean value for this proton, observed in structures containing the fucosylated *N*, *N'*-diacetylchitobiose moiety (see Table XVII). This is an additional proof of the influence of GlcNAc-2 on the chemical shift of H-1 of Fuc α-(1→6)-linked to GlcNAc-1, as described for compound **4** and extensions thereof (see Tables XV and XVI).

As compared to **55**, attachment of Fuc in α-(1→6) linkage to GlcNAc-1 hardly affects the chemical shifts of the structural-reporter groups of GlcNAc-1, whereas the H-1 signal of Gal-a undergoes a significant, downfield shift ($\Delta\delta$ 0.057 p.p.m.). The latter shift is in perfect agreement with that observed for H-1 of GlcNAc-2 in the case of attachment of Fuc in α-(1→6) linkage to GlcNAc-1 in the normal core-structure (compare Table XVII).

Compound **57** is a linear glyco-asparagine containing a terminating Fuc group α-(1→2)-linked to Gal of the *N*-acetyllactosamine moiety (a-1). This compound was isolated from the urine of a patient with aspartylglucosaminuria,[89] as the minor (10%) component of a mixture containing **55** as the main constituent. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 39; the relevant, n.m.r.-spectral parameters for **57** are summarized in Table XXIII.

Compound **57** may be considered to be an extension of **55** with Fuc α-(1→2)-linked to Gal-a. The α-(1→2)-linked Fuc group can be recognized from the following set of structural-reporter-group signals: δH-1 5.312, δH-5 4.224, and δCH_3 1.240 (compare Table XXI). The shift effects brought about by the introduction of this Fuc group, as compared to **55**, are similar to those described for the step from **18**β to **54** (see Table XXI). The H-1 signal of Gal-a undergoes a considerable downfield shift ($\Delta\delta$ 0.071 p.p.m.); also, the *N*-acetyl signal of GlcNAc-1 is shifted downfield ($\Delta\delta$ 0.004 p.p.m.), whereas the chemical shift of the H-1 signal of GlcNAc-1 changes slightly upfield.

Compound **58** is a sialo glyco-asparagine containing NeuAc in α-(2→6) linkage to Gal-a; it has been isolated from the urine of a patient suffering from aspartylglucosaminuria.[51,52,89] The 500-MHz, ^{1}H-n.m.r. spectrum of **58** is given in Fig. 41; its pertinent n.m.r.-spectral parameters are compiled in Table XXIII.

The spectrum of **58** shows the characteristic features of the peripheral part of a mono-antennary *N*-acetyllactosamine type of structure bearing α-(2→6)-linked NeuAc (compare compounds **21**–**24**). The presence of NeuAc in α-(2→6) linkage to Gal is obvious from the set of

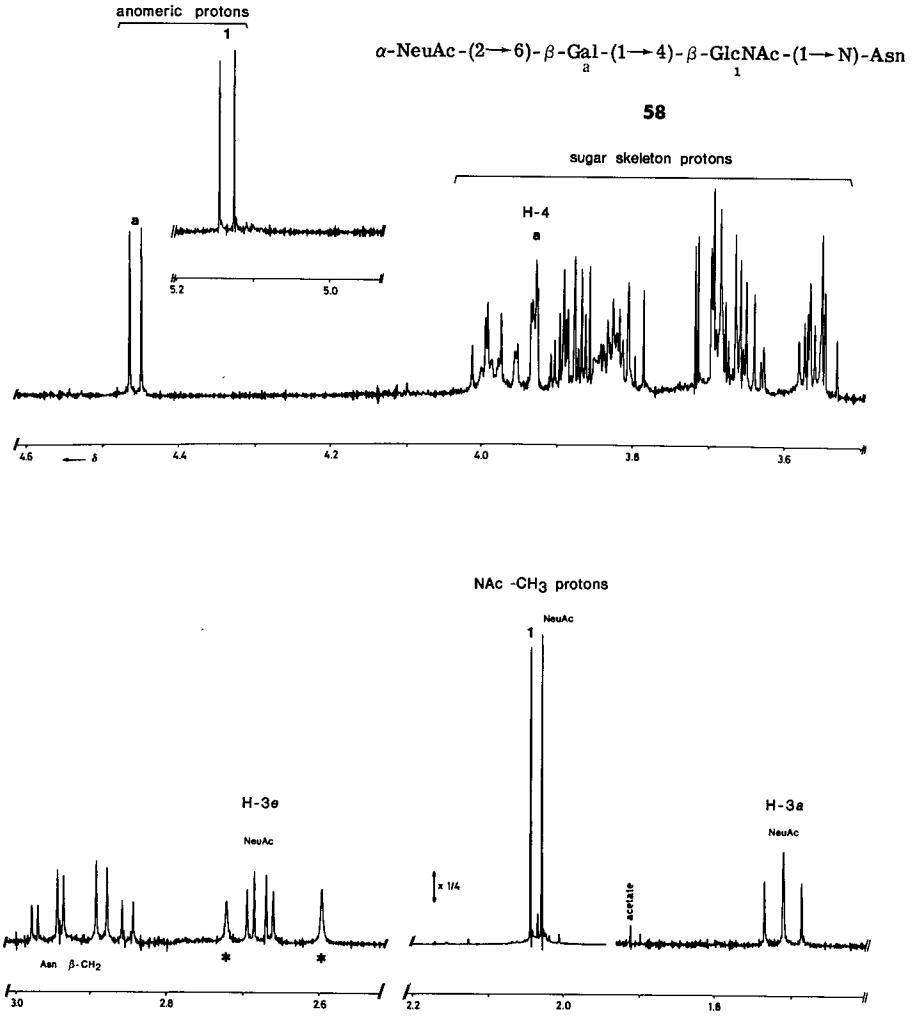


FIG. 41.—Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **58**. [The bold number and letter in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The signals marked by asterisks stem from impurities of unknown origin.]

chemical shifts of its H-3*a* (δ 1.709), H-3*e* (δ 2.677), and *N*-acetyl-group protons (δ 2.028). These values deviate slightly from those given for NeuAc in α-(2→6) linkage to an *O*-glycosylically linked *N*-acetyl-lactosamine unit (see Table XIII); this illustrates that these parameters are very sensitive to the complete structure of the branch to which NeuAc is attached.

The introduction of NeuAc at O-6 of Gal-a of **55** gives rise to a shift decrement for the H-1 signal of Gal-a ($\Delta\delta$ −0.027 p.p.m.), and to shift increments for H-1 ($\Delta\delta$ +0.032 p.p.m.) and the *N*-acetyl protons of GlcNAc-1 ($\Delta\delta$ +0.024 p.p.m.). The line widths of these signals are not affected. The changes in chemical shifts due to the introduction of α-(2→6)-linked NeuAc, observed for neighboring, structural-reporter groups, are essentially identical to those deduced for the attachment of NeuAc in α-(2→6) linkage to asialo oligosaccharides or glycopeptides of the *N*-acetyllactosamine type (see Table XIII).

Compound **59** is a sialo glyco-asparagine, containing NeuAc in α-(2→3) linkage to Gal-a. This compound has been isolated from the urine of a patient suffering from aspartylglucosaminuria.[51,52,89] The 500-MHz, ^{1}H-n.m.r. spectrum of **59** is depicted in Fig. 42; its n.m.r.-spectral parameters are listed in Table XXIII.

The presence of NeuAc in α-(2→3) linkage to Gal is evident from the chemical shifts of the NeuAc structural-reporter groups (δH-3*a* 1.796, δH-3*e* 2.759, and δNAc 2.031; compare Table XIV).

The observed, downfield shift of the H-1 signal of Gal-a in comparison to the asialo analog of **59**, namely, **55**, is $\Delta\delta$ +0.076 p.p.m., and this, together with the resonance position of H-3 of Gal-a, far apart from the bulk of the skeleton proton signals (δH-3 4.113), gives additional proof of the substitution of Gal at O-3 by NeuAc. The resonance positions of the structural-reporter groups of GlcNAc-1 are slightly, but significantly, affected ($\Delta\delta$H-1 −0.005 p.p.m., $\Delta\delta$NAc −0.003 p.p.m., as compared to **55**), which is in agreement with the limited effects, due to introduction of α-(2→3)-linked NeuAc, on the chemical shifts of structural-reporter groups of neighboring residues, as summarized in Table XIV.

Compound **60** is a sialo glyco-asparagine of the *N*-acetyllactosamine type containing NeuAc in α-(2→3) linkage to Gal-c. The latter residue forms part of an additional *N*-acetyllactosamine unit (b-c) which is β-(1→3)-linked to the fundamental structure (**55**). Compound **60** has been isolated from the urine of a patient with aspartylglucosaminuria[89] as the major component of a mixture also containing small amounts of **58** and of the asialo analog of **60**. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig. 43; the pertinent n.m.r.-spectral parameters of **60** are compiled in Table XXIII.

The presence of the additional, sialylated *N*-acetyllactosamine unit in β-(1→3) linkage to Gal-a is revealed by the occurrence of the following structural-reporter-group resonances: (*i*) the H-1 doublet of the β-(1→3)-linked GlcNAc-b at δ 4.699, (*ii*) the additional singlet of GlcNAc-b in the *N*-acetyl-proton region of the spectrum, at δ 2.032, and (*iii*) the H-1 doublet of Gal-c at δ 4.554.

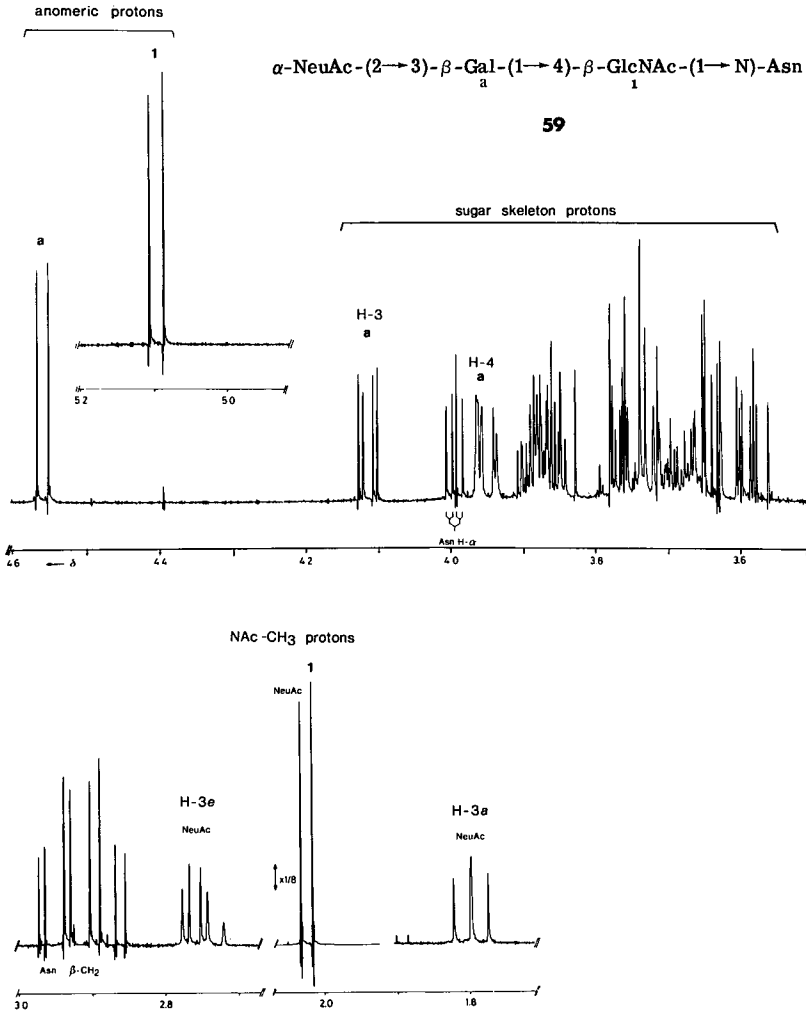


FIG. 42.—Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **59**. [The bold number and letter in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

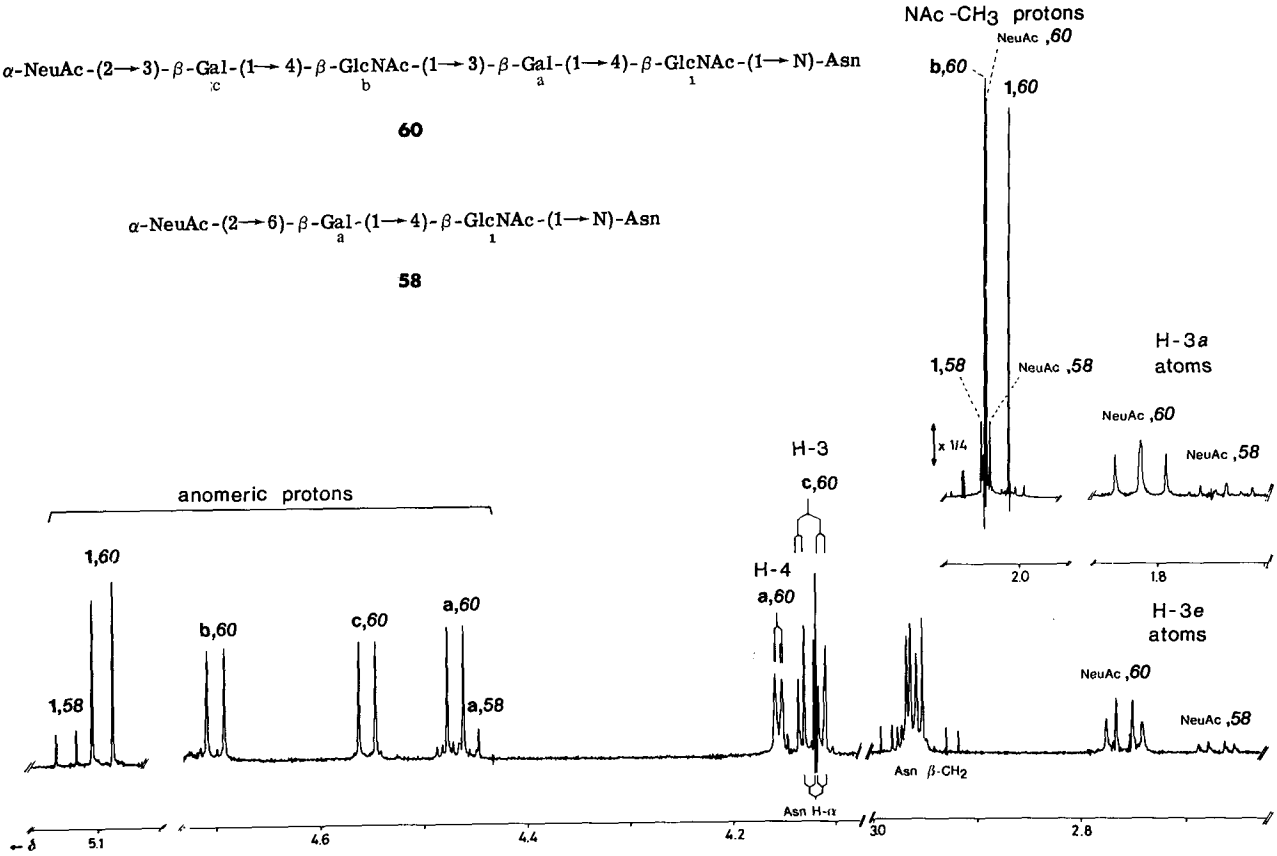
The introduction of a (sialylated) *N*-acetyllactosamine unit β-(1→3)-linked to Gal-a gives rise to very few alterations in chemical shifts, coupling constants, and line widths of structural-reporter groups of neighboring residues that can be used for the localization of additional *N*-acetyllactosamine units in more-complex structures (compare Refs. 91–92a). The influences of this unit are almost exclusively re-

α-NeuAc-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→3)-β-Gal-(1→4)-β-GlcNAc-(1→N)-Asn
c b a 1

60

α-NeuAc-(2→6)-β-Gal-(1→4)-β-GlcNAc-(1→N)-Asn
a 1

58



stricted to the protons of Gal-a. The considerable, downfield shift of the H-4 signal, from δ 3.927 for **55** to δ 4.154 for **60**, is most striking. For the assignment of this signal in the spectrum, its typical shape, namely, a relatively broad-lined, virtual doublet ($J_{4,5}$ of Gal < 1 Hz) is extremely helpful. The combination of a characteristic pattern and a resonance position clearly resolved from the bulk (δ ~4.15) makes the Gal H-4 an excellent, structural-reporter group for the recognition of a β-(1→3) substitution of this Gal residue by, for example, an *N*-acetyllactosamine unit (see also Ref. 12). Furthermore, the resonance positions of H-1 (Δδ ~ −0.013 p.p.m.) and of H-3 of Gal-a (Δδ ~ 0.3 p.p.m.) are significantly influenced by this extension. Finally, it should be noted that the *N*-acetyl signal of GlcNAc-1 undergoes a small upfield shift (Δδ −0.004 p.p.m.) due to the elongation of the glycan chain.

In **60**, the additional *N*-acetyllactosamine unit is terminated by NeuAc in α-(2→3) linkage to Gal-c; this can be inferred from comparison of the set of chemical shifts of the NeuAc structural-reporter groups for **60** with those reported to be characteristic for the αNeuAc-(2→3)βGal(1→4)βGlcNAc(1→•) sequence (compare Table XIV). As the influences of sialic acid in α-(2→3) linkage to Gal are well described (see Table XIV), the values of chemical shift of the structural-reporter-group signals for an additional, asialo *N*-acetyllactosamine unit can be predicted.

2. Carbohydrate Chains of the Oligomannoside Type

In the second family of *N*-glycosylically linked carbohydrate chains, the core pentasaccharide αMan(1→3)[αMan(1→6)]βMan(1→4)-βGlcNAc(1→4)βGlcNAc(1→N)Asn is substituted by mannose residues.[1]

The n.m.r.-spectral study of such glycopeptides and oligosaccharides is hampered by the similarity of the constituent units. Only the H-1 and H-2 resonances of each of the Man residues can serve as markers for the primary structure. Obviously, this demands high spectral resolution on increase in the number of Man residues.

Crucial features for spectral assignments for this type of carbohy-

FIG. 43.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **60** and **58** in the Ratio of 9:1. [In contravention of the usual experimental conditions (see Section IV), the glycoasparagine solution used for recording this spectrum had pD ~2. The bold number and letters in the spectrum refer to the corresponding residues in the structures, and the italic numbers refer to the compounds in the mixture. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other parts of the spectrum, as indicated.]

drate structure are the parameters of the structural-reporter groups of α-(1→2)-linked Man residues, as well as the manifestation of the second branching-point, which is a Man-4′ residue, substituted in the same way as Man-3. The complete assignment of all structural-reporter groups for structures containing up to nine Man residues will be discussed.

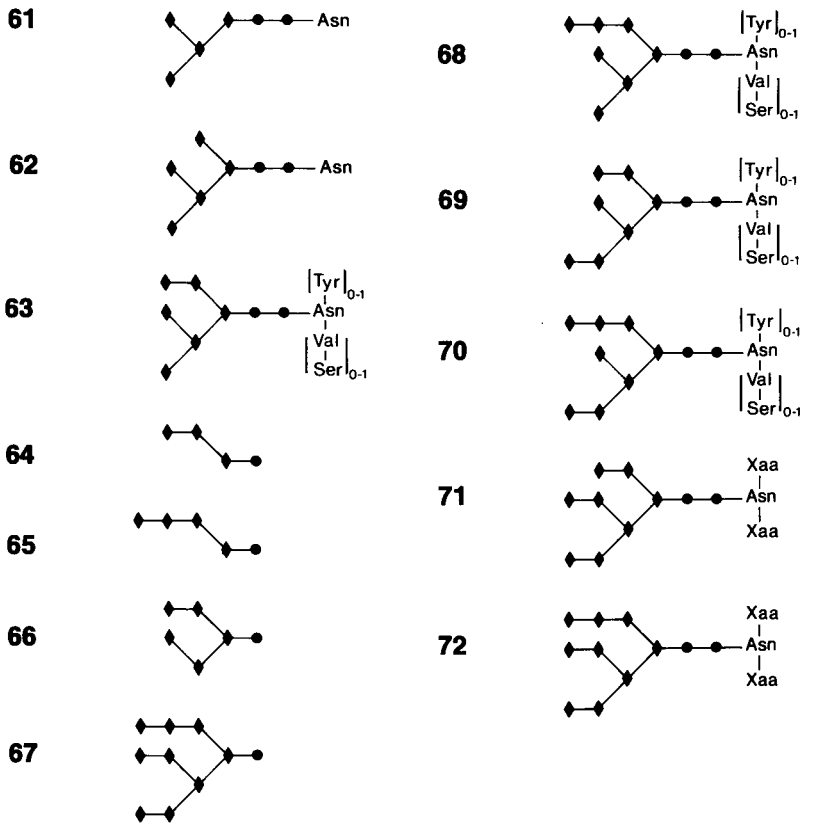
Symbols employed for compounds **61–72** are depicted in Chart 5.

Compound **61** is a glyco-asparagine of the oligomannoside type containing four Man residues that has been isolated[58,59] from the urine of a patient with Gaucher's disease, a glucocerebrosidase deficiency.[94]

(94) E. G. Brunngraber, in Ref. 51, pp. 143–144.

CHART 5

Symbols Employed[a] for Compounds 61–72



[a] For the key to the symbolic notation, see Chart 1.

The 500-MHz, ^{1}H-n.m.r. spectrum of **61** is presented in Fig. 44, and its n.m.r. parameters are listed in Table XXIV.

For the spectral interpretation, **61** is conceived of as an extension of compound **3** with Man-A and -B. The latter groups, together with Man-4′, constitute the second branching-point, which is characteristic for oligomannoside-type structures.[1] Compared to the spectrum of **3**, the resonance positions of H-1 and H-2 of Man-3 are found essentially unaltered; this indicates that **61** contains a mono-α-(1→6)-substituted Man-3 residue (compare the chemical shifts of H-1 and H-2 of Man-3 in other mono-antennary, lower-branch, glyco-asparagines and oligosaccharides, for example, **4**, **18**, **22**, **47**, and **54**). The H-1 and *N*-acetyl proton signals of GlcNAc-2 are shifted somewhat upfield, to δ 4.608 and 2.061, respectively, due to extension of the α-(1→6)-linked Man-4′ by Man groups only. In this type of compound, these values are not markers for the type of mono-substitution of Man-3, in contrast to structures of the *N*-acetyllactosamine type, but now, these values proved to be

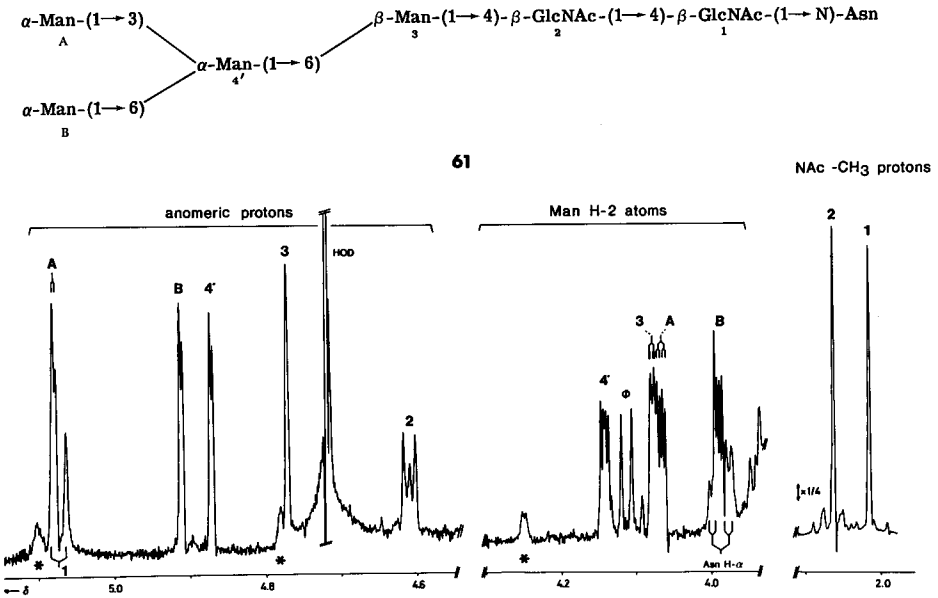
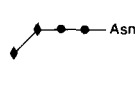
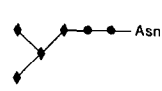
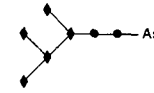
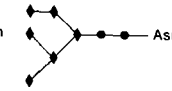


FIG. 44.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **61**. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The signal marked by φ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure. In addition to **61**, the sample contained a small proportion of glyco-asparagine **62**, as can be inferred from the signals marked by asterisks.]

TABLE XXIV

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Small Glyco-asparagines of the Oligomannoside Type (Compounds 3 and 61–63)

		Compound and schematic structure			
		3	**61**	**62**	**63** (ov)
					
Reporter group	**Residue**				
H-1 of	1	5.071	5.069	5.071	5.070
	2	4.618	4.608	4.606	4.605
NAc of	1	2.014	2.013	2.012	2.011
	2	2.076	2.061	2.060	2.061
H-1 of	3	4.767	4.770	4.781	4.770
	4	—	—	5.099	5.342
	4′	4.915	4.870	4.872	4.871
	A	—	5.076	5.093	5.095
	B	—	4.909	4.908	4.908
	C	—	—	—	5.052
H-2 of	3	4.080	4.076	4.251	4.229
	4	—	—	4.077	4.114
	4′	3.968	4.140	4.144	4.143
	A	—	4.064	4.066	4.066
	B	—	3.988	3.985	3.990
	C	—	—	—	4.066

characteristic for the GlcNAc-2 core residue in glycopeptides of the oligomannoside type.

Concerning the characteristics of the peripheral part of the structure, the well resolved doublet at δ 4.909 ($J_{1,2}$ 1.8 Hz) and the narrow doublet of doublets at δ 3.988 are attributed to H-1 and H-2, respectively, of the terminal, α-(1→6)-linked Man-B group. These values closely resemble those of H-1 and H-2 of the terminal α-(1→6)-linked Man-4′ in **3**, **4**, **19**β, and **23**β. The H-1 doublet at δ 5.076 ($J_{1,2}$ 1.8 Hz), in conjunction with the H-2 resonance at δ 4.064, is characteristic for the terminal, α-(1→3)-linked Man-A group, as can be derived from comparison with the n.m.r.-spectral data for compounds containing a terminal, α-(1→3)-linked Man-4, for example, **5**, **20**β, and **24**β.

The H-1 signal of the disubstituted Man-4′ is found at δ 4.870 ($J_{1,2}$ 1.8 Hz). The shift decrement (Δδ −0.045 p.p.m.) observed for this proton in comparison to **3** is ascribed to the substitution of Man-4′ at O-6, as the presence of an α-(1→3)-linked Man group hardly influences the resonance position of H-1 of the Man residue to which it is attached

(compare the steps from **18** to **20**, Table VIII; and from **22** to **24**, Table IX). A similar shift-decrement is observed for H-1 of Man-4′ in the steps from di- to tri′-antennary, and from tri- to tetra-antennary, structures of the *N*-acetyllactosamine type (see Table VI), involving extension with a βGal(1→4)GlcNAc moiety in β-(1→6) linkage to Man-4′. The H-2 signal of Man-4′ is found at δ 4.140, showing a shift increment (Δδ 0.172 p.p.m.) as compared to **3**. This downfield shift can be attributed almost exclusively to the attachment of the α-(1→3)-linked Man-A, as an α-(1→6) substitution of one Man residue by another hardly affects the H-2 chemical shift of the former (compare the steps from **17** to **19**, Table VIII; and from **21** to **23**, Table IX).

Compound **62** is a glyco-asparagine of the oligomannoside type containing five Man residues that has been isolated from the urine of a patient with Gaucher's disease[58,59]; it was also obtained from hen-egg albumin.[95] The 500-MHz, ^{1}H-n.m.r. spectrum of **62** is presented in Fig. 45; its n.m.r. parameters are compiled in Table XXIV.

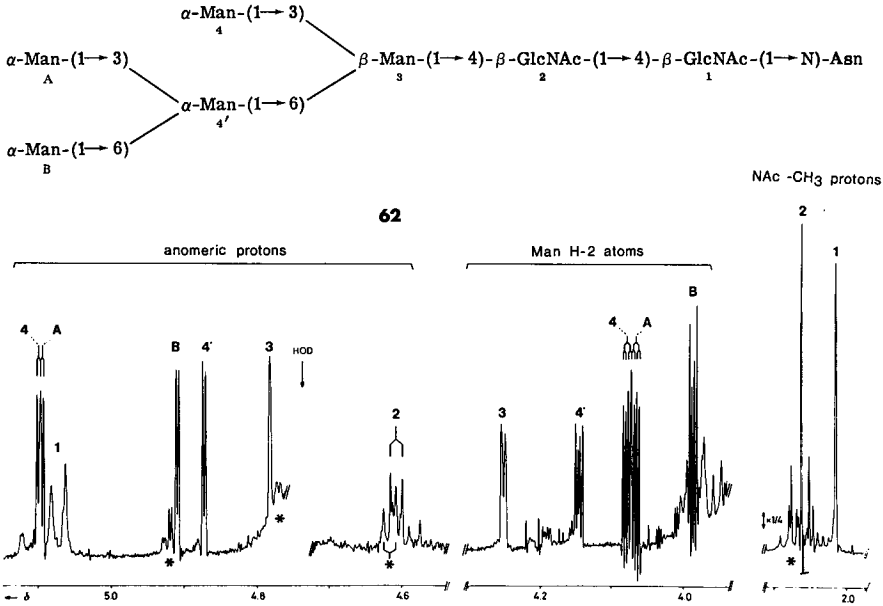


FIG. 45.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **62**. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated. The HOD resonance has been omitted from the spectrum; its position is indicated by an arrow. In addition to **62**, the sample contained small proportions of structurally related, oligomannoside-type glyco-asparagines, for example, compound **3**; some characteristic, reporter-group signals of the latter compound are marked by asterisks.]

For the interpretation of the spectrum, **62** is conceived of as an extension of **61** with Man-4 in α-(1→3) linkage to Man-3, also completing the first branching-point. The additional Man-4 is characterized by its H-1 (δ 5.099) and H-2 signal (δ 4.077). The chemical shifts of these H-1 and H-2 signals are essentially the same as those described for the terminal Man-4 in compounds **5**β (see Table III), **20**β (see Table VIII), and **24**β (see Table IX).

Owing to the introduction of Man-4, the characteristically shaped H-2 signal of Man-3 in particular undergoes a downfield shift, from δ 4.076 for **61** to δ 4.251, whereas the chemical shift of H-1 of Man-3 is only slightly affected ($\Delta\delta$ ~0.01 p.p.m.) (compare the influence ascribed to the introduction of Man-A upon the reporter groups of Man-4′, in the step from **3** to **61**). The effects of introduction of Man-4 into **61** are identical to those described for the steps from **18** to **20** (see Table VIII) and from **22** to **24** (see Table IX).

As compared to **61**, the chemical shifts of all other structural-reporter groups (including those of GlcNAc-2) remain essentially unaltered, except for the H-1 signal of Man-A. The change in the chemical shift of the latter proton may reflect a spatial effect. It should be noted that, at lower magnetic field strength, the H-1 signals of Man-4 and -A for **62** are indistinguishable from each other.[46,50]

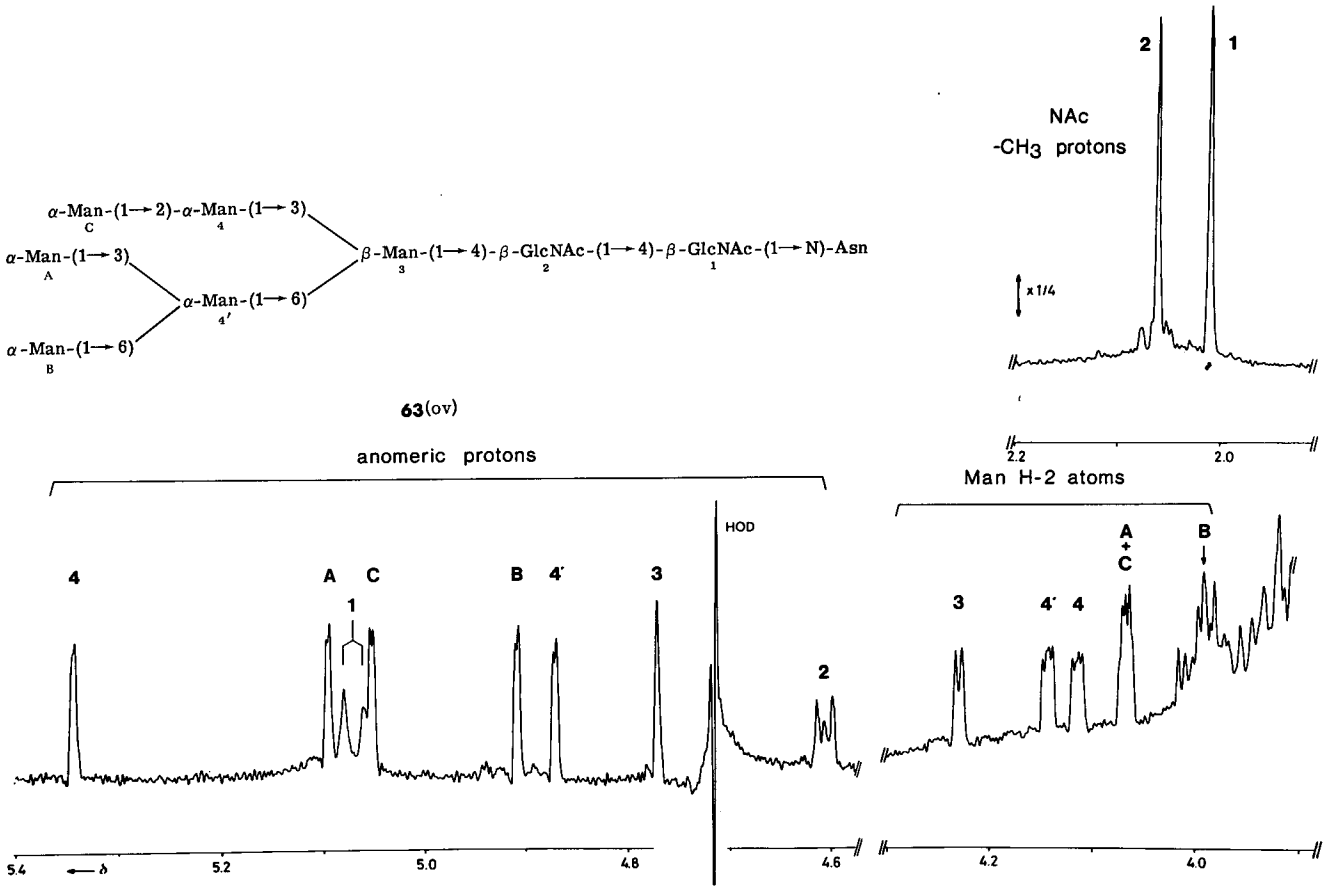
Compound **63** (ov) is a glyco-asparagine of the oligomannoside type containing six Man residues that has been obtained from hen-egg albumin.[95] The 500-MHz, ^{1}H-n.m.r. spectrum of **63** (ov) is given in Fig. 46, and its n.m.r.-spectral parameters are included in Table XXIV.

For the spectral interpretation, **63** is conceived of as an extension of **62**, with Man-C in α-(1→2) linkage to Man-4. The additional Man-C is characterized by its H-1 (δ 5.052) and H-2 (δ 4.066) signals. This set of chemical shifts for a Man residue has not been observed for the foregoing compounds of the oligomannoside type (**61** and **62**); it proves to be typical for a terminal, α-(1→2)-linked Man group.

Owing to the introduction of Man-C, the structural-reporter-group signals of Man-4 undergo downfield shifts ($\Delta\delta$H-1, 0.243; $\Delta\delta$H-2, 0.037 p.p.m.). The H-1 and H-2 signals of Man-3 are both shifted slightly up-

(95) T. Tai, K. Yamashita, A.-M. Ogata, N. Koide, T. Muramatsu, S. Iwashita, Y. Inoue, and A. Kobata, *J. Biol. Chem.*, 250 (1975) 8569–8575.

FIG. 46.—*Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz,* ^{1}H-N.m.r. *Spectrum of Compound* **63**. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region (see insertion) differs from that of the other parts of the spectrum, as indicated. For solvent-peak suppression, a w.e.F.t. pulse-sequence was used.]


α-Man-(1→2)-α-Man-(1→3)
C 4
α-Man-(1→3)
A
α-Man-(1→6)
4'
α-Man-(1→6)
B
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
3 2 1
63(ov)
anomeric protons
4
A C
1
B
4'
3
HOD
2
5.4
δ
5.2
5.0
4.8
4.6
NAc
-CH3 protons
2
1
x1/4
2.2
2.0
Man H-2 atoms
3
4'
4
A
+
C
B
4.2
4.0

field ($\Delta\delta$H-1 −0.011 and $\Delta\delta$H-2 −0.022 p.p.m.). The chemical shifts of the structural-reporter groups of the other residues remain essentially unaltered, in comparison to **62**.

Compound **64** is a reducing tetrasaccharide that has been isolated from the urine of a patient suffering from mannosidosis.[52,61] The 500-MHz, ^{1}H-n.m.r. spectrum of **64** is presented in Fig. 47, and its n.m.r.-spectral parameters are summarized in Table XXV.

For the spectral interpretation, **64** can be conceived of as an extension of trisaccharide **5**, with Man-C in α-(1→2) linkage to Man-4. The structural-reporter groups, H-1 and H-2, of Man-C give rise to relatively sharp-lined signals, at δ 5.050 and 4.069, respectively. This set of chemical shifts is typical for a terminal, α-(1→2)-linked Man group (compare compound **63**, see Table XXIV).

Upon attachment of Man-C to Man-4, the H-1 and H-2 signals of Man-4 undergo downfield shifts ($\Delta\delta$ 0.245 and 0.034 p.p.m., respectively) and are slightly broadened. The chemical shifts of the structural-reporter groups of the other residues, Man-3 and GlcNAc-2, are unaffected by the presence of Man-C; together, they reflect the α-(1→3) type of mono-substitution of Man-3 (compare **5**, Table III; **17**, Table VIII; **21**, Table IX; **35**, Table XI; **40**, Table XII; and **46**, Table XVI). The effect of anomerization on the chemical shift of structural-

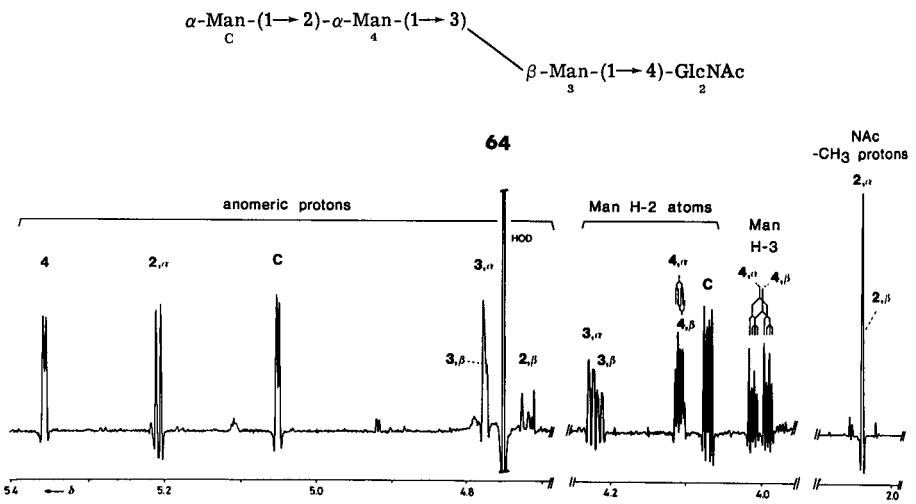


FIG. 47.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **64**. [The bold numbers and the letter in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **64**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated.]

TABLE XXV

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Oligosaccharides of the Oligomannoside Type (Compounds 5, and 64–67)

Reporter group	Residue	Anomer of compound	Compound and schematic structure 5	64	65	66[a]	67
H-1 of	2	α	5.209	5.206	5.207	5.250	5.231
		β	4.718	4.718	4.719	4.716	4.714
NAc of	2	α	2.043	2.041	2.043	2.044[a]	2.050
		β	2.041		2.041		2.046
H-1 of	3	α	4.787	4.776	4.774	4.776[a]	4.776
		β	4.783	4.772	4.771		4.772
	4	α,β	5.111	5.356	5.343	5.348	5.337
	4′	α	—	—	—	4.899	4.872
		β					4.869
	A	α	—	—	—	5.098	5.398
		β				5.126	5.407
	B	α,β	—	—	—	—	5.142
	C	α,β	—	5.050	5.296	5.056	5.308
	D_1	α,β	—	—	5.046	—	5.048
	D_2	α	—	—	—	—	5.058
		β					5.063
	D_3	α,β	—	—	—	—	5.040
H-2 of	3	α	4.244	4.224	4.220	4.238	4.239
		β	4.233	4.213	4.209	4.227	4.229
	4	α	4.075	4.108	4.087	4.109[a]	4.089
		β	4.071	4.105	4.082		
	4′	α	—	—	—	4.129	4.158
		β				4.127	4.155
	A	α	—	—	—	4.068	4.106
		β				4.065	4.103
	B	α,β	—	—	—	—	4.025
	C	α,β	—	4.069	4.105	4.068	4.109
	D_1	α,β	—	—	4.066	—	4.069[b]
	D_2	α	—	—	—	—	4.073[b]
		β					
	D_3	α,β	—	—	—	—	4.066[b]

[a] Measured at 360 MHz, T = 295 K. [b] Assignments may have to be interchanged.

reporter groups decreases, going from the reducing residue to the nonreducing end-group; at 500 MHz, neither H-1 nor H-2 of Man-C shows a difference in chemical shift between the α and β anomer of **64**.

Compound **65** is a linear, reducing pentasaccharide containing four Man residues that has been isolated from the urine of a patient suffering from mannosidosis.[52,61] The 500-MHz, ^{1}H-n.m.r. spectrum of **65** is depicted in Fig. 48, and its relevant, n.m.r.-spectral parameters are listed in Table XXV.

For the spectral interpretation, **65** can be conceived of as an extension of **64** with another α-(1→2)-linked Man group, denoted as Man-D_1, at the nonreducing end. The chemical shifts, coupling constants, and line widths of the H-1 and H-2 signals of Man-D_1 are essentially identical to those of the terminal, α-(1→2)-linked Man-C in **64**.

The α-(1→2) substitution of Man-C by Man-D_1 leads to shift increments of the H-1 and H-2 signals of Man-C: $\Delta\delta$ 0.246 and 0.036 p.p.m., respectively. These increments are identical to the effects described for H-1 and H-2 of Man-4 due to elongation of **5** to **64** with Man-C (see earlier). Attachment of Man-D_1 in α-(1→2) linkage to Man-C also influences the chemical shifts of the H-1 and H-2 signals of Man-4 ($\Delta\delta$ −0.013 and −0.022 p.p.m., respectively).

The line widths of the H-1 doublets of the α-linked Man residues in **65** decrease on going from the internal Man-4 by way of Man-C to the terminal Man-D_1. Doubling of signals due to anomerization of **65** is manifested only in the signals of GlcNAc-2 and Man-3. The more remote a residue from the reducing end, the less the effect of the ano-

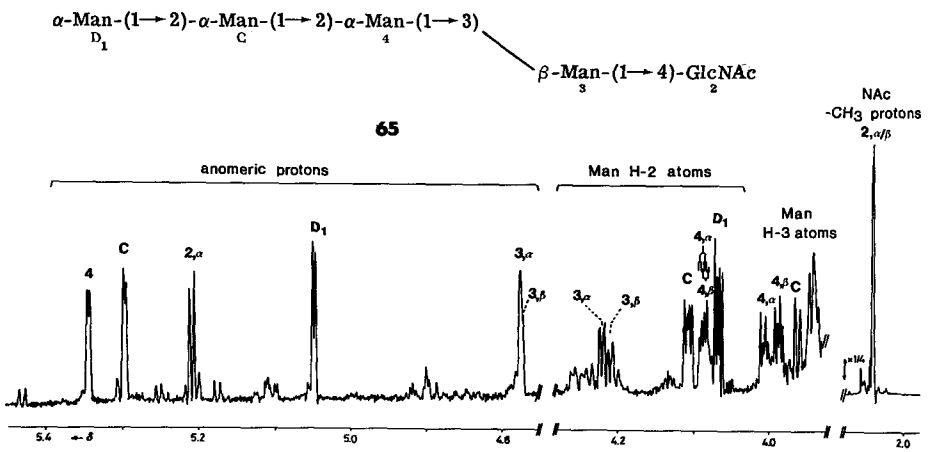


FIG. 48.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **65**. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **65**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

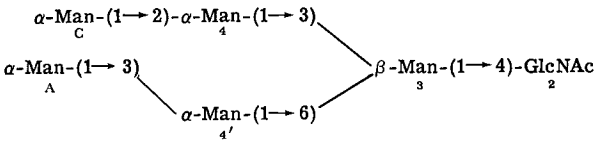
meric configuration of GlcNAc-2 on the chemical shift of its structural-reporter groups.

Compound **66** is a branched-chain, reducing oligosaccharide containing five Man residues that has been isolated from the urine of a patient with mannosidosis. The 360-MHz, ^{1}H-n.m.r. spectrum of **66** is depicted in Fig. 49; its n.m.r.-spectral features are listed in Table XXV.

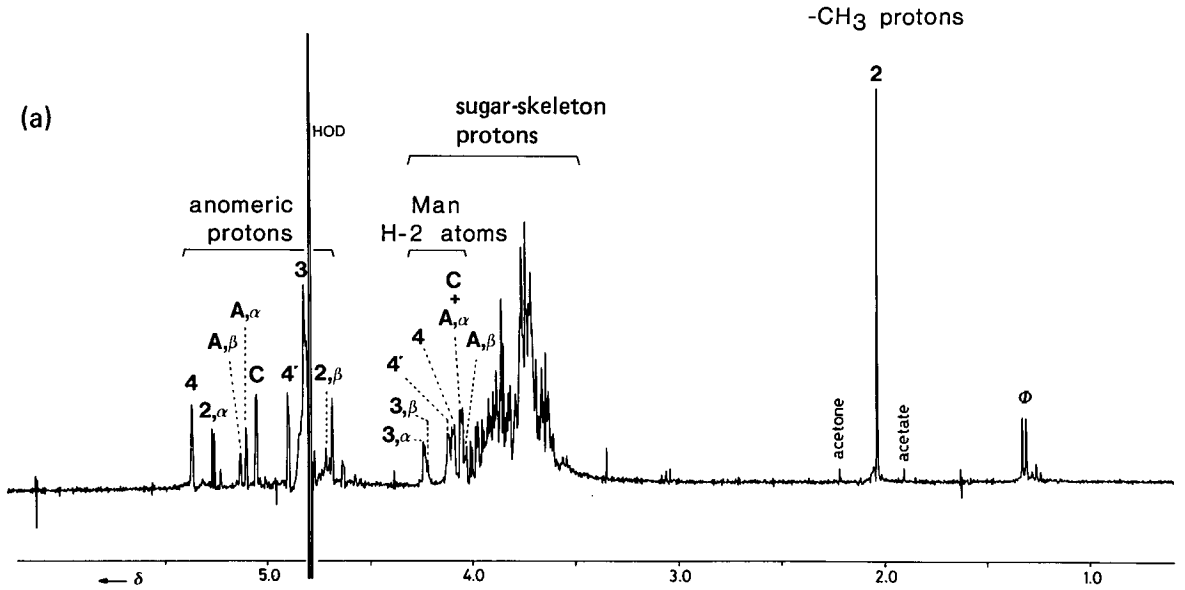
For the interpretation of the spectrum, **66** is considered to be an extension of **64**, with an αMan(1→3)Man unit (A-4′) in α-(1→6) linkage to Man-3. For H-1 of the terminal, α-(1→3)-linked Man-A group, two well-separated doublets are found, at δ 5.098 and 5.126, with relative intensities in the anomeric ratio, $\alpha:\beta = 2:1$. The H-2 signal of Man-A is also doubled, due to anomerization. The set of chemical-shift values for H-1 and H-2 of Man-A is very similar to the sets observed for terminal, α-(1→3)-linked Man groups in compounds **5** (Man-4; see Table III), **20** (Man-4; see Table VIII), **24** (Man-4; see Table IX), **61** (Man-A; see Table XXIV), **62** (Man-4 and -A; see Table XXIV), and **63** (Man-A; see Table XXIV). However, a value for the anomerization effect as large as $|\Delta\delta_{\alpha-\beta}|$ for H-1 of Man-A in **66** (0.028 p.p.m.) had not been observed before; it indicates a very close spatial proximity of this H-1 to the anomeric center of GlcNAc-2. This is mutually proved by the shift increment (Δδ ~0.04 p.p.m.) observed for H-1 of GlcNAc-2 in the α anomer of **66** in comparison to that in related structures, for example, **5**, **64**, or **65** (see Table XXV).

The H-1 doublet of Man-4′ is observed at δ 4.899; its position is only little affected by the attachment of Man-A in α-(1→3) linkage [compare with compounds **3**, **4**, **19**, **23**, and with **61**, **62**, and **63**, having in common a terminal, α-(1→6)-linked Man group, namely, 4′ and B, respectively; see Tables III, VIII, IX, and XXIV]. However, the chemical shift of the H-2 signal of Man-4′ (δ ~4.128) differs considerably from that of H-2 of the terminal α-(1→6)-linked group in the aforementioned series of compounds (δ ~3.98; signal almost hidden in the bulk resonance of skeleton protons). The shift increment observed must be attributed to the substitution of Man-4′ by Man-A; it is essentially identical to that described for H-2 of Man-3 upon introduction of Man-4 (in the steps from **18** to **20**, from **22** to **24**, and from **61** to **62**; see Tables VIII, IX, and XXIV). Thus, the spectral data for compound **66** offered the possibility of unraveling the influences of the attachment of Man-A on the chemical shifts of the structural-reporter groups of neighboring residues, apart from those of attachment of Man-B (compare **66** as against **61**–**63**), as **66** is the only compound reported herein having an incomplete, second branching-point.

As described for the step from **17** to **19**, the introduction of the α-


α-Man-(1→2)-α-Man-(1→3)
C 4
α-Man-(1→3)
A
β-Man-(1→4)-GlcNAc
3 2
α-Man-(1→6)
4'

66


-CH3 protons
(a)
HOD
sugar-skeleton protons
anomeric protons
Man H-2 atoms
2
3
C + A,α
A,α
A,β
C
4
4'
2,β
2,α
4
4'
3,β
3,α
A,α
A,β
acetone
acetate
Φ
← δ
5.0
4.0
3.0
2.0
1.0

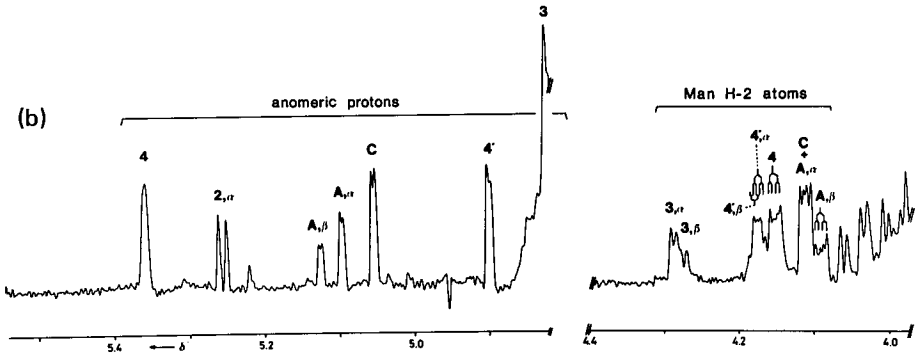


FIG. 49.—(a) Resolution-enhanced, Overall, 360-MHz, ^{1}H-N.m.r. Spectrum of Compound **66**. (b) Expanded, Structural-reporter-group Regions of Spectrum (a). [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **66**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The signal marked by ϕ originates from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure.]

(1→6)-linked Man-4′ mainly influences the chemical shift of H-2 of Man-3 ($\Delta\delta$ 0.014 p.p.m.); the chemical shift of H-1 of Man-4 is also slightly affected (see Table XXV) (compare compound **19**). The presence of Man-A apparently does not alter the magnitude of these effects. The chemical shifts of the structural-reporter groups of Man-C are unaltered with respect to those of the corresponding protons in **64**. As already described (compare compounds **61** and **62**), disubstitution of Man-3 does not result in a significant alteration of the chemical shift of the *N*-acetyl signal of GlcNAc-2. In contrast to the situation with compounds of the *N*-acetyllactosamine type (see earlier), it is impossible to deduce, for glyco-asparagines and oligosaccharides of the oligomannoside type, the type of mono-substitution from the chemical shift of this particular signal.

Compound **67** is a double-branched, reducing oligosaccharide of the oligomannoside type containing nine Man residues; it is the largest oligosaccharide thus far isolated from the urine of patients with mannosidosis.[61,96,97] The 500-MHz, ^{1}H-n.m.r. spectrum of **67** is depicted in Fig. 50, and its pertinent, n.m.r.-spectral features are included in Table XXV.

(96) K. Yamashita, Y. Tachibana, K. Mihara, S. Okada, H. Yabuuchi, and A. Kobata, *J. Biol. Chem.*, 255 (1980) 5126–5133.

(97) F. Matsuura, H. A. Nunez, G. A. Grabowski, and C. C. Sweeley, *Arch. Biochem. Biophys.*, 207 (1981) 337–352; P. F. Daniel, D. F. De Feudis, and I. T. Lott, *Eur. J. Biochem.*, 114 (1981) 235–237; H. Egge, J.-C. Michalski, and G. Strecker, *Arch. Biochem. Biophys.*, 213 (1982), 318–326.

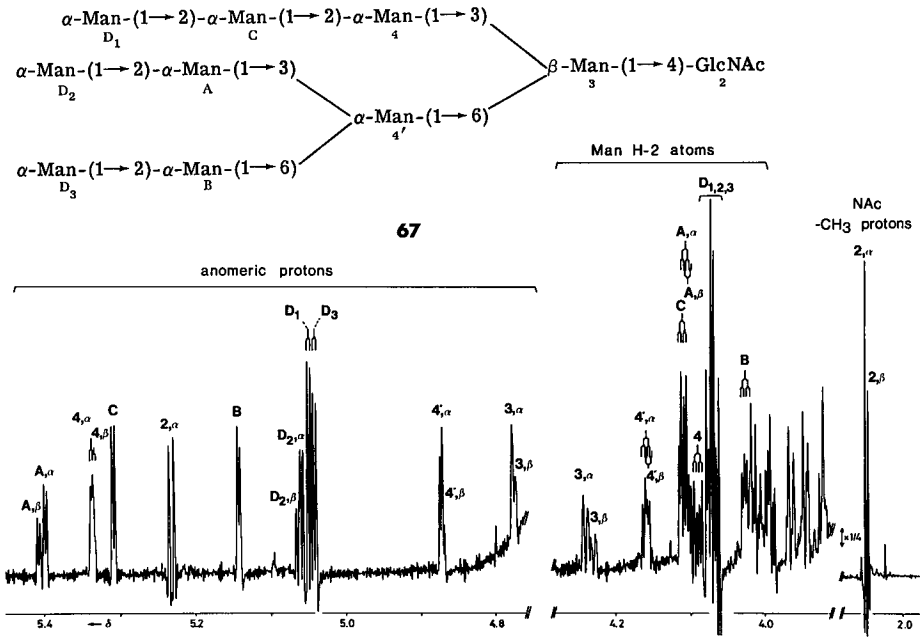


FIG. 50.—Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **67**. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structure. Signals of corresponding protons in the α and β anomer of **67**, occurring in this anomeric mixture in the ratio of 2:1, coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

The presence of the first branching-point (4–3–4′) in **67** is evident from the chemical shifts of the H-1 and, especially, of the H-2 signal of Man-3 (compare **66**). The changes in the chemical shifts of H-1 and of the *N*-acetyl protons of GlcNAc-2, as compared to **66**, reflect the extension of the second branching-point (A–4′–B) with α-(1→2)-linked Man groups, in particular with Man-D_2 (see later).

Analogous to the data for **65**, the H-1 signals at δ 5.337, 5.308, and 5.048 are ascribed to the upper-branch Man residues 4, C, and D_1, respectively; also, the H-2 resonances of these residues are found at the same positions as described for **65** (see Table XXV).

The H-1 and H-2 signals of Man-4′ are found at $\delta \sim 4.87$ and ~ 4.16, respectively; these positions are in accord with those of the corresponding protons in **61–63**, also possessing a Man-4′ residue disubstituted by A and B. Thus, for the spectral interpretation, the lower-branch part of **67** can be conceived of as an extension of **61–63** with two α-(1→2)-linked Man groups, D_2 and D_3. As is usual for terminal,

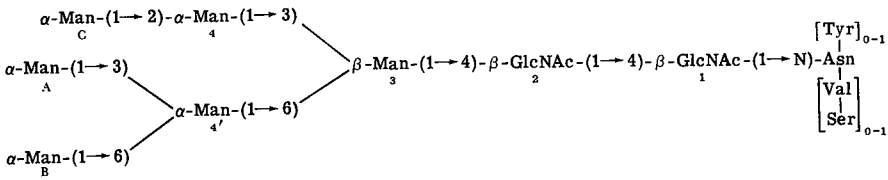
α-(1→2)-linked Man groups, the H-1 signals of Man-D_2 and -D_3 are observed at δ ~5.05, whereas their H-2 resonances are found at δ ~4.07.

Owing to the substitution of Man-B by Man-D_3 in α-(1→2) linkage, the signals of both H-1 and H-2 of the former residue undergo a downfield shift (for H-1, $\Delta\delta$ 0.234 p.p.m. and for H-2, $\Delta\delta$ 0.040 p.p.m.; compare Table XXVII). The H-1 and H-2 signals of Man-A also show downfield shifts, due to the attachment of Man-D_2. The shift increment for H-2 ($\Delta\delta$ ~0.04 p.p.m.) is identical to that just described as being typical for the introduction of an α-(1→2)-linked Man. However, the shift increment for H-1 ($\Delta\delta$ ~0.31 p.p.m.) deviates, and this may arise from a spatial effect (see later).

Doubling of the signals of the structural-reporter groups due to anomerization is clearly observable for Man-A ($|\Delta\delta_{\alpha-\beta}|$ for its H-1 signal is again larger than $|\Delta\delta_{\alpha-\beta}|$ for H-1 of Man-4'; however, the value is lessened as compared to **66** having a terminal, Man-A group), but anomerization effects upon the H-1 and H-2 signals of Man-B are apparently absent. Therefore, the set of doublets at δ 5.058 and 5.063 (relative intensity 2:1, the anomeric ratio) are attributed to H-1 of Man-D_2 in the α and β anomer of **67**, respectively. Consequently, the single H-1 doublet at δ 5.040 belongs to Man-D_3.

Comparison of the n.m.r.-spectral parameters of the three branches of **67** reveals that the anomeric configuration of GlcNAc-2 exerts its influence far more pronouncedly on the A-D_2 branch than on the other two branches, suggesting that this branch occurs in the sphere of influence of the anomeric center of GlcNAc-2, in contrast to the other two. On the basis of these n.m.r.-spectral features, it is proposed that the favored conformation of **67** in solution is as depicted in Fig. 51. This conformation might offer an explanation for the unexpectedly large, downfield shift of the H-1 signal of Man-A due to elongation of this branch with Man-D_2, and also for the observed shifts of the H-1 signals of GlcNAc-2 in the α anomers of **66** and **67** and for that of the *N*-acetyl protons of this residue in **67**, in comparison to **64** and **65**.

Compounds **63** (IgM) and **68–70** are examples of the oligomannoside type of glycopeptide that differ in the content of α-(1→2)-linked Man residues. These compounds were isolated in a mixture from glycosylation site Asn-563 of an immunoglobulin M obtained from blood



63 (IgM)

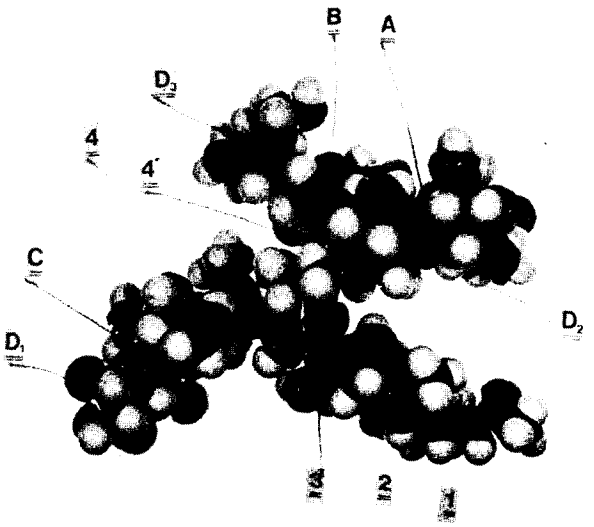
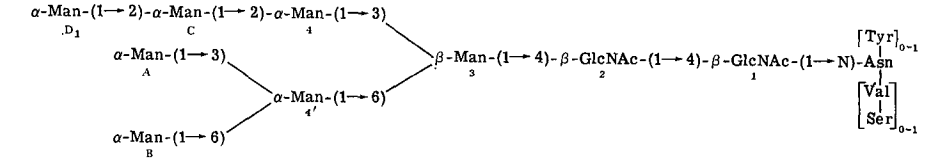
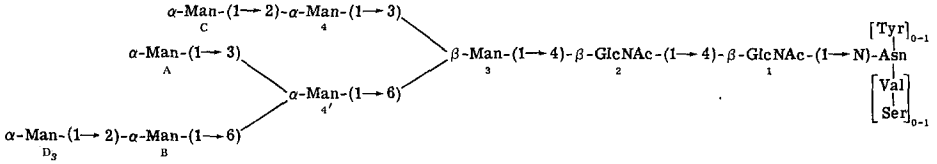


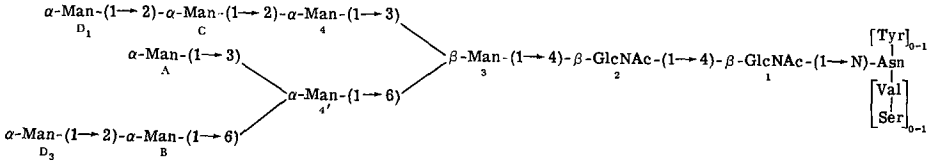
FIG. 51.—Space-filling, Molecular Model of the Oligomannoside Type of Carbohydrate Chain that Contains 9 Mannose Residues. [Numbers and letters correspond to the coding used in Figs. 50, 53, and 54 (see also, footnote on page 221). The relatively close, spatial proximity of the A–D_2 branch and the *N*,*N*′-diacetylchitobiose core-region (2–1) is clearly illustrated.]



68



69



70

plasma of a patient (Du) with Waldenström's macroglobulinemia.[98,99] The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is presented in Fig.

(98) J. Jouanneau and R. Bourrillon, *Biochem. Biophys. Res. Commun.*, 91 (1979) 1057–1061.

(99) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, J. Jouanneau, and R. Bourrillon, *Biochem. Biophys. Res. Commun.*, 99 (1981) 886–892.

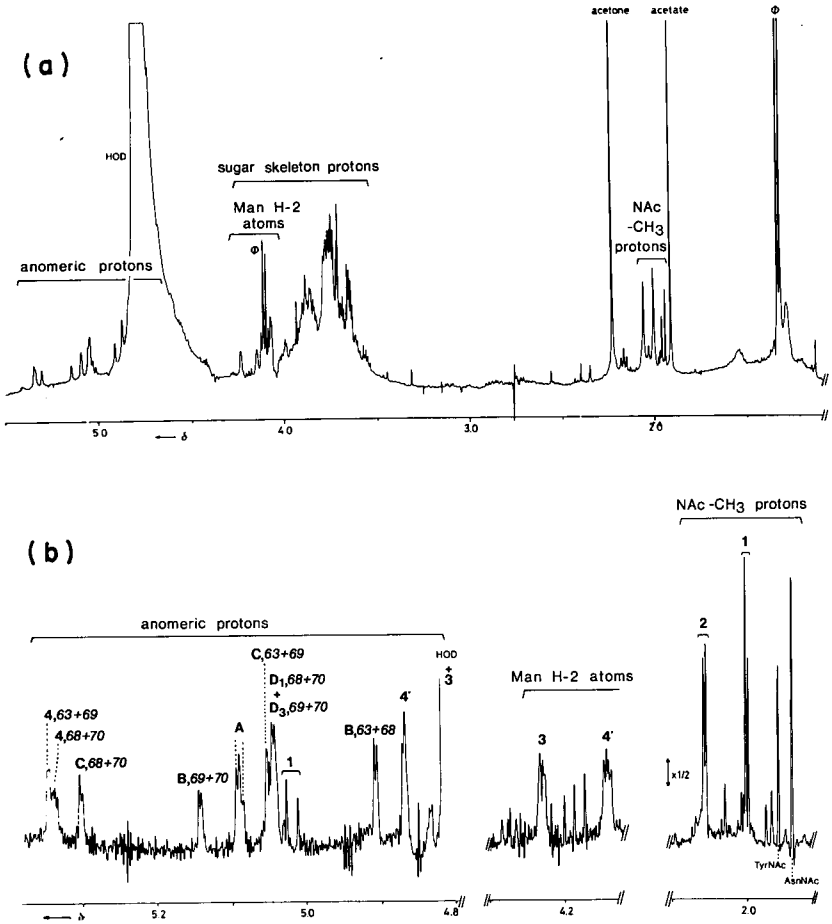
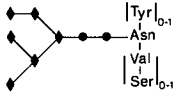
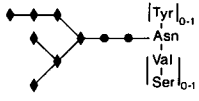
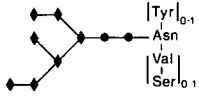
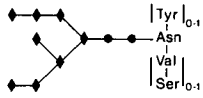
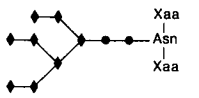
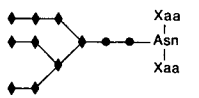


FIG. 52.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **63, 68, 69,** and **70.** [The signals marked by ϕ originate from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure.] (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of the Aforementioned Mixture. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structures, and the italic numbers to the compounds in the mixture. Signals of corresponding protons in the various components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

TABLE XXVI

^{1}H Chemical Shifts of Structural-reporter Groups of Constituent Monosaccharides for Glycopeptides of the Oligomannoside Type (Compounds 63 and 68–72)

		Compound and schematic structure						
		63 (IgM)	68	69	70	71	72 (LTF)	72 (SBA)
Reporter group	Residue							
H-1 of	1	5.02[a]	5.02[a]	5.02[a]	5.02[a]	5.092	5.092	5.092
	2	4.621	4.621	4.621	4.621	4.608	4.608	4.610
NAc of	1	2.00[a]	2.00[a]	2.00[a]	2.00[a]	2.007	2.007	2.015
	2	2.054	2.054	2.054	2.054	2.066	2.066	2.067
H-1 of	3	~4.78	~4.78	~4.78	~4.78	~4.77	~4.77	~4.77
	4	5.345	5.336	5.345	5.336	5.345	5.333	5.334
	4′	4.869	4.869	4.869	4.869	4.868	4.868	4.869

	A	5.09[a]	5.09[a]	5.09[a]	5.09[a]	5.401	5.401	5.404
	B	4.908	4.908	5.145	5.145	5.141	5.141	5.143
	C	5.052	5.304	5.052	5.304	5.059	5.308	5.308
	D_1	—	5.044	—	5.044	—	5.047	5.049
	D_2	—	—	—	—	5.059	5.059	5.061
	D_3	—	—	5.044	5.044	5.040	5.040	5.042
H-2 of	3	4.232	4.232	4.232	4.232	4.228	4.228	4.228
	4	4.10[b]	4.09[b]	4.10[b]	4.09[b]	4.10[c]	4.10[c]	4.098
	4′	4.145	4.145	4.145	4.145	4.15[c]	4.15[c]	4.156
	A	4.07[b]	4.07[b]	4.07[b]	4.07[b]	4.10[c]	4.10[c]	4.109
	B	3.98[b]	3.98[b]	4.02[b]	4.02[b]	4.02[c]	4.02[c]	4.023
	C	4.07[b]	4.10[b]	4.07[b]	4.10[b]	4.07[c]	4.10[c]	4.109
	D_1	—	4.07[b]	—	4.07[b]	—	4.07[c]	4.073[d]
	D_2	—	—	—	—	4.07[c]	4.07[c]	4.073[d]
	D_3	—	—	4.07[b]	4.07[b]	4.07[c]	4.07[c]	4.066[d]

[a] Values could not be determined more accurately (±0.01 p.p.m.), due to heterogeneity of the peptide moiety (see also, legend to Fig. 52). [b] Values could not be determined more accurately (±0.01 p.p.m.), due to the complexity of the mixture of **63**, and **68–70.** [c] Values could not be determined more accurately (±0.01 p.p.m.), due to the complexity of the pattern in the spectral region $4.00 < \delta < 4.15$ of the mixture of **71** and **72** (LTF). [d] Assignments may have to be interchanged.

52, and the relevant, n.m.r.-spectral parameters of the compounds are listed in Table XXVI.

From the spectrum of this complex mixture of glycopeptides, it may be inferred that all of the constituents have, in common, Man residues 3, 4, 4′, A, and B, and the *N*,*N*′-diacetylchitobiose unit. [The smallest component of this mixture differs from **63** (ov), obtained from ovalbumin, only in the peptide part.] The *N*-acetyl signals of GlcNAc-1 and -2 are both split into two singlets; this reflects the heterogeneity of the peptide moiety. Nevertheless, the chemical shift of the *N*-acetyl signal of GlcNAc-2 (δ ~2.055) is typical for extension of the core pentasaccharide with Man residues only [compare **61, 62, 63** (ov), and **67**; see Tables XXIV and XXV].

Besides the β-linked Man-3, clearly characterized by its H-2 signal at δ 4.232, only α-linked Man residues occur in the peripheral part of the glycan chain.This is derived from the chemical shifts of their H-1 signals, in combination with their $J_{1,2}$ values.[66] The set of chemical shifts of H-1 (δ 4.869) and of H-2 (δ 4.145) of Man-4′ is indicative of a disubstitution of this residue at O-3 and O-6 by Man-A and -B, respectively [compare **61, 62, 63** (ov), and **67**, Tables XXIV and XXV]. The H-1 atom of Man-A gives rise to two doublets, at δ 5.093 and 5.088, in the ratio of 2:1, resonating in an area that is characteristic for a terminal, nonreducing position of Man-A [compare **61, 62, 63** (ov), **66**, and Table XXVII]. The doubling of the signal can be explained in terms of heterogeneity of the sample. It is tempting to correlate the intensity ratio of the two signals of the anomeric proton of Man-A with the molar ratio of the *N*-terminal amino acids (see Fig. 52). This suggestion would fit the proposal for the spatial conformation of carbohydrate chains of the oligomannoside type (compounds **66** and **67**; see Fig. 51).

The H-1 atom of Man-B also gives rise to two doublets, at δ 4.908 (terminal B) and at δ 5.145 (O-2-substituted B), in the ratio of 2:1. However, from these values, it may be concluded that 33% of the Man-B residues in the glycopeptide sample bear an α-(1→2)-linked Man-D_3 (δH-1 5.044), whereas, in the remaining part, Man-B occupies a terminal position in the chain.

For H-1 of Man-4 in the upper branch, two doublets are observed, at δ 5.345 and 5.336, in the ratio of 1:1. The approximate value of the chemical shift (5.34 p.p.m.) indicates that, in all components of the mixture, Man-4 bears the α-(1→2)-linked Man-C [see **63** (ov), and **64–67**]. Also, H-1 of Man-C gives rise to two doublets, at δ 5.304 (O-2-substituted C) and 5.052 (terminal C). These features can be explained by the presence of Man-D_1 in 50% of the structures in the glycopeptide mixture. The presence of Man-D_1 (δH-1 5.044) gives rise to a downfield shift of the H-1 signal of Man-C (Δδ 0.252 p.p.m.) and to a slight,

upfield shift for H-1 of Man-4 ($\Delta\delta \sim -0.01$ p.p.m.), as compared to the 4-C branch without D_1 (compare the step from **64** or **66** to **65** or **67**; see Table XXV).

The high-resolution, ^{1}H-n.m.r. spectroscopy conducted on this glycopeptide mixture shows that compounds of the oligomannoside type that differ in the number and the position of α-(1→2)-linked Man residues can be identified, even if heterogeneity is present in the peptide part.

Compounds **71** and **72** (LTF) are glycopeptides of the oligomannoside type containing eight and nine Man residues, respectively, that were obtained from bovine lactotransferrin[100] in a mixture of them in the ratio of 2:3. The 500-MHz, ^{1}H-n.m.r. spectrum of this mixture is depicted in Fig. 53, and the pertinent, n.m.r.-spectral parameters of **71** and **72** (LTF) are included in Table XXVI.

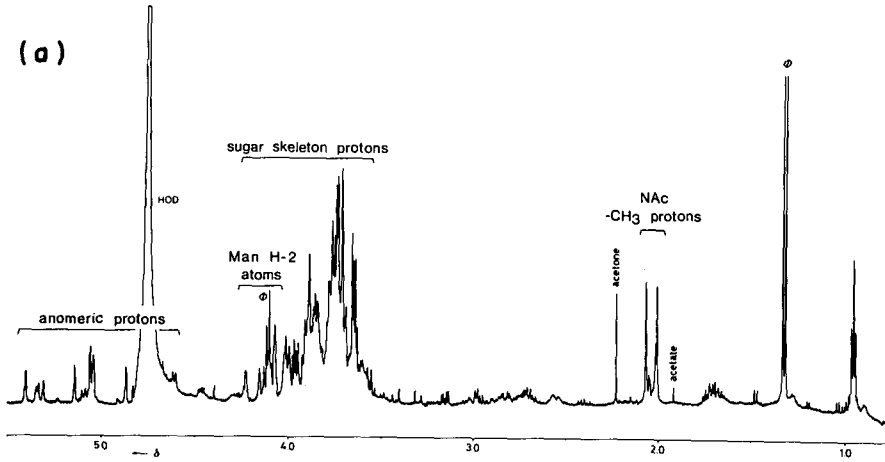
For the spectral interpretation, **71** and **72** can be conceived of as extensions of **63** with two, or three, terminal α-(1→2)-linked Man groups, respectively. The spectral features of the core residues, GlcNAc-1 and -2, and Man-3, -4, and -4′, in **71** and **72** are essentially identical with those of the corresponding residues in compounds **63** and **68–70** (see Table XXVI). Heterogeneity of the peptide moiety produces multiple resonances for the *N*-acetyl protons, and also relatively broad-lined, anomeric doublets of GlcNAc-1 and of GlcNAc-2.

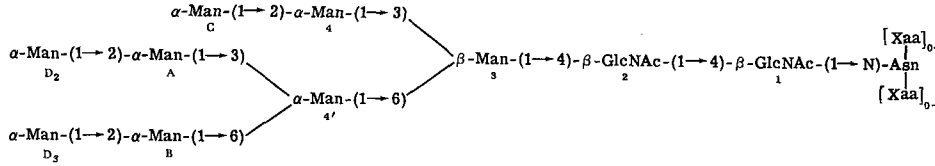
The anomeric signal of Man-A, -B, and -4′ at δ 5.401, 5.141, and 4.868, respectively, each appears as a single doublet in the spectrum of the mixture (relative intensities 1:1:1). These values of chemical shift are in good agreement with those of the corresponding residues in the β anomer of oligosaccharide **67**, indicating that the lower branches of **71** and **72** are terminated with Man-D_2 and -D_3.

However, for H-1 of Man-4, two doublets are observed, at δ 5.333 and 5.345, in the intensity ratio of 3:2. The sum of these intensities equals that of the H-1 signal of Man-A. This feature indicates the presence of a terminal Man-C, in α-(1→2) linkage to Man-4, in the minor component (**71**) of the mixture, whereas, in the major component (**72**), Man-C is substituted at O-2 by another Man residue (D_1). This effect of the attachment of Man-D_1 has been described in the step going from oligosaccharide **64** to **65** (see Table XXV). In accordance with this interpretation, the H-1 signal of Man-C in **72** at δ 5.308 has an intensity equal to that of Man-4 at δ 5.333. The H-1 signal of the terminal, α-(1→2)-linked Man-C in **71** is found at $\delta \sim 5.05$, as would be expected from the data for **63**, **64**, **66**, and **69**.

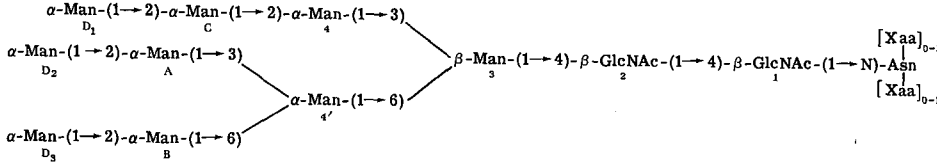
Obviously, the presence of Man-D_1 in one of the two compounds

(100) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, G. Spik, A. Chéron, and J. Montreuil, *Biochim. Biophys. Acta*, 675 (1981) 293–296.

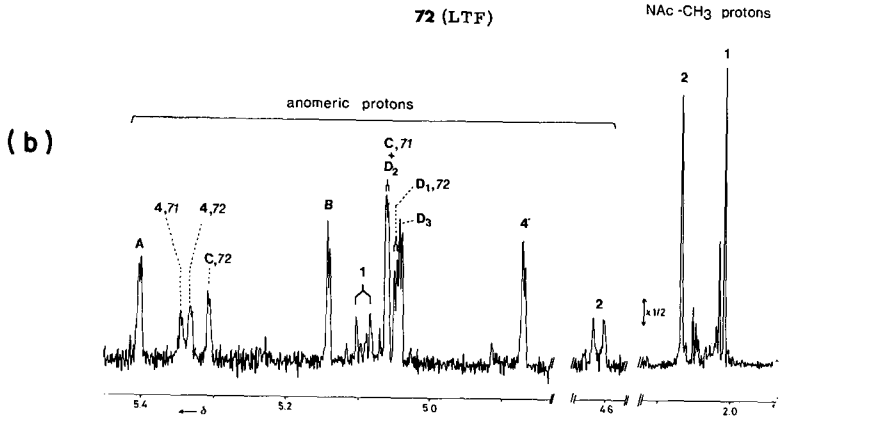

(a)
HOD
sugar skeleton protons
Man H-2 atoms
anomeric protons
NAc -CH3 protons
acetone
acetate
Φ
50
δ
40
30
20
10


α-Man-(1→2)-α-Man-(1→3)
C 4
α-Man-(1→2)-α-Man-(1→3)
D2 A
α-Man-(1→6)
4'
α-Man-(1→2)-α-Man-(1→6)
D3 B
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
3 2 1
[Xaa]0-1
[Xaa]0-1

71


α-Man-(1→2)-α-Man-(1→2)-α-Man-(1→3)
D1 C 4
α-Man-(1→2)-α-Man-(1→3)
D2 A
α-Man-(1→6)
4'
α-Man-(1→2)-α-Man-(1→6)
D3 B
β-Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc-(1→N)-Asn
3 2 1
[Xaa]0-1
[Xaa]0-1

72 (LTF)


NAc -CH3 protons
(b)
anomeric protons
A
4,71
4,72
C,72
B
C,71
D2
D1,72
D3
1
4'
2
2
1
x 1/2
5.4
δ
5.2
5.0
4.6
2.0

present in the mixture is reflected in the intensity ratio of the anomeric signals of terminal α-(1→2)-linked Man residues (δ ~5.05). Three signals are observed, at δ 5.059, 5.047, and 5.040 in the ratios of 7:3:5. Based on its relatively low intensity, the signal at δ 5.047 is ascribed to Man-D_1 in **72**. The signal at δ 5.059 partly belongs to Man-D_2 in both **71** and **72**, as may be deduced from comparison with **67** (see Table XXV). The anomeric signal of the terminal Man-C in **71** (compare **64** and **66**) also contributes to the intensity of the signal at δ 5.059, thereby making this the most intense signal in this area. Finally, the doublet at δ 5.040 is attributed to H-1 of Man-D_3.

Compound **72** (SBA) is a glycopeptide of the oligomannoside type containing nine Man residues that was obtained from soybean agglutinin.[101,102] The 500-MHz, ^{1}H-n.m.r. spectrum of **72** (SBA) is depicted in Fig. 54, and its n.m.r.-spectral parameters are listed in Table XXVI.

The chemical shifts, coupling constants, and line widths of all structural-reporter-group signals of **72** (SBA) are identical to those of the corresponding signals of **72** (LTF), except for the *N*-acetyl signals of GlcNAc-1 and -2. The latter signals show a multiplicity that is due to heterogeneity of the peptide moiety.[67] The carbohydrate part is homogeneous; its spectral data are in excellent agreement with those for the β anomer of oligosaccharide **67** (see Table XXV).

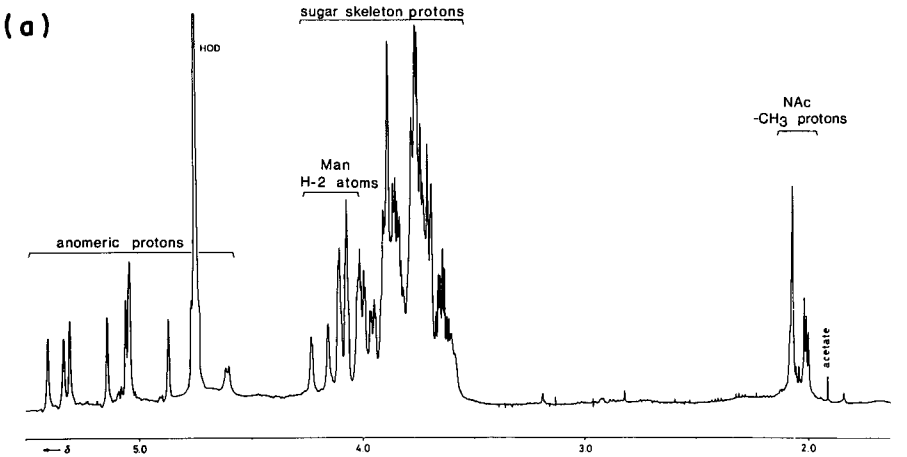
3. Additivity Rules

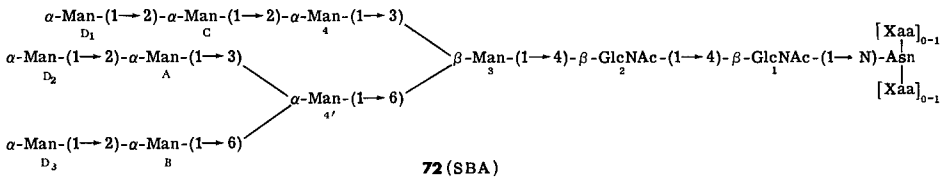
In Sections II, 1 and 2, have been described the ^{1}H-n.m.r.-spectral parameters, namely, the chemical shifts, coupling constants, and line widths of the structural-reporter groups of the monosaccharide constituents, of a wide diversity of *N*-glycosylic carbohydrate chains of glycoproteins.

The coupling constant of the anomeric proton of a given sugar residue ($J_{1,2}$) provides essential information on the configuration of its

(101) H. Lis and N. Sharon, *J. Biol. Chem.*, 253 (1978) 3468–3476.
(102) L. Dorland, H. van Halbeek, J. F. G. Vliegenthart, H. Lis, and N. Sharon, *J. Biol. Chem.*, 256 (1981) 7708–7711.

FIG. 53.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of a Mixture Containing Compounds **71** and **72** in the Ratio of 2:3, Obtained from Bovine Lactotransferrin. [The signals marked by φ originate from a frequently occurring, nonprotein, noncarbohydrate contaminant of unknown structure.] (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of the Aforementioned Mixture. [The bold numbers and letters in the spectrum refer to the corresponding residues in the structures. Signals of corresponding protons in the two components of this mixture coincide, unless otherwise indicated. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]





72 (SBA)

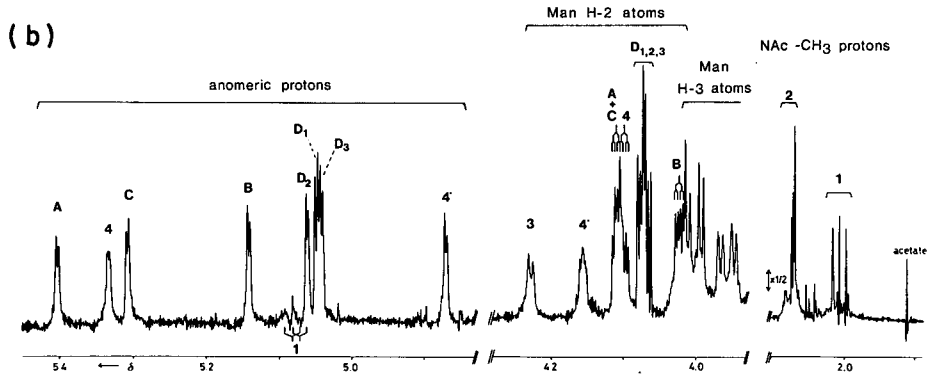


FIG. 54.—(a) Overall, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **72**, Obtained from Soybean Agglutinin. (b) Expanded, Structural-reporter-group Regions of the Resolution-enhanced, 500-MHz, ^{1}H-N.m.r. Spectrum of Compound **72**. [The bold numbers and letters refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl-proton region differs from that of the other parts of the spectrum, as indicated.]

glycosylic linkage. With H-2 in axial position, a $J_{1,2}$ value of 2–4 Hz indicates an α, and that of 7–9 Hz, a β, anomer; this applies to such monosaccharides as Gal, GlcNAc, and L-Fuc. For Man, having its H-2 atom in equatorial position, the difference is more subtle: $J_{1,2}$ 1.6 Hz

for an α, and 0.8 Hz for a β, linkage. In principle, the complete set of vicinal coupling-constants of sugar-ring protons allows the deduction of the conformation of the residue in question, as was shown for compounds **1**, **2**, and **55**.

The line widths of the structural-reporter-group signals bear information on the "flexibility" of the protons involved and, by extrapolation, on the mobility of individual sugar residues and of branches in the chain. Furthermore, coupling constants, in combination with line widths, give the structural-reporter-group signals their characteristic shapes, which make them recognizable: These parameters facilitate the assignments of resonances belonging to structural-reporter groups.

However, it is the chemical shift that makes a certain proton a structural-reporter group. The large amount of information hidden in a chemical shift is readily accessible for purposes of structure determination if a signal occurs outside the bulk of overlapping resonances and thus can be found by inspection. Among the structural-reporter groups, the chemical shift of an anomeric proton is primarily sensitive to the type of sugar, the anomeric configuration of its linkage, the site of glycosidic linkage, the sequence of the component sugars, and the position of the residue in the carbohydrate chain. However, steric factors are also important determinants of the chemical shift of an anomeric proton. The chemical shifts of other structural-reporter groups also show a dependence on many primary, structural parameters. In theory, it is possible to deduce the complete, primary structure of carbohydrate chains from the n.m.r.-spectral parameters of the structural-reporter groups. The availability of consistent series of compounds is helpful in interpretation of the spectra. This is well illustrated by the series of structurally related *N*-glycosylic carbohydrate chains of glycoproteins, which have been arranged herein by increasing complexity for those of the *N*-acetyllactosamine and the oligomannoside type.

Based on the sets of data listed in Tables I–V, VII–XII, XV, XVI, XVIII–XX, and XII–XXVI, it is possible to establish empirical rules to describe the influence of structural changes on the chemical shifts of the structural-reporter groups involved. These rules permit reliable prediction of the ^{1}H-n.m.r. spectrum for every partial structure in the most complex one thus far observed. Less certain may be predictions of the spectra of structures built up from the same units as those occurring in the set of reference compounds, but constituting a novel family of structures.

Comparison of the available data shows that additivity rules can readily be deduced for the chemical shifts of structural-reporter groups in those cases in which the mannotriose core is complete in the starting compound, as well as in the elongated chain. For compounds

possessing an incomplete, mannotriose branching-unit, the chemical shifts of structural-reporter groups can less readily be rationalized; the influences, on the chemical shifts of neighboring residues, due to the attachment of Man-4 (for example, **17**→**19**) or Man-4′ (**18**→**20**) are partly attributable to the drastic, steric changes accompanying these extensions. However, the type of mono-substitution of Man-3 by (a substituted) Man-4, or by (a substituted) Man-4′, is obviously characterized by the presence of either the structural-reporter-group signals of Man-4 (δH-1 ~5.12; δH-2 ~4.19 for mono-antennary structures) or of Man-4′ (δH-1 ~4.92; δH-2 ~4.12). The set of chemical shifts of the Man-3 H-2 signal (δ ~4.24 and 4.08, respectively), and of the GlcNAc-2 H-1 (for example, for the α anomers of reducing oligosaccharides, δ 5.207 and 5.216, respectively) and *N*-acetyl protons (δ 2.042 and 2.062, respectively) gives independent information in this respect. Recognition of the type of branching in the *N*-acetyllactosamine type of structures possessing an intact mannotriose core unit substituted by one or more GlcNAc residues is possible on the basis of the set of chemical shifts of the H-1 and H-2 signals of Man-3, -4, and -4′, as summarized in Table VI.

The chemical shift of the H-1 signal of a terminal, β-linked GlcNAc residue is dependent on the type of its linkage to Man. For the β-(1→2)-linked GlcNAc residues 5 and 5′ in diantennary structures, δH-1 ~4.56 is found. For the β-(1→4)-linked GlcNAc-7, δH-1 4.52 is observed (for example, for **16**), whereas, for the β-(1→6)-linked GlcNAc-7′, δH-1 4.53 can be expected (see later). For the chemical shift of the H-1 signal of the intersecting GlcNAc-9 residue, β-(1→4)-linked to Man-3, in *N*-acetyllactosamine-type structures possessing at least GlcNAc-5 and -5′, the value δ 4.47 is observed. The relatively high-field, resonance position of this proton is probably due to the special, steric requirements for the attachment of GlcNAc-9, which also cause a disturbance of the total pattern of the Man H-1 and H-2 signals (compare Table VII with VI), rather than to the β-type of linkage of the substituted Man-3. This is a very illustrative example of changes in chemical shifts that are attributable to conformational differences between starting compound and extended chain. As such conformational changes are hardly predictable, it is evident that the prediction of chemical shifts of structural-reporter groups in such cases will also be difficult. This problem is nicely illustrated by the impossibility of predicting the chemical shifts of some structural-reporter groups in hybrid structures containing GlcNAc-9, starting from data for compounds of the *N*-acetyllactosamine type.[50] The *N*-acetyl-proton resonances of GlcNAc residues in terminal position are observed at δ 2.053 (GlcNAc-5 in diantenna), 2.050 (GlcNAc-5′ in diantenna), 2.078

(GlcNAc-7), and 2.066 (GlcNAc-9), whereas δ 2.041 can be expected for GlcNAc-7′ (see later).

Attachment of Gal in β-(1→4) linkage to GlcNAc, completing an *N*-acetyllactosamine unit, causes a downfield shift for the signal of GlcNAc H-1 (Δδ 0.025 p.p.m.) and a small, upfield shift for the GlcNAc *N*-acetyl protons (Δδ ~ −0.003 p.p.m.), regardless of the type of linkage of the latter residue. The resonance position of H-1 of such a terminal Gal residue primarily reflects the type of linkage of the *N*-acetyllactosamine unit, containing Gal, to Man-4 or -4′ (δH-1 4.47 for Gal-6 and -6′, δH-1 4.46 for Gal-8, and δH-1 4.48 for Gal-8′). The small difference in chemical shifts (Δδ ~0.005 p.p.m.) between the H-1 signals of Gal-6 and -6′ can be ascribed to the different type of linkage between Man-4 and Man-3, in comparison to that between Man-4′ and Man-3.

The influences of extension of structures of the *N*-acetyllactosamine type with terminating NeuAc, in α-(2→6) or α-(2→3) linkage to a Gal residue, on the chemical shifts of structural-reporter groups of neighboring residues are compiled in Tables XIII and XIV. These Tables also contain the characteristics of the structural-reporter groups (H-3*a*, H-3*e*, and *N*-acetyl protons) of NeuAc itself. Similar data have been listed for the effects of attachment of Fuc, as well as for the structural-reporter groups (H-1, H-5, and -CH_3) of Fuc in α-(1→6) linkage to GlcNAc-1 (in Table XVII), for Fuc in α-(1→3) linkage to a peripheral GlcNAc residue (in Table XIX), and for Fuc in α-(1→2) linkage to Gal (in Table XXI).

Concerning structures of the oligomannoside type, the influences of substitution of the Man-4 and Man-4′ residues of the core pentasaccharide by Man-A, -B, and -C, and, in turn, those of substitution of the latter by terminating Man-D residues, have been summarized in Table XXVII. Also, the characteristics of the H-1 and H-2 signals of terminal α-(1→2)-, α-(1→3)-, or α-(1→6)-linked Man residues have been given.

In the foregoing, it has been emphasized that chemical shifts and chemical-shift differences are the most appropriate parameters for correlating the primary, structural characteristics with the set of n.m.r.-spectral data. The fact that the chemical shifts of structural-reporter groups are so sensitive to changes in the sequence of the oligosaccharide chain opens the possibility of discriminating between compounds that are very closely related structurally. Under favorable conditions, all components of mixtures of isomers can even be characterized (see, for example, **49–51**). However, the chemical shifts of structural-reporter groups are also sensitive to such conformational effects as through-space interactions, or changes in linkage orienta-

TABLE XXVII

Recognition of α-(1→2)-, α-(1→3)-, or α-(1→6)-Linked, Terminal or Nonterminal Man in the Sequence αMan(1→*x*)*y*Man(1→*z*)Gly

^{1}H Chemical Shifts of Structural-reporter Groups of α-Linked, Terminal Man[a]

Residue	Reporter group	δ ±s.d.
Man-D_1	H-1 (x = 2; y = α; z = 2)	5.047 ±0.002
Man-C or -D_2	H-1 (x = 2; y = α; z = 3)	5.057 ±0.003
Man-D_3	H-1 (x = 2; y = α; z = 6)	5.042 ±0.002
Man-4	H-1 (x = 3; y = β; z = 4)	5.105 ±0.005
Man-A	H-1 (x = 3; y = α; z = 6)	5.09 ±0.02[b]
Man-4′	H-1 (x = 6; y = β; z = 4)	4.919 ±0.003
Man-B	H-1 (x = 6; y = α; z = 6)	4.908 ±0.001
Man-C, -D_1, -D_2, or -D_3	H-2 (x = 2; y = α; z = 2, 3, or 6)	4.068 ±0.005
Man-4	H-2 (x = 3; y = β; z = 4)	4.073 ±0.003
Man-A	H-2 (x = 3; y = α; z = 6)	4.065 ±0.003
Man-4′	H-2 (x = 6; y = β; z = 4)	3.97 ±0.01
Man-B	H-2 (x = 6; y = α; z = 6)	3.99 ±0.01

Principal Influences of α-Linked Man on the Chemical Shifts of Structural-reporter Groups of Neighboring Residues[c]

Neighboring residue	Site of substitution	Reporter group	Δδ (p.p.m.) ±s.d. (p.p.m.)
*y*Man(1→*z*)	x = 2	H-1 (y = α; z = 2, 3, or 6)	+0.244 ±0.007[d]
		H-2 (y = α; z = 2, 3, or 6)	+0.037 ±0.007
Gly = αMan(1→3)	x = 2	H-1 (y = α; z = 2)	−0.012 ±0.003
		H-2 (y = α; z = 2)	−0.02 ±0.01
*y*Man(1→*z*)	x = 3	H-1 (y = β; z = 4)	+0.016 ±0.005
		H-1 (y = α; z = 6)	−0.017 ±0.003
		H-2 (y = β; z = 4)	+0.175 ±0.005
		H-2 (y = α; z = 6)	+0.160 ±0.003
*y*Man(1→*z*)	x = 6	H-1 (y = β; z = 4)	~0.0
		H-1 (y = α; z = 6)	−0.027 ±0.004[e]
		H-2 (y = β; z = 4)	+0.021 ±0.005
		H-2 (y = α; z = 6)	+0.015 ±0.004[e]

[a] Mean values ±s.d. at T = 300 K; see Tables II, III, VIII, IX, XXIV, XXV, and XXVI. [b] δ Value for H-1 of Man-A is also influenced by spatial structure (see **61**, **62**, and **66**). [c] Values of Δδ are mean values ±s.d. at T = 300 K, calculated for extensions with differently linked Man residues (see Tables II, III, VIII, IX, XXIV, XXV, and XXVI). [d] Except for Man-A; due to steric effects, ΔδH-1 ~0.31 p.p.m. for this residue (see compound **67**). [e] Values of Δδ could not be established separately; only their combination with the influences of attachment of αMan(1→3) could be deduced, for example, from comparison of **3** with **61**. However, see the description of compound **66**.

tions due to elongation of the carbohydrate chain (see δH-1 of Man-A in **67**). Obviously, these effects are not comprised within the additivity rules. Except for this limitation, the additivity rules, correlating the chemical shifts of structural-reporter groups with primary, structural changes, are very useful for the deduction of ^{1}H-n.m.r. spectra of (partial) structures belonging to the class of compounds already discussed.

III. Concluding Remarks

This article describes the application of high-resolution, ^{1}H-n.m.r. spectroscopy for the characterization of the primary structure of 72 carbohydrate chains derived from glycoproteins. These structures constitute sets of *N*-glycosylic compounds of the *N*-acetyllactosamine type and of the oligomannoside type. Most of the ^{1}H-n.m.r. spectra of these compounds were recorded at 500 MHz. The chemical shifts, coupling constants, and line widths of their structural-reporter-group signals were established as far as possible, and the following remarks are relevant.

(*i*) The ^{1}H-n.m.r. spectrum is, as a whole, typical for the compound investigated. Provided that it is recorded at a relatively strong magnetic field, such a spectrum can serve as an identity card, which, even without interpretation, is at least suitable for comparison with spectra of other compounds, allowing a conclusion as to whether the compounds are identical or not. The visual inspection of spectra, besides comparing data listed in tables, is extremely helpful for this purpose, as the foregoing atlas of spectra may illustrate.

(*ii*) Integration of structural-reporter-group signals permits estimation of the purity of the sample. In this way, it is often easy to decide whether a sample consists of a single compound or a mixture of several components. This is particularly interesting for the detection of natural and artificial (micro)heterogeneity.

(*iii*) The n.m.r.-spectral parameters of the structural-reporter groups bear the essential information necessary for the elucidation of the primary structure of the carbohydrate chain. If consistent series of structurally related compounds are available for both types of *N*-glycosylic, carbohydrate structure, a number of empirical, additivity rules for the chemical shifts of the structural-reporter groups can be deduced, as outlined in Section II,3.

(*iv*) Differences in chemical shifts, and also in line widths, of structural-reporter-group signals allow, in several instances, qualitative deduction of spatial structures, in particular regarding the orientation of different branches with respect to each other (for example, the influences of the configuration of the anomeric center of GlcNAc-2 on the

chemical shifts of structural-reporter groups of reducing oligosaccharides). To obtain more details about the structure of the carbohydrate chains in solution, a theoretical approach, such as that given by Lemieux and coworkers,[83,103] is necessary. There is no doubt that, if, by means of more-advanced, n.m.r.-spectral techniques, more details about resonances located in the bulk of the spectrum become available, these data will be useful in elucidating secondary structures.

(*v*) Performance of the ^{1}H-n.m.r.-spectral measurements at 500 MHz has the advantage of enabling a more precise determination to be made of the chemical shifts of the structural-reporter groups, and of affording more details of the splitting patterns of their signals. Also, the sensitivity of an n.m.r. apparatus operating at such a strong magnetic field is relatively high, making possible the analysis of 25 nmoles of glycopeptide or oligosaccharide sample.[99]

(*vi*) The high-resolution, 360-MHz and 500-MHz, ^{1}H-n.m.r.-spectral data described herein can be used for the interpretation of spectra obtained with n.m.r. apparatus operating at lower magnetic-field strengths (as has been illustrated in, for example, Refs. 50 and 104–110). For the analysis of new types of compound, it is advisable to record spectra at the highest magnetic fields available; this is particularly relevant for *O*-glycosylic carbohydrate structures of glycoproteins, where only a beginning has as yet been made in the characterization of corresponding oligosaccharide-alditols by means of high-resolution, ^{1}H-n.m.r. spectroscopy.[12,13,32,111–115]

(103) R. U. Lemieux, *Chem. Soc. Rev.*, 7 (1978) 423–452.
(104) J. P. Carver and A. A. Grey, *Biochemistry*, 20 (1981) 6607–6616.
(105) J. Hakimi, J. P. Carver, and P. H. Atkinson, *Biochemistry*, 20 (1981) 7314–7319.
(105a) R. R. Townsend, E. Hilliker, Y.-T. Li, R. A. Laine, W. R. Bell, and Y. C. Lee, *J. Biol. Chem.*, 257 (1982) 9704–9710.
(106) M. Kuriyama, T. Ariga, S. Ando, M. Suzuki, T. Yamada, and T. Miyatake, *J. Biol. Chem.*, 256 (1981) 12,316–12,321.
(107) J. Arnarp, M. Haraldsson, and J. Lönngren, *Carbohydr. Res.*, 97 (1981) 307–313.
(108) T. Ogawa and K. Sasajima, *Carbohydr. Res.*, 97 (1981) 205–227.
(109) T. G. Warner and J. S. O'Brien, in T. Yamakawa, T. Osawa, and S. Handa (Eds.), *Glycoconjugates, Proc. Int. Symp. Glycoconjugates, 6th*, Japan Scientific Societies, Tokyo, 1981, pp. 10–11.
(110) Y. Inoue and H. Nomoto, in Ref. 109, pp. 505–506.
(111) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, A.-M. Fiat, and P. Jollès, *Biochim. Biophys. Acta*, 623 (1980) 295–300.
(112) H. van Halbeek, L. Dorland, J. Haverkamp, G. A. Veldink, J. F. G. Vliegenthart, B. Fournet, G. Ricart, J. Montreuil, W. D. Gathmann, and D. Aminoff, *Eur. J. Biochem.*, 118 (1981) 487–495.
(113) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, A.-M. Fiat, and P. Jollès, *FEBS Lett.*, 133 (1981) 45–50.

IV. Experimental

Samples of carbohydrate structures dealt with herein stemmed from a wide variety of glycoproteins, and were isolated by a number of research groups (see Acknowledgments). The detailed sources of the compounds generously supplied for our ^{1}H-n.m.r.-spectral investigations are indicated in the text (see also, footnote 53).

The pD of solutions of glycopeptides and oligosaccharides in D_2O was adjusted to ~7, if necessary. Deuterium-exchanged samples were prepared by five dissolutions in D_2O, and lyophilizations of the solution, finally using 99.96 atom% deuterated water (Aldrich, Milwaukee, WI, USA). For n.m.r.-spectral analysis, 0.1 to 3.0 m*M* solutions of the compounds in 0.4 mL of D_2O were generally used.

The 500-MHz, ^{1}H-n.m.r. spectra were recorded with a Bruker WM-500 spectrometer operating in the pulsed, Fourier-transform mode and equipped with a Bruker Aspect 2000 computer having an 80k memory-capacity. The D resonance of D_2O was used as the field-frequency lock-signal. The spectra were obtained by using a 90° pulse-width, and accumulated into 16k addresses with an acquisition time of 3.28 s and a spectral width of 2.5 kHz. Resolution enhancement was achieved by Lorentzian to Gaussian transformation from quadrature-phase detection, followed by employment of a 32k-point, complex, Fourier transformation. In general, a few hundred acquisitions were accumulated for each sample. The indicated probe temperature was 300 K, kept constant within ±0.1 K. At this temperature, the HOD resonance is at δ ~4.75; the value depends slightly on the concentration of the sample.

The 360-MHz, ^{1}H-n.m.r. spectra were recorded with a Bruker HX-360 spectrometer operating in the pulsed, Fourier-transform mode and equipped with a Bruker B-NC 12 computer having a 16k memory-capacity. The D resonance of D_2O was used as the field-frequency lock-signal. The spectra were obtained by using a 90° pulse-width, and accumulated into 16k addresses with an acquisition time of 3.28 s and a spectral width of 2.5 kHz. Resolution enhancement was achieved by Lorentzian to Gaussian transformation from quadrature-phase detection. In general, a few hundred acquisitions were accumulated for each sample. Spectra were recorded at probe temperatures of 295 to 300 K.

---

(114) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, W. E. Hull, G. Lamblin, M. Lhermitte, A. Boersma, and P. Roussel, *Eur. J. Biochem.*, 127 (1982) 7–20.

(115) H. van Halbeek, L. Dorland, J. F. G. Vliegenthart, N. K. Kochetkov, N. P. Arbatsky, and V. A. Derevitskaya, *Eur. J. Biochem.*, 127 (1982) 21–29.

The chemical shifts (δ) are expressed in p.p.m. downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS), but were actually measured by reference to internal acetone (δ 2.225) with an accuracy of 0.001 p.p.m. at 500 MHz, or of 0.003 p.p.m. at 360 MHz.

ACKNOWLEDGMENTS

The authors are much indebted to Profs. J. Montreuil, G. Strecker, B. Fournet, G. Spik (Lille, France), K. Schmid (Boston, MA, U. S. A.), N. Sharon (Rehovoth, Israel), A. Kobata (Kobe, Japan), R. Bourrillon (Paris, France), D. H. van den Eijnden (Amsterdam, The Netherlands), and their coworkers for fruitful collaboration. The authors express their gratitude to Bruker Analytische Messtechnik GmbH, Rheinstetten, FRG, and Netherlands Bruker Spectrospin NV, Wormer, The Netherlands, for making the 500-MHz facility in Rheinstetten, FRG, accessible for their studies, and, in particular, to Dr. W. E. Hull for recording most of the 500-MHz spectra depicted herein. The remaining 500-MHz spectra and the 360-MHz spectra described in this article were recorded at the Dutch SON-n.m.r. facilities in Nijmegen and Groningen, The Netherlands, respectively. The authors also acknowledge the contribution of Dr. C. A. G. Haasnoot, Prof. R. Kaptein, and their coworkers for keeping the instruments employed in excellent condition. The authors thank Dr. G. A. Veldink for his continuous interest and for his valuable discussions, Miss J. H. van Hout and Miss C. L. E. de Lint for their expert secretarial assistance, and Mr. J. L. den Boesterd for preparing the photographs.

This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), and by the Netherlands Foundation for Cancer Research (KWF), grant UUKC-OC 79-13.