Pure & Appl. Chem., Vol.53, pp.45-77. 0033-4545/81/0101-0045$02.00/0
Pergamon Press Ltd. 1981. Printed in Great Britain.
©IUPAC

THE APPLICABILITY OF 500-MHz HIGH-RESOLUTION ^{1}H-NMR SPECTROSCOPY FOR THE STRUCTURE DETERMINATION OF CARBOHYDRATES DERIVED FROM GLYCOPROTEINS

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

Department of Bio-Organic Chemistry, State University of Utrecht,
Croesestraat 79, 3522 AD Utrecht, The Netherlands

Abstract - This report describes the application of 500-MHz ^{1}H-NMR spectroscopy for the determination of the primary structures of the carbohydrate chains of glycoproteins. For this purpose partial structures, *i.e.*, glycopeptides, oligosaccharides or oligosaccharide-alditols, of the biopolymers were investigated. The spectra, recorded in deuteriumoxide at ambient temperature, contain easily accessible information suitable for the elucidation of primary structures of virtually pure compounds as well as for the analysis of mixtures of closely related components. The key to this information is found in the resonances of the so-called structural reporter groups which are the signals of individually resonating protons. The chemical shifts and coupling constants of the structural reporter groups and the line widths of their resonances are the characteristic parameters which reflect the sugar composition and substitution pattern of the constituting monosaccharides.
The suitability of this approach is illustrated for a wide variety of carbohydrate chains. With regard to the *N*-glycosidically linked type of chains the asialo afuco *N*-acetyllactosaminic mono-, bi-, tri- and tetra-antennary structures were studied, as well as several of their analogues terminated with different types of sialic acid or fucose residues. The structure of oligomannoside type asparagine-bound carbohydrate chains can now completely be defined in terms of NMR parameters of their structural reporter groups. Furthermore, the characterization of several representatives of the *O*-glycosidically linked mucin type of carbohydrate chains could be achieved. Finally it should be stressed that a high-resolution ^{1}H-NMR spectrum is highly valuable even in cases that no complete assignment of the primary structure is possible. The spectrum can *i.a.* be used as an identity card to compare carbohydrate chains obtained from different sources. ^{1}H-NMR spectroscopy appears to be significantly more powerful at 500 MHz than at 360 MHz. By this non-destructive method samples as small as 25 nmoles could be adequately analyzed.

INTRODUCTION

During the last decade the interest in the structure and function of glycoproteins increased greatly. It could be shown that the carbohydrate chains of these biopolymers 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 have to be mentioned.
Significant progress has been made in the characterization of the primary structures of the carbohydrate chains. This holds for the chains *N*-glycosidically linked to asparagine, as well as for the *O*-glycosidically linked moieties which may be attached to serine, threonine or hydroxylysine. The refinement of enzymic methods and chemical techniques, like permethylation analysis, for the structure determination contributed to a large extent to these results. Nevertheless, the classical approach has several disadvantages and shortcomings as pointed out before (Ref. 1). Especially it is highly time and material consuming and has severe limitations in the analysis of mixtures of closely related compounds. It should be noted that the structure analysis is carried out at the level of partial structures like glycopeptides, oligosaccharides and oligosaccharide-alditols. It is the (micro-)heterogeneity of a single carbohydrate chain of the glycoprotein which may give rise to a mixture of partial structures difficult to fractionate.
In 1977 we introduced the application of high-resolution ^{1}H-NMR spectroscopy as a new approach for the structure elucidation of underivatized carbohydrate chains obtained from glycoproteins. First of all the ^{1}H-NMR spectrum of the compound provides us with a structural

identity card which, even if the spectrum can not be interpreted, renders possible comparison with the spectra obtained from other sources, allowing the conclusion whether or not compounds are identical. The relative intensities of the signals in the NMR spectrum can be used as markers for the purity of the compound. Often it can be deduced from the spectrum whether or not the sample consists of more than one carbohydrate structure and in which molar ratios the components occur.

It could be shown that 360-MHz ^{1}H-NMR spectroscopy in conjunction with methylation analysis is excellently suited for elucidation of the primary structures of *N*- or *O*-glycosidically linked carbohydrate chains of glycoproteins (Ref. 1 - 17). For the interpretation of the NMR spectrum the concept of structural reporter groups was developed (Ref. 1), *i.e.*, the chemical shifts and coupling constants of protons resonating at clearly distinguishable positions in the spectrum bear the essential information to assign the primary structure. Recently a 500-MHz ^{1}H-NMR spectrometer became available for our studies, concomitant with a generally applicable computer resolution enhancement routine (Ref. 18). The increased sensitivity of this spectrometer system with respect to the 360-MHz apparatus affords spectra with a more favourable signal to noise ratio. Therefore the gain in spectral resolution, inherent to the stronger magnetic field, can be optimally utilized in Lorentzian to Gaussian transformation. As could be demonstrated, also the line widths of the structural reporter group signals can in several cases be used to derive pertinent information on primary structures (Ref. 18).

In this report the 500-MHz ^{1}H-NMR spectra of some *N*- and *O*-glycosidically linked carbohydrate chains will be discussed in detail.

RESULTS AND DISCUSSION

N-acetyllactosamine type glycopeptides and oligosaccharides derived from *N*-glycoproteins

The best defined asialo representatives of the *N*-acetyllactosamine type asparagine-bound carbohydrate chains of glycoproteins are shown in Fig. 1.

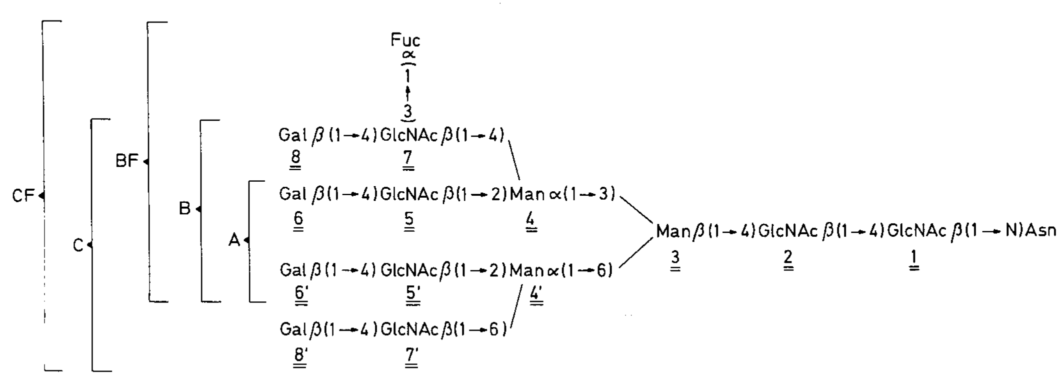


Fig. 1. The primary structures of the classes A, B, BF, C and CF asialo carbohydrate units of *N*-type glycoproteins. An additional Fuc residue may be attached in an α(1→6) linkage to GlcNAc-1. In sialylated analogues of these chains one or more NeuAc residues occur α(2→3) and/or α(2→6) linked to Gal residues.

The core structure, *i.e.*, the trimannosyl-*N*,*N*'-diacetylchitobiose moiety, is virtually invariant. In all glycopeptide samples derived from *N*-glycoproteins we observed this intact core, which is in agreement with the biosynthetic pathway for these compounds (Ref. 19 - 23). The asialo afuco bi-antennary compound (class A) can be conceived as a basic element of this type of carbohydrate chain. The ^{1}H-NMR spectral features of this compound can be used for the interpretation of more complex structures. The overall 500-MHz ^{1}H-NMR spectrum of such a glycopeptide, compound 1, together with the essential assignments, is given in Fig. 2. The chemical shifts of its structural reporter groups, *viz.* the anomeric protons, the Man H-2's and the *N*-acetyl group protons, are compiled in Table 1. In comparison to the earlier published 360-MHz ^{1}H-NMR spectrum of this glycopeptide (Ref. 3 & 8) the resolution enhanced 500-MHz ^{1}H-NMR spectrum reveals a large improvement. This can be clearly illustrated by the expanded structural reporter group regions and detailed assignments presented in Fig. 3. The resonance pattern of the three Man H-2's is as a whole characteristic for the bi-antennary type of glycopeptide (*cf.* Fig. 16), whereas the individual patterns contain the information on the α- or β-linkages of the Man residues (Ref. 18). The H-1 doublets of the terminal Gal

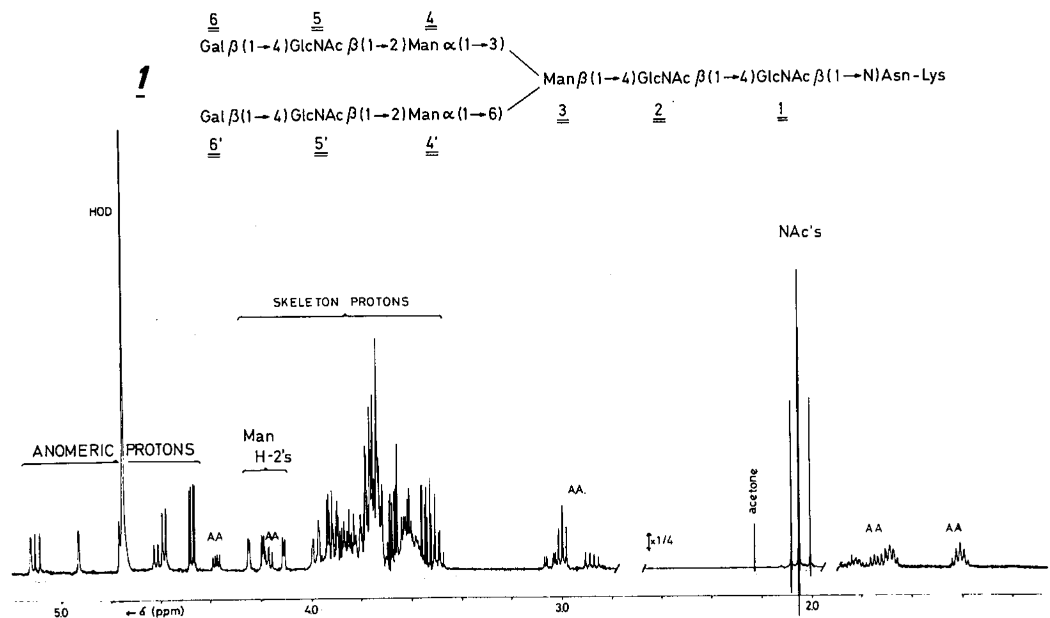


Fig. 2. Structure and resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 1. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. A.A. = amino acid protons.

residues 6 and 6'are well separated and could be assigned (see Fig. 3 and Table 1). It should be noted that their line widths are relatively small, which is in accordance with the expected mobility of terminal residues. The anomeric signals of the more internally located residues show broader lines. However, some variation exists, probably reflecting more or less segmental motion of the chain. The line widths of the *N*-acetyl signals of the core GlcNAc residues 1 and 2 are larger than those of the peripheral GlcNAc residues 5 and 5'. Addition of two NeuAc residues α(2→6) linked to Gal-6 and -6' of compound 1, respectively, affords compound 2. The structure of this glycopeptide is depicted in Fig. 4, together with the structural reporter group regions of its 500-MHz ^{1}H-NMR spectrum. The relevant NMR parameters of compound 2 are given in Table 1. New structural reporter groups are the H-3ax and H-3eq of NeuAc. The resonances of the H-3ax's of both residues coincide whereas those of the H-3eq's are well seperated. The assignment of the latter signals was carried out on the basis of the NMR spectrum of an enzymically prepared monosialo bi-antennary glycopeptide having the NeuAc residue α(2→6) linked to Gal-6 (Ref. 15). It is confirmed by the 500-MHz ^{1}H-NMR spectrum of a sialo bi-antennary glycopeptide mixture, the main component of which bears one NeuAc α(2→6) linked to Gal-6 and another one α(2→3) linked to Gal-6' (Ref. 10 & 18). Furthermore, in the spectrum of compound 2 two relatively sharp *N*-acetyl singlets are observed at δ = 2.028 ppm and δ = 2.029 ppm belonging to the two NeuAc residues. Owing to the attachment of NeuAc in α(2→6) linkage to Gal-6 and -6' their anomeric proton signals undergo identical upfield shifts, *viz.* $\Delta\delta$ = -0.025 ppm. The influence of α(2→6) linked NeuAc upon the chemical shifts of the anomeric protons of GlcNAc-5 and -5' and of Man-4 and -4' has been discussed in full detail earlier (Ref. 1, 5 & 13). Comparison of the *N*-acetyl regions of the spectra of this compound and the foregoing shows that the signals of GlcNAc-5 and -5' are shifted downfield and considerably broadened, which is typical for the introduction of α(2→6) linked NeuAc residues.

The NMR analysis of reducing oligosaccharides is highly relevant since they can occur in urine and other body fluids of patients with inborn errors of glycoconjugate metabolism (see *e.g.*, Ref. 24). However, investigation of such substances is more difficult than that of glycopeptides because in fact mixtures of compounds have to be studied. This will be illustrated for compounds related to the *N*-glycosidically linked type of carbohydrate chains. The mixtures of oligosaccharides secreted by the aforementioned patients contain *i.a.* compounds ending on GlcNAc-2, which are formed through the action of an *endo*-*N*-acetyl-β-D-hexosaminidase. In these substances wherein GlcNAc-2 is substituted at C-4, under the applied NMR measuring conditions an approximately constant ratio of its α- and β-pyranoses is found regardless the length of the carbohydrate chain, *viz.* α : β = 2 : 1. By consequence the NMR spectrum of such an oligosaccharide is a superposition of the subspectra of its two anomers containing GlcNAc-2 in the α- and β-pyranose forms, respectively.

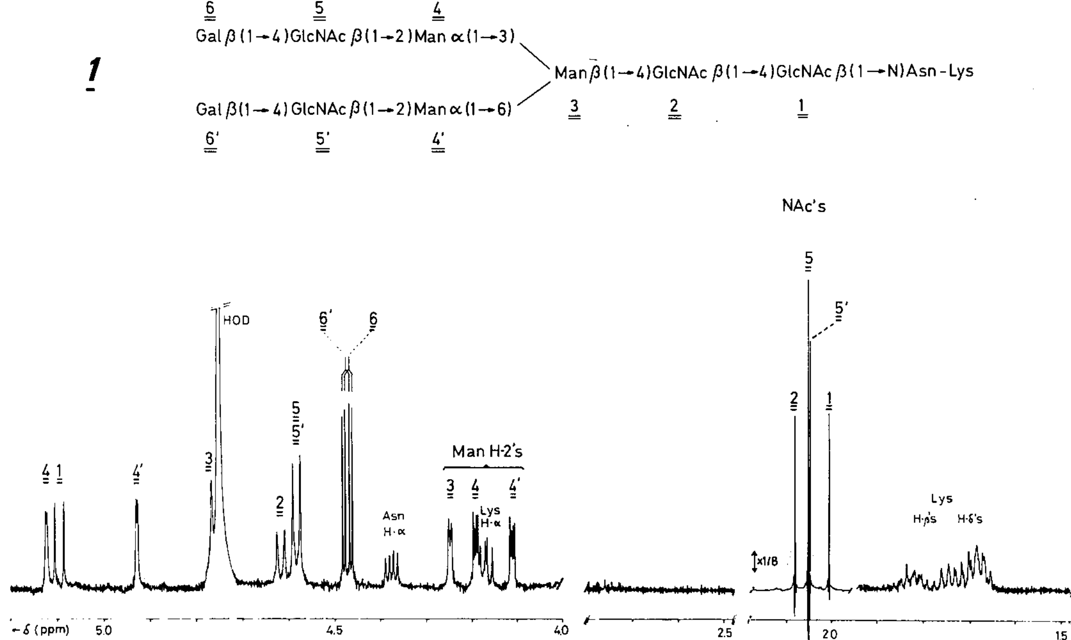
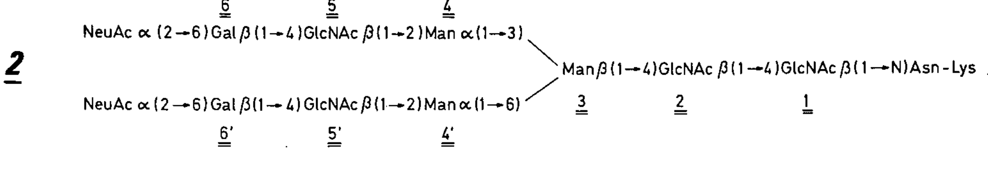


Fig. 3. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 1, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.



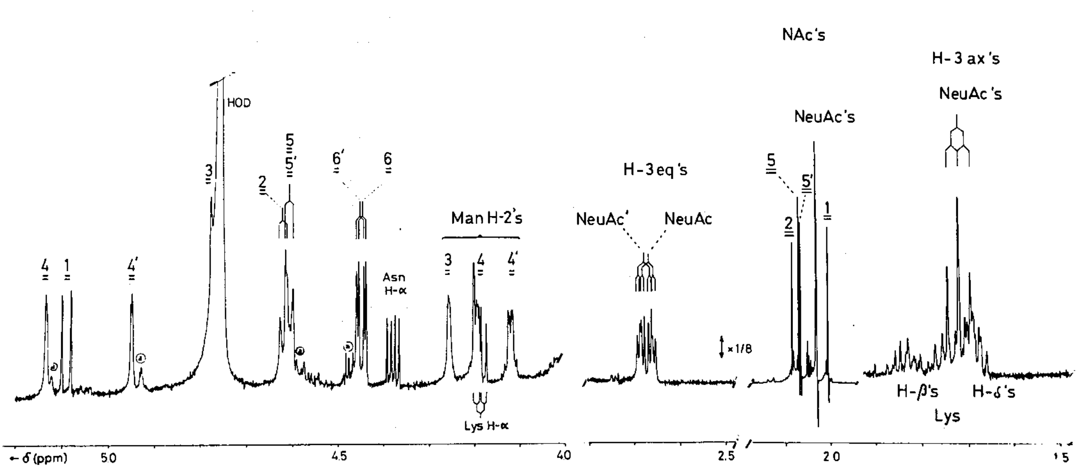


Fig. 4. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 2, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The signals of relatively low intensity, marked by *a* , point to the presence of small amount of partially and/or complete asialo analogues of compound 2 in the sample.

TABLE 1. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for asialo (compound 1) and bi-α(2→6)-sialo (compound 2) bi-antennary *N*-acetyllactosamine type glycopeptides [a]

reporter group	residue	compound 1 †	compound 2
H-1 of	1	5.094	5.088
	2	4.616	4.616
	3	4.765	4.773
	4	5.121	5.133
	4'	4.928	4.949
	5	4.582	4.603
	5'	4.582	4.603
	6	4.467	4.442
	6'	4.473	4.447
H-2 of	3	4.246	4.254
	4	4.190	4.195
	4'	4.109	4.116
H-3ax of	NeuAc	-	1.716
	NeuAc'	-	1.716
H-3eq of	NeuAc	-	2.666
	NeuAc'	-	2.672
NAc of	1	2.004	2.002
	2	2.079	2.081
	5	2.050	2.067
	5'	2.046	2.063
	NeuAc	-	2.029 [b]
	NeuAc'	-	2.028 [b]

[a] For complete structures and numbering of monosaccharide residues see Figs. 2, 3 and 4.
[b] Assignments may be interchanged.

The structure of compound 3, the oligosaccharide analogue of the asialo bi-antennary glycopeptide 1, together with relevant regions of its 500-MHz ^{1}H-NMR spectrum, is shown in Fig. 5. The chemical shifts of the structural reporter groups of 3 are compiled in Table 2. The effect of anomerization is most clearly recognizable from the structural reporter group signals of GlcNAc-2 (δ H-1α = 5.212 ppm; δ H-1β = 4.721 ppm; δ NAcα = 2.060 ppm and δ NAcβ = 2.057 ppm; intensity ratio of the signals α : β = 2 : 1). However, this effect is not limited to the reducing end sugar; it is also observable for more remote residues. The rough pattern of the Man H-1 and H-2 signals points to a bi-antennary type of structure (*cf.* Fig. 16). The chemical shifts of H-1 and H-2 of Man-3 in the β-anomer of the oligosaccharide 3 are essentially identical to those of the corresponding protons in the glycopeptide 1. The H-1 signal is accompanied by a resonance 0.01 ppm downfield, with similar shape and twice the intensity, belonging to H-1 of Man-3 in the oligosaccharide's α-anomer. The same feature exists for the H-2 of Man-3. For H-1 of Man-4' two well separated doublets occur which are assigned on the basis of their relative intensities (Fig.5). Comparison of the

Note †: In column headings of Table 1 and of other tables in this report compounds are represented by schematic structures; •— = neutral or amino sugar residue; △— = NeuAcα(2→3); ○— = NeuAcα(2→6)[only in compound 24 NeuGlα(2→6)]; □— = Fucα(1→•); ×— = GlcNAcα(1→4).

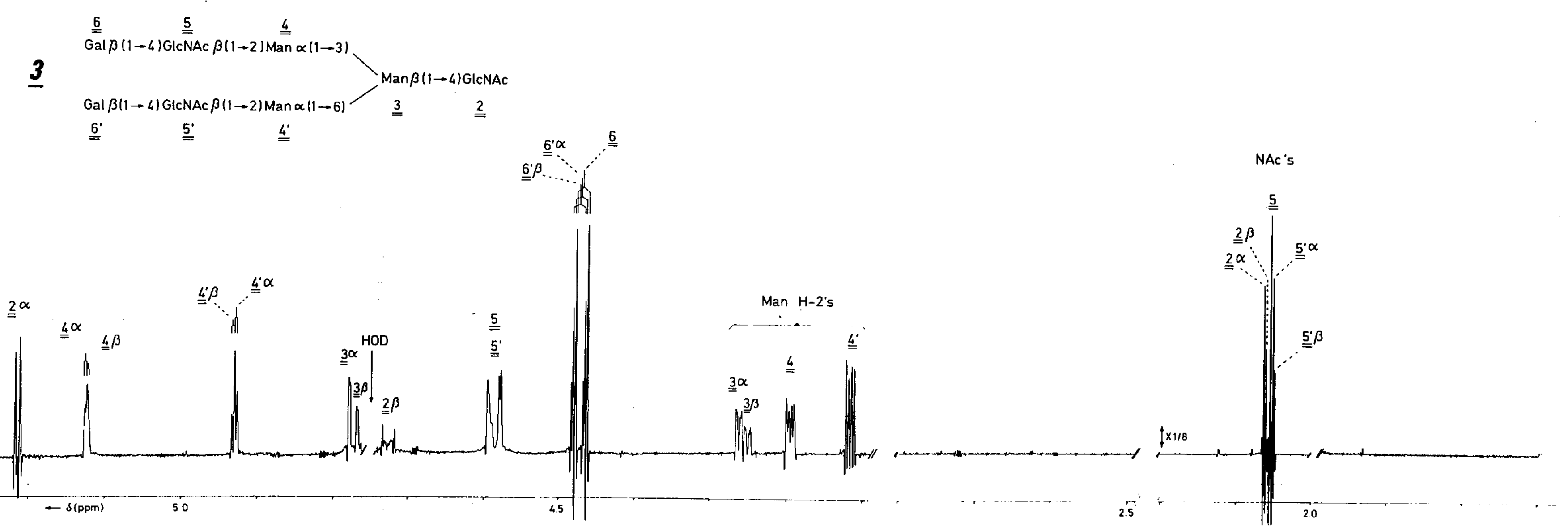


Fig. 5. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 3, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow.

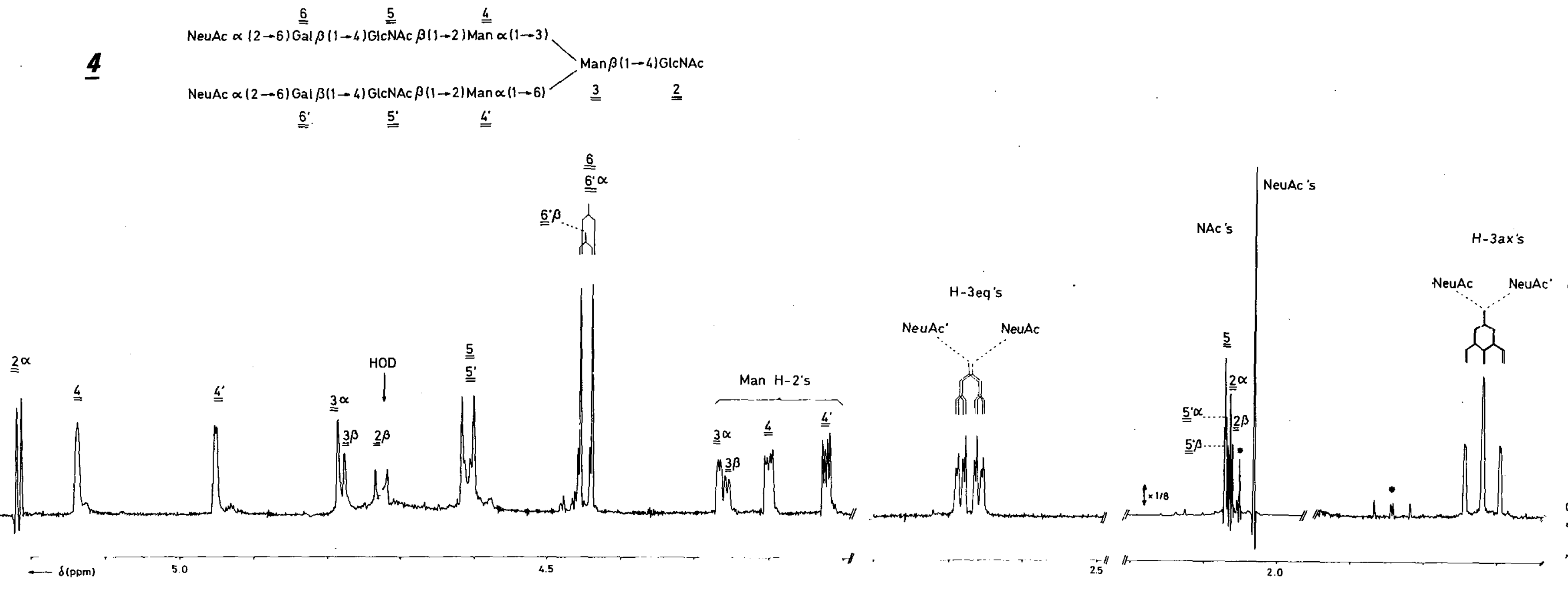


Fig. 6. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 4, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow. The doublet of doublets at δ = 1.842 ppm and the singlet at δ = 2.050 ppm, both indicated by an asterisk, belong to H-3ax and the *N*-acetyl methyl protons of (the β-anomer) of free NeuAc, respectively (Ref. 25), present in small amount in this sample.

TABLE 2. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for asialo (compound 3), bi-α(2→6)-sialo (compound 4), mono-α(2→6)-sialo (compounds 5 and 6) and bi-α(2→3)-sialo (compound 7) bi-antennary *N*-acetyllactosamine type oligosaccharides [a]

reporter group	residue	compound's anomer	compound 3	compound 4	compound 5	compound 6	compound 7
H-1 of	2	α	5.212	5.216	5.215	5.215	5.213
		β	4.721	≈4.72 [b]	4.725	4.725	≈4.72 [b]
	3	α	4.775	4.786	4.782	4.782	4.775
		β	4.765	4.777	4.772	4.772	4.763
	4	α	5.123	5.137¶	5.138¶	5.124	5.122
		β	5.121			5.122	5.120
	4'	α	4.927	4.952¶	4.929	4.950¶	4.923
		β	4.930		4.933		4.928
	5	α,β¶	4.583	4.608	4.608	4.583	4.578
	5'	α,β	4.586	4.608	4.586	4.608	4.578
	6	α,β	4.468	4.445	4.446	4.468	4.544
	6'	α	4.471	4.445	4.471	4.446	4.546
		β	4.473	4.449	4.473	4.449	4.547
H-2 of	3	α	4.259	4.266	4.265	4.260	4.256
		β	4.249	4.255	4.254	4.249	4.244
	4	α,β	4.193	4.199	4.198	4.194	4.192
	4'	α,β	4.114	4.121	4.115	4.119	4.104
H-3ax of	NeuAc	α,β	-	1.721	1.720	-	1.797
	NeuAc'	α,β	-	1.719	-	1.719	1.799
H-3eq of	NeuAc	α,β	-	2.669	2.669	-	2.758
	NeuAc'	α,β	-	2.672	-	2.672	2.758
NAc of	2	α	2.060	2.063	2.061	2.063	2.061
		β	2.057	2.060	2.058	2.060	2.058
	5	α,β	2.052	2.071	2.070	2.053	2.049
	5'	α	2.048	2.069	2.050	2.068	2.046
		β	2.046	2.066	2.048	2.066	2.043
	NeuAc	α,β	-	2.030	2.030	-	2.032
	NeuAc'	α,β	-	2.030	-	2.030	2.032

[a] For complete structures and numbering of monosaccharide residues see Figs. 5, 6, 7 and 8.
[b] The H-1 doublet of the β-anomer of GlcNAc-2 is partially hidden under the HOD-line at 300 K. Therefore its δ-value can not be determined more accurately (± 0.01 ppm).

Note ¶: If for a certain proton $|\Delta\delta(\alpha-\beta)| \leq 0.001$ ppm, only an average δ-value has been listed in this and other tables in this paper, since such a difference in chemical shift due to anomerization, if any, is not detectable at 500 MHz, unless the signal's lines are extremely sharp.

chemical shifts of H-1 of Man-4' in the α- and β-forms of the reducing oligosaccharide 3 (Table 2) shows that the signal belonging to the β-form is found at lower field than that of the α-form. Also for Man-4 anomerization gives rise to two signals for H-1, but in reversed order as compared to Man-4'. The $|\Delta\delta(\alpha-\beta)|$ effect upon H-1 of Man-4 is less outspoken than on H-1 of Man-4', and somewhat masked by the broader lines. The difference in line width between Man-4 and Man-4' anomeric proton signals can also be observed for the corresponding H-2 signals. For the GlcNAc-5 and -5' anomeric protons a relatively complex set of signals is obtained (Fig. 5), needing further study for detailed interpretation. With regard to their *N*-acetyl signals only a doubling of the singlet of GlcNAc-5' is observed due to anomerization. Concerning the anomeric proton signals of the Gal residues in 3 it has to be noted that three doublets occur with relative intensities of 1:2:3. The highest doublet belongs to H-1 of Gal-6. Its chemical shift is unaltered with respect to the glycopeptide 1. The chemical shift of the lowest doublet matches exactly that of H-1 of Gal-6' in the glycopeptide 1, therefore leaving the remaining signal to H-1 of Gal-6'in the α-form of the oligosaccharide 3.
In general the reporter groups of the lower branch residues, *i.e.*, 4', 5' and 6', appear to be more sensitive to anomerization of the oligosaccharide than those of the upper branch residues. Furthermore, the closer a residue is to the reducing end, the greater are the absolute differences between the chemical shifts of its (reporter group) protons, comparing α- and β-anomer of the compound. The reporter groups of residues 3 to 6/6' of the β-anomer of oligosaccharide 3 have identical chemical shifts as in glycopeptide 1, wherein GlcNAc-2 is β-linked to GlcNAc-1. Obviously for GlcNAc-2 the situation is completely different.
Compound 4 is the bisialo analogue of oligosaccharide 3, bearing two α(2→6) linked NeuAc residues. The structure of compound 4 and the reporter group regions of its 500-MHz ^{1}H-NMR spectrum are depicted in Fig. 6. Relevant NMR parameters of oligosaccharide 4 are summarized in Table 2. Essentially the NMR features of NeuAc and the influences of NeuAc on the chemical shifts of structural reporter groups of other residues are identical to those described above (*cf.* compound 2). However, with regard to the NeuAc residues, a small difference in chemical shift between the H-3ax's of the upper and lower branch residue exists. Furthermore, the H-3eq of the lower branch NeuAc gives rise to a doublet of doublets with broader lines than that of the upper branch NeuAc, whereas both NeuAc *N*-acetyl signals coincide. Anomerization of oligosaccharide 4 does not give rise to doubling of any of the structural reporter group signals of NeuAc. The difference between the chemical shifts of H-1 of Man-4 in the α- and β-form of the oligosaccharide is reduced in comparison to compound 3; the same holds for H-1 of Man-4' (see Table 2, and Footnote ¶ at that place).
To demonstrate the potency of high-resolution ^{1}H-NMR spectroscopy to analyzing complex mixtures of closely related compounds, the spectral data of a mixture of two reducing oligosaccharides, compounds 5 and 6, in a ratio 55:45, respectively, will be discussed. For the structures of these compounds and the 500-MHz ^{1}H-NMR spectrum of the mixture see Fig. 7. The NMR parameters of the structural reporter groups of oligosaccharides 5 and 6 are summarized in Table 2. Comparison of the data of this mixture of bi-antennary oligosaccharides with those of compounds 3 and 4 (Table 2) shows that for both the upper and lower branch the α(2→6)-sialo and the asialo form occur. The conclusion that a mixture of monosialo bi-antennary structures is involved rather than a mixture of a bisialo (4) and an asialo (3) oligosaccharide is based *i.a.* on the absence of signals of H-1 of Man-3 at δ = 4.775 ppm and δ = 4.765 ppm (see compound 3, Table 2). It should be noted that for the interpretation of the NMR spectrum of a mixture the relative intensities of the signals are helpful, as in this case in particular those of the H-3ax and also of the H-3eq signals of the NeuAc residues. It is evident that a 1:1 mixture is more difficult to analyze by NMR.
Sialic acid also occurs in α(2→3) linkage to Gal in oligosaccharides of the *N*-acetyllactosamine type,as for example in compound 7, the structure and 500-MHz ^{1}H-NMR spectrum of which are given in Fig. 8. The NMR parameters of the structural reporter groups of 7 are given in Table 2. As reported previously (Ref. 1, 5, 7, 10, 13 & 14) this type of NeuAc linkage to Gal is characterized by its own set of chemical shifts of the H-3's of NeuAc. The H-3eq resonances of both NeuAc residues in 7 coincide but those of the H-3ax's are separated. In this bi-antennary oligosaccharide the *N*-acetyl methyl protons of an α(2→3) linked NeuAc residue resonate at the same frequency as those of an α(2→6) linked residue (Table 2). The *N*-acetyl signals of GlcNAc-5 and 5' shift slightly upfield upon substitution of both Gal-6 and -6' at C-3 by a NeuAc residue (*cf.* compound 3). The effects of α(2→3) linked NeuAc upon the chemical shifts of the remaining structural reporter groups have been discussed earlier (Ref. 5). In line with the absence of a long-distance effect of α(2→3) linked NeuAc on the chemical shifts of the H-1's of Man-4 and -4' no changes are detectable on the doubling of their signals due to anomerization.

Extension of the bi-antennary structure with an *N*-acetyllactosamine unit β(1→4) linked to Man-4 furnishes the normal tri-antennary structure (class B, see Fig. 1). NMR consequences of this extension are illustrated for a glycopeptide, compound 8, and for an oligosaccharide, compound 9. Their structures and 500-MHz ^{1}H-NMR spectra are given in Figs. 9 and 11, respectively. The NMR parameters of 8 and 9 are compiled in Tables 3 and 4, respectively. The type of branching can be deduced from the chemical shifts of the structural reporter groups of the Man residues 3, 4 and 4' (Ref. 9 & 18). At a sample temperature of 300 K the H-1 signal of Man-3 in *N*-acetyllactosamine type glycopeptides and oligosaccharides is often

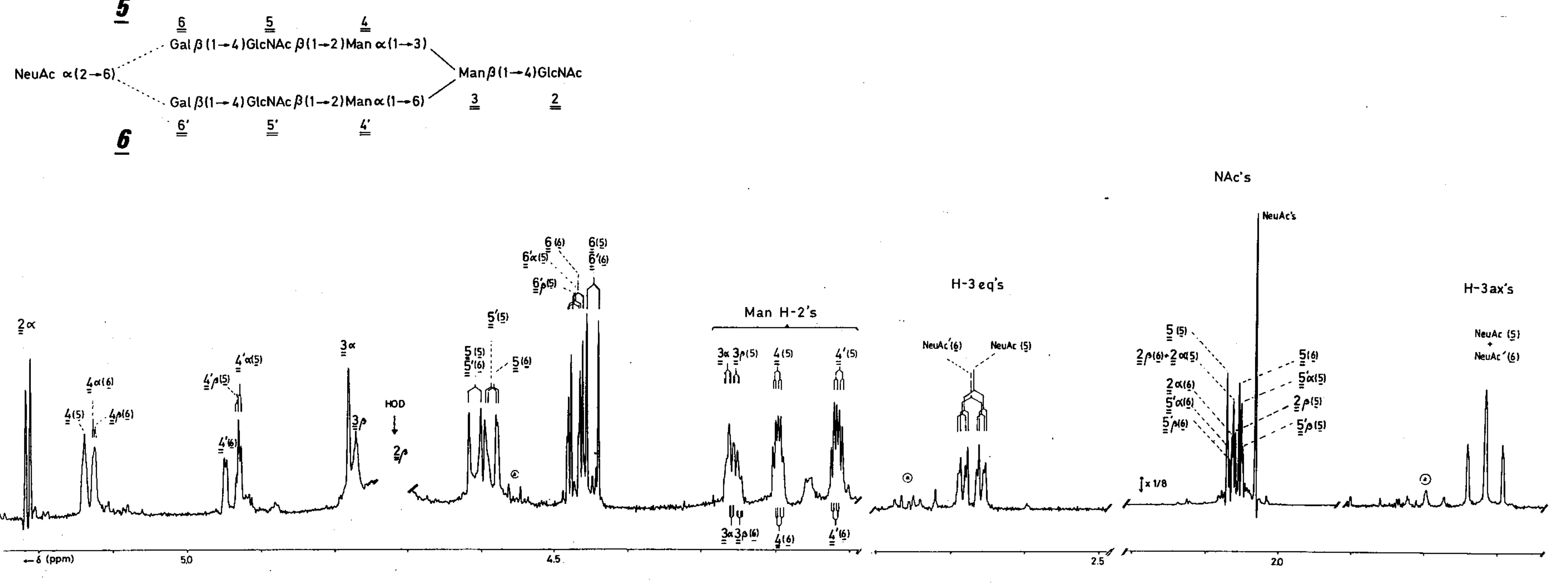


Fig. 7. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of a mixture of monosialo compounds 5 and 6, in ratio 55:45, together with their structures. The double-underlined numbers in the spectrum refer to the corresponding residues in the structures. Signals stemming from corresponding protons of compounds 5 and 6 coincide, unless otherwise indicated (see Table 2). The relative intensity of the *N*-acetyl proton region differs from that of other parts of the spectrum as indicated. The HOD-resonance, as well as the H-1 signal of the β-anomer of GlcNAc-2 which at a sample temperature of 300 K is partially hidden under this line, are left out from the spectrum; their rough position is indicated by an arrow. The sample is contaminated with a small amount of a similar oligosaccharide, bearing NeuAc α(2→3) linked to Gal-6 and/or 6'. This conclusion is based upon the presence of the H-3 signals of this type of NeuAc residue, and of the H-1 doublet of Gal at δ ≈ 4.54 ppm, all indicated by ⓐ (*cf.* Fig. 8).

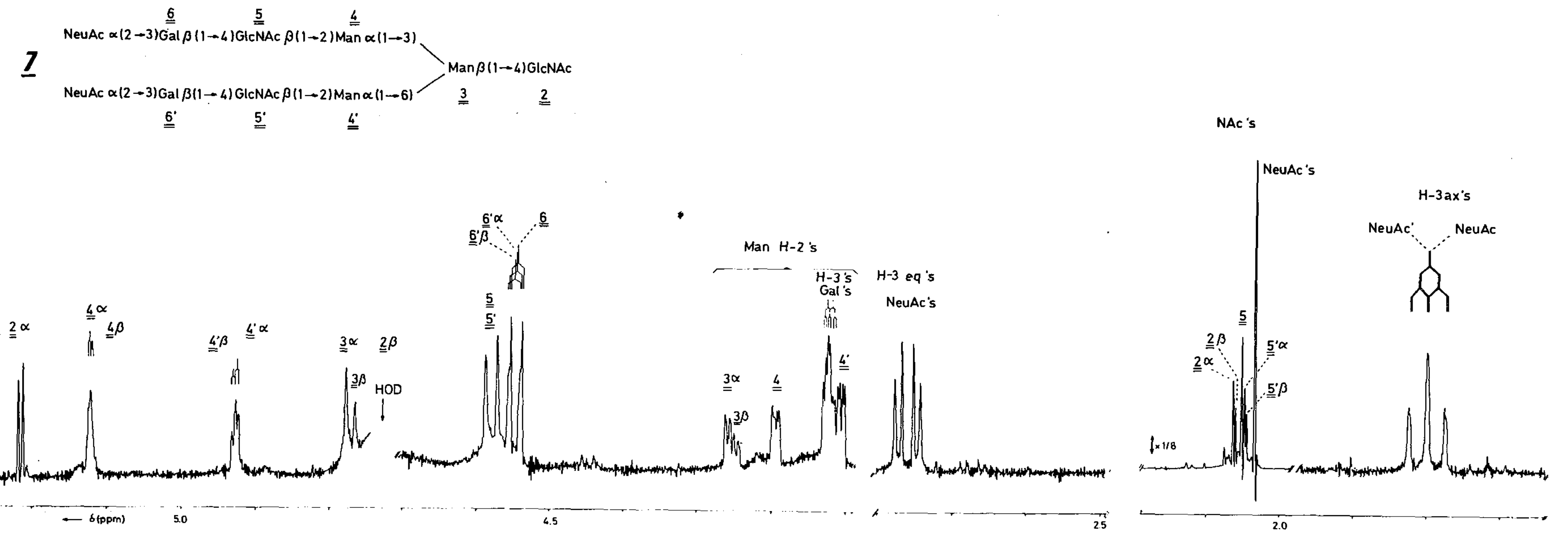


Fig. 8. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 7, together with its structure. *The numbers in the spectrum refer to the corresponding residues in the structure.* The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance and the H-1 doublet of the β-anomer of GlcNAc-2, almost coinciding at a sample temperature of 300 K, are left out from the spectrum; their rough position is indicated by an arrow.

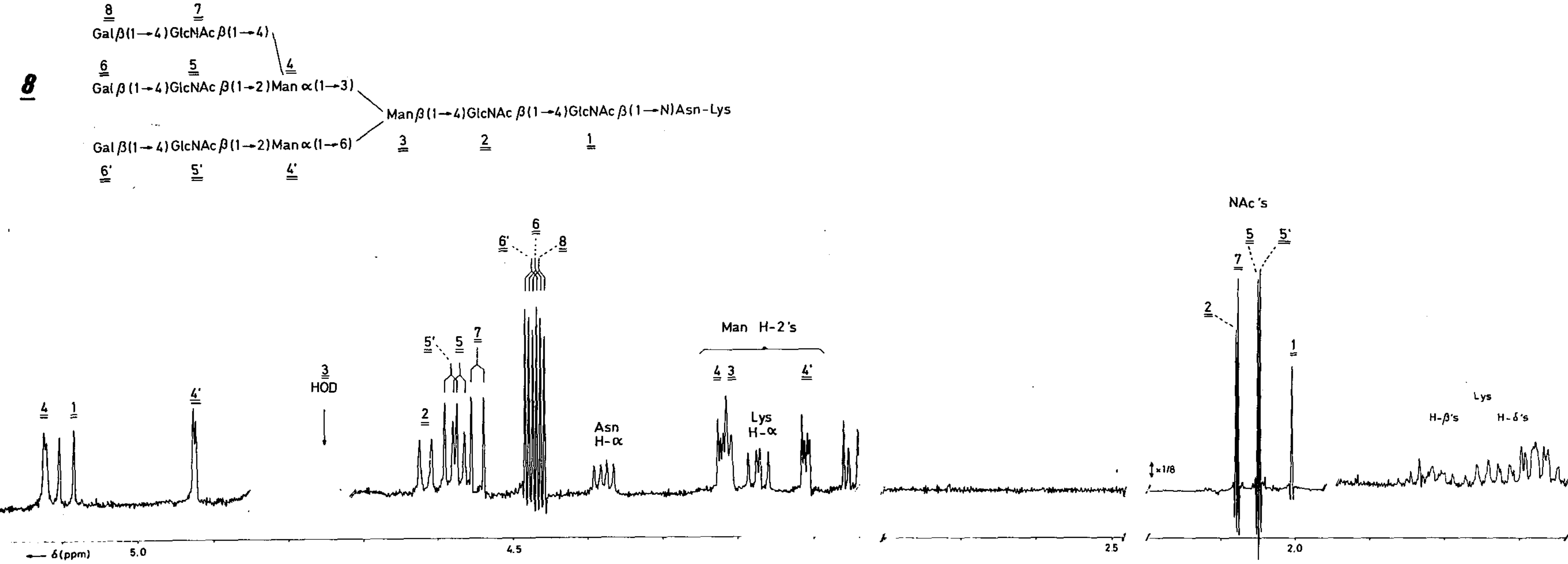


Fig. 9. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 8, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance, as well as the H-1 signal of Man-3, which at a sample temperature of 300 K is hidden under this line, are left out from the spectrum; their position is indicated by an arrow.

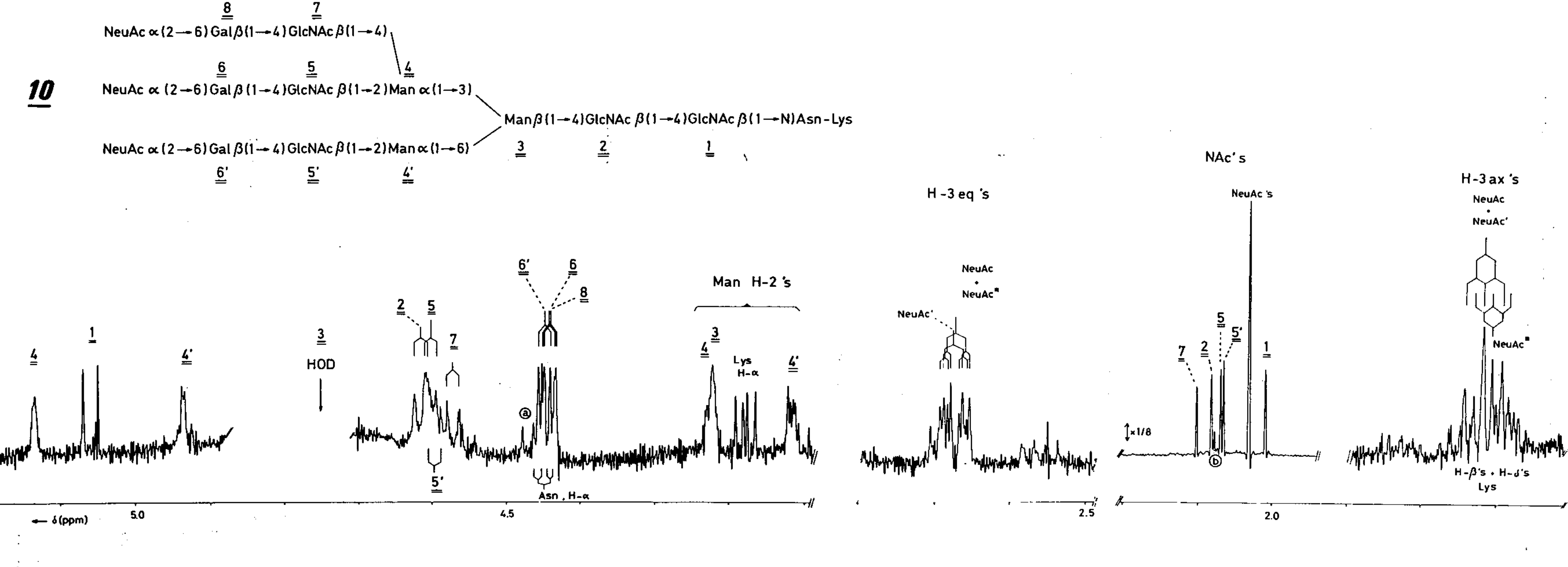


Fig. 10. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 10, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance, as well as the H-1 signal of Man-3, which at a sample temperature of 300 K is hidden under this line, are left out from the spectrum; their position is indicated by an arrow. The presence of a bi-α(2→6)-sialo tri-antennary glycopeptide possessing a terminal Gal-8 residue, as a minor constituent (≃ 10%) of this glycopeptide can be concluded from the Gal-8 H-1 doublet at $\delta \simeq 4.465$ ppm (ⓐ) and in particular from the *N*-acetyl singlet of GlcNAc-7 at $\delta = 2.077$ ppm (ⓑ), both indicative of an asialo third *N*-acetyllactosamine branch.

TABLE 3. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for asialo (compound 8) and tri-α(2→6)-sialo (compound 10) tri-antennary *N*-acetyllactosamine type glycopeptides [a]

reporter group	residue	compound 8	compound 10
H-1 of	1	5.092	5.060
	2	4.614	4.615
	3	4.755	≈4.76 [b]
	4	5.120	5.133
	4'	4.924	4.936
	5	4.570	4.594
	5'	4.580	4.602
	6	4.468	4.440
	6'	4.473	4.448
	7	4.545	4.571
	8	4.462	4.439
H-2 of	3	4.209	4.220
	4	4.218	4.225
	4'	4.108	4.116
H-3ax of	NeuAc	-	1.717
	NeuAc'	-	1.717
	NeuAc*	-	1.706
H-3eq of	NeuAc	-	2.670
	NeuAc'	-	2.674
	NeuAc*	-	2.670
NAc of	1	2.003	2.008
	2	2.078	2.082
	5	2.048	2.069
	5'	2.045	2.065
	7	2.075	2.101
	NeuAc	-	2.030
	NeuAc'	-	2.030
	NeuAc*	-	2.030

[a] For complete structures and numbering of monosaccharide residues see Figs. 9 and 11.

[b] The H-1 signal of Man-3 is hidden under the HOD-line at 300 K. Therefore its δ-value can not be determined more accurately (± 0.01 ppm).

hidden under the HOD-line. In this case small temperature changes ($|\Delta T| \approx 5$ K) can visualize this signal. The introduction of the third branch, β(1→4) linked to Man-4, gives rise to a characteristic upfield shift of H-1 of Man-3 in comparison to the bi-antenna (compounds 1 and 3, Tables 1 and 2). Interestingly the signal of H-3 of Man-4 becomes well resolved from the bulk upon the introduction of the third branch. This proton can be conceived as a new structural reporter group. In the NMR spectrum of oligosaccharide 9 the signal of H-3 of Man-4 is doubled with intensities in the anomeric ratio (*cf.* Fig. 16). The assignment of the remaining structural reporter groups (see Figs. 9 and 11) has been reported earlier (Ref. 18).

In the spectrum of oligosaccharide 9 anomerization does not furnish doubling of the structural reporter group signals of GlcNAc-7 and Gal-8.
Attachment of NeuAc's α(2→6) linked to each of the three branches of compound 8, leads to glycopeptide 10. The structure and 500-MHz ^{1}H-NMR spectrum of the latter compound are presented in Fig. 10; its relevant NMR data are summarized in Table 3. In comparison to the bi-α(2→6)-sialo bi-antennary glycopeptide 4, an additional set of H-3 signals of NeuAc is observed in the spectrum of 10 (δ H-3ax = 1.706 ppm; δ H-3eq = 2.670 ppm), which is typical for NeuAc α(2→6) linked to Gal-8 (indicated as NeuAc* in Table 3 and Fig. 10). The influences of this NeuAc residue on the chemical shifts of the structural reporter groups of Gal-8 and GlcNAc-7 are analogous to the effects described for NeuAc α(2→6) linked to Gal-6 or -6' *(vide supra)*. In particular the shift increment (Δδ = 0.026 ppm; Table 3) and the line broadening of the *N*-acetyl signal of GlcNAc-7 are highly characteristic. However, there is no effect of NeuAc α(2→6) linked to Gal-8 on the chemical shifts of H-1 and H-2 of Man-4 (*cf.* compounds 4 and 8). Surprisingly H-1 of Man-4' undergoes an upfield shift in comparison to the bisialo bi-antennary glycopeptide 4. It is not yet clear whether NeuAc* attached to Gal-8 exerts a direct through-space action on Man-4' or reduces the influence of NeuAc' attached to Gal-6' on this Man residue.
Previously we reported on the 360-MHz ^{1}H-NMR spectroscopy of compound 11, a tri-antennary oligosaccharide bearing two α(2→6) and one α(2→3) linked NeuAc residues (Ref. 5 & 13). Reinvestigation of this sample by 500-MHz ^{1}H-NMR spectroscopy revealed that in fact it consists of a mixture of components, the major one being compound 11, the minor one (≈ 8 %) compound 12. The latter one is the oligosaccharide analogue of glycopeptide 10. The structures of compounds 11 and 12 and the 500-MHz ^{1}H-NMR spectrum of the mixture are presented in Fig. 12. In addition to the spectral features of oligosaccharide 11 discussed earlier (Ref. 5), significant new information can be provided. The H-3ax signals and also the H-3eq signals of NeuAc and NeuAc', α(2→6) linked to Gal-6 and -6' respectively, are well separated. The assignments are consistent with those of compound 4 (Table 2). The *N*-acetyl region of the spectrum has now been interpreted in full detail (see Fig. 12). The shift increments observed for the *N*-acetyl signals of GlcNAc-5 and -5' in the step from compound 3 to 4 (see Table 2) are similar to those in the step from compound 9 to 11 or 12 (Table 4). The NeuAc* α(2→3) linked to Gal-8 in compound 11 introduces an upfield shift for the *N*-acetyl signal of GlcNAc-7 in comparison to the asialo tri-antennary oligosaccharide 9 (*cf.* the *N*-acetyl signals of GlcNAc-5 and -5' in the step from compound 3 to 7, Table 2, *vide supra*). The NeuAc* residue α(2→6) linked to Gal-8, which occurs in the minor component 12, can be recognized from its set of H-3 chemical shifts together with the chemical shifts of the *N*-acetyl methyl protons of GlcNAc-7 (δ = 2.102 ppm)(*cf.* compound 10, Table 3). Finally it is worth mentioning that the *N*-acetyl protons of this α(2→6) linked NeuAc* residue resonate at δ = 2.028 ppm, whereas the *N*-acetyl signals of all other NeuAc residues of 11 and 12 are found at δ = 2.030 ppm.

In many *N*-acetyllactosamine type carbohydrate chains, *N*-glycosidically linked to the peptide backbone, one or more Fuc's may occur as terminal residues. As reported previously (Ref. 1 & 14) each of such Fuc residues has in general three structural reporter groups *viz.* its H-1, H-5 and H-6's. The set of their chemical shifts together with the coupling constant $J_{1,2}$ are indicative of the type of linkage of Fuc and of the nature and environment of the monosaccharide residue to which Fuc is attached.
In case a Fuc residue is linked to the trimannosyl-*N,N'*-diacetylchitobiose core, solid evidence only exists for an α(1→6) linkage to GlcNAc-1 (*cf.* Ref. 2). Compound 13 contains this structural element. Its structure and 500-MHz ^{1}H-NMR spectrum are given in Fig. 13. The chemical shifts of the structural reporter groups of compound 13 are summarized in Table 5. The interpretation of the 360-MHz ^{1}H-NMR spectrum of this glycopeptide has been reported earlier (Ref. 6). Especially the effects of introduction of Fuc α(1→6) linked to GlcNAc-1 upon the chemical shifts of the structural reporter groups of GlcNAc-2 have to be emphasized (*cf.* compound 1, Table 1). The chemical shifts of H-1 and H-2 of Man-3 are characteristic for a mono-α(1→6) substitution of this residue.
Another well defined type of Fuc linkage is the α(1→3) attachment to a peripheral GlcNAc residue. Earlier we have described the 360-MHz ^{1}H-NMR characteristics of such a structural element occurring in a partially degraded glycopeptide, compound 14 (Ref. 6). Upon reconsideration of this sample with 500-MHz ^{1}H-NMR spectroscopy it appeared to be a mixture of at least three structurally related compounds *viz.* 14, 15 and 16 in a ratio 6:1:1. The structures of these glycopeptides and the 500-MHz ^{1}H-NMR spectrum of the mixture are depicted in Fig. 14. The NMR data of 14, 15 and 16 are compared in Table 5. Inspection of the Fuc CH_3 region of the spectrum shows that the signals at that place can be divided into four doublets at δ = 1.234 ppm, δ = 1.206 ppm, δ = 1.177 ppm and δ = 1.172 ppm in relative intensities 1:8:6:1. The doublet at δ = 1.206 ppm belongs to the CH_3-group protons of Fuc α(1→6) linked to GlcNAc-1. This assignment is in accordance with the resonance positions of the other structural reporter groups: δ H-1 = 4.876 ppm and δ H-5 = 4.126 ppm (*cf.* compound 13). The doublet at δ = 1.177 ppm is ascribed to Fuc α(1→3) linked to GlcNAc-5' in the major component of the mixture, compound 14. The other structural reporter groups of this Fuc residue, *viz.* H-1 and H-5, resonate at δ = 5.127 ppm and δ = 4.830 ppm, respectively. The remaining CH_3 doublets and including part of the doublet at δ = 1.206 ppm are derived from the minor components.

TABLE 4. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for asialo (compound 9), mixed α(2→3)- and α(2→6)-sialo (compound 11) and tri-α(2→6)-sialo (compound 12) tri-antennary *N*-acetyllactosamine type oligosaccharides [a]

reporter group	residue	compound's anomer	compound 9	compound 11	compound 12
H-1 of	2	α	5.211	5.215	5.215
		β	4.723	≈4.72 [b]	≈4.72 [b]
	3	α	4.767	≈4.77 [b]	≈4.77 [b]
		β	4.757	≈4.76 [b]	≈4.76 [b]
	4	α	5.119	5.134	5.134
		β	5.118		
	4'	α	4.923	4.943	4.943
		β	4.925		
	5	α,β	4.569	4.594	4.594
	5'	α,β	4.584	4.605	4.605
	6	α,β	4.468	4.443	4.443
	6'	α	4.471	4.443	4.443
		β	4.473	4.447	4.447
	7	α,β	4.546	4.551	4.573
	8	α,β	4.463	4.545	4.441
H-2 of	3	α	4.223	4.228	4.228
		β	4.212	4.217	4.217
	4	α,β	4.218	4.221	4.221
	4'	α,β	4.114	4.119	4.119
H-3ax of	NeuAc	α,β	-	1.719	1.719
	NeuAc'	α,β	-	1.716	1.716
	NeuAc*	α,β	-	1.801	1.705
H-3eq of	NeuAc	α,β	-	2.670	2.670
	NeuAc'	α,β	-	2.672	2.672
	NeuAc*	α,β	-	2.757	2.670
NAc of	2	α	2.060	2.062	2.062
		β	2.057	2.059	2.059
	5	α,β	2.050	2.067	2.069
	5'	α	2.048	2.067	2.067
		β	2.046	2.065	2.065
	7	α,β	2.078	2.074	2.102
	NeuAc	α,β	-	2.030	2.030
	NeuAc'	α,β	-	2.030	2.030
	NeuAc*	α,β	-	2.030	2.028

[a] For complete structures and numbering of monosaccharide residues see Figs. 10 and 12.
[b] Values can not be determined more accurately (± 0.01 ppm) due to interference of the relatively broad HOD-line in the spectrum of the mixture of compounds 11 and 12.

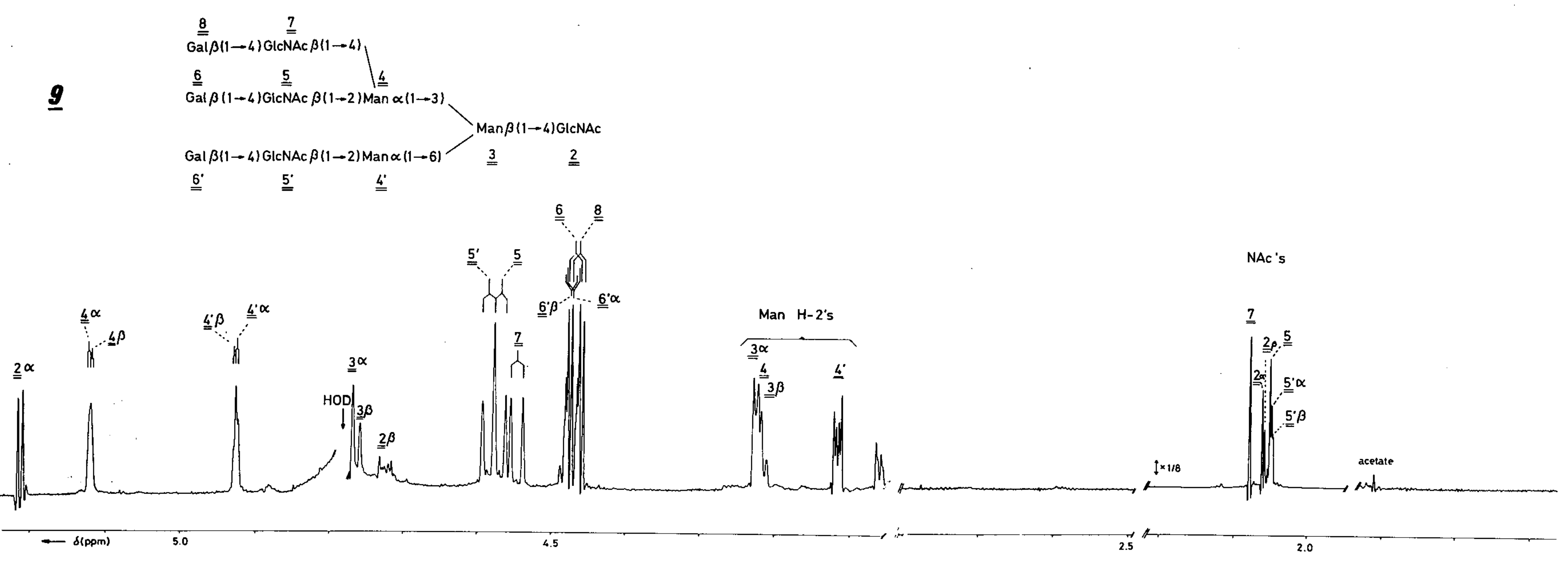


Fig. 11. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 9, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow.

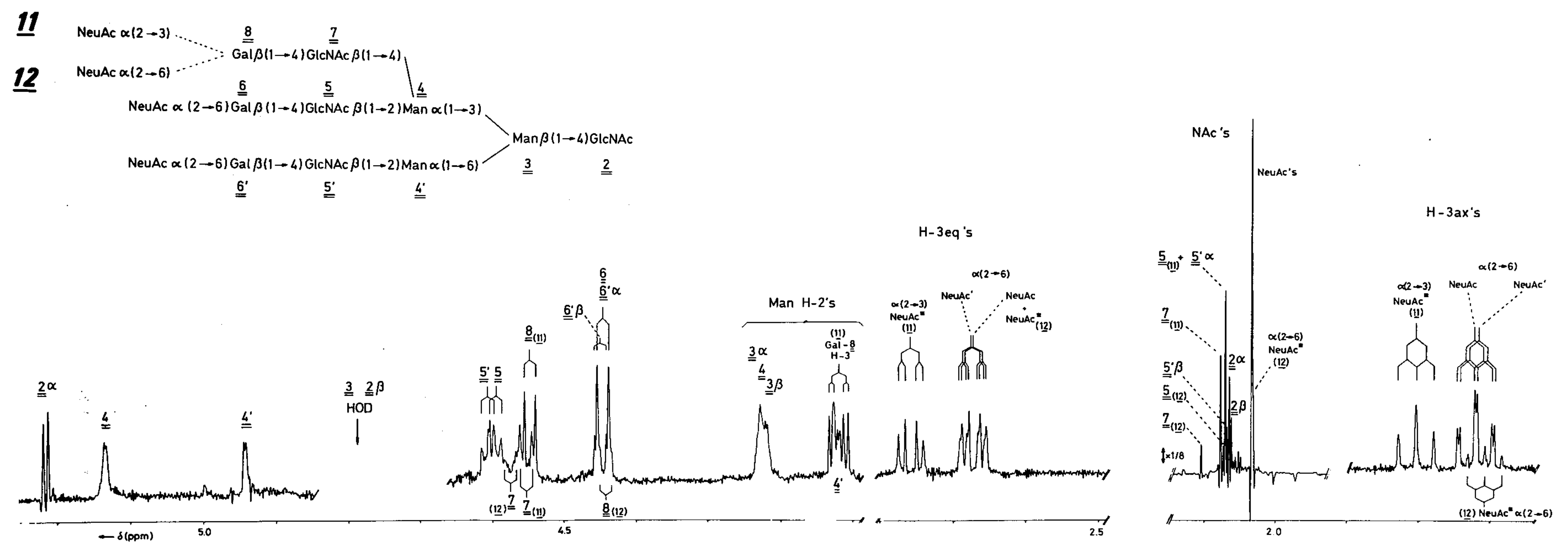


Fig. 12. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of a mixture of trisialo compounds 11 and 12, differing in the type of linkage of NeuAc* to Gal-8, in ratio 92:8, together with the structures. The double-underlined numbers in the spectrum refer to the corresponding residues in the structures. Signals stemming from corresponding protons of compounds 11 and 12 coincide, unless otherwise indicated (see Table 4). 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 and GlcNAc-2 (β) which at a sample temperature of 300 K are partially hidden under this line, are left out from the spectrum; the position of the HOD-line is indicated by an arrow.

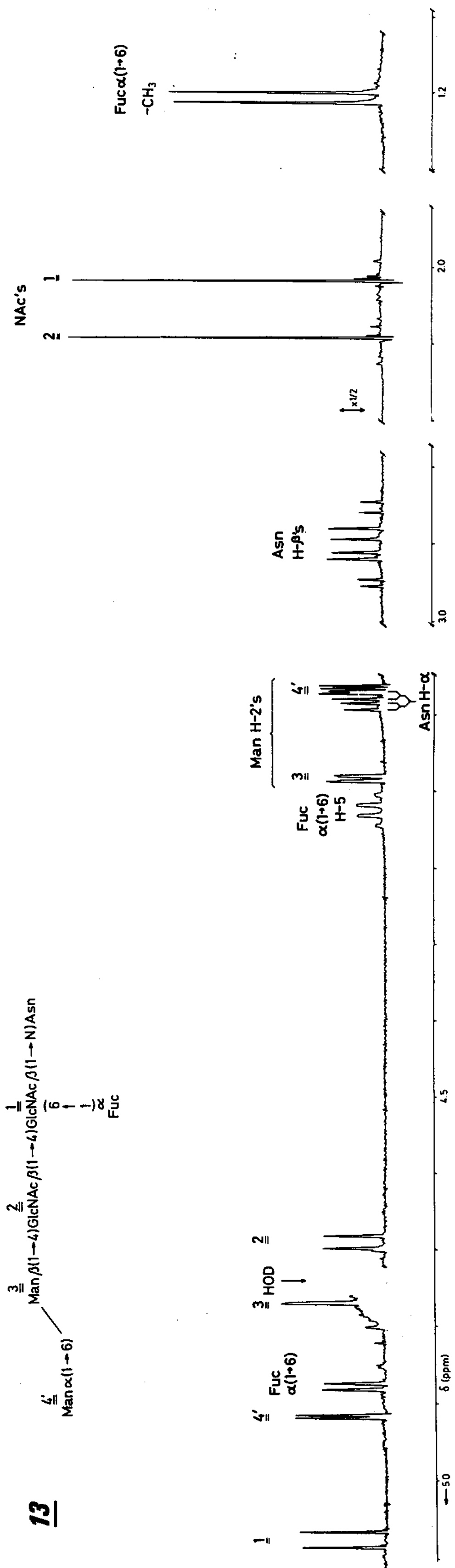


Fig. 13. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 13, together with its structure. The numbers in the spectrum refer to the corresponding residues in the structure. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow.

TABLE 5. [1]H Chemical shifts of structural reporter groups of constituent monosaccharides for mono-antennary *N*-acetyllactosamine type glycopeptides containing only α(1→6) linked Fuc (compound 13), both α(1→6) and α(1→3) linked Fuc (compounds 14 and 15) or both α(1→6) and α(1→2) linked Fuc (compound 16) [a]

reporter group	residue	compound 13	compound 14	compound 15	compound 16
H-1 of	1	5.076	5.073	5.073	5.073
	2	4.690	4.687	4.677	4.687
	3	4.770	4.764	4.775	4.764
	4	-	-	5.110	-
	4'	4.916	4.906	-	4.918
	5	-	-	4.585	-
	5'	-	4.585	-	≈4.58 [b]
	6	-	-	4.442	-
	6'	-	4.449	-	4.541
H-2 of	3	4.083	4.081	4.228	4.081
	4	-	-	4.187	-
	4'	3.967	4.087	-	4.087
NAc of	1	2.018	2.016	2.016	2.016
	2	2.095	2.095	2.090	2.095
	5	-	-	2.043	-
	5'	-	2.040	-	2.072
H-1 of	Fucα(1→6)	4.877	4.876	4.876	4.876
H-5 of	Fucα(1→6)	4.125	4.126	4.126	4.126
CH_3 of	Fucα(1→6)	1.209	1.206	1.206	1.206
H-1 of	Fucα(1→3)	-	5.127	5.126	-
H-5 of	Fucα(1→3)	-	4.830	4.830	-
CH_3 of	Fucα(1→3)	-	1.177	1.172	-
H-1 of	Fucα(1→2)	-	-	-	5.309
H-5 of	Fucα(1→2)	-	-	-	≈4.22 [b]
CH_3 of	Fucα(1→2)	-	-	-	1.234

[a] For complete structures and numbering of monosaccharide residues see Figs. 13 and 14.
[b] Values can not be determined more accurately (± 0.01 ppm) due to the low amount of compound 16 present in the mixture together with the relatively broad lines of the H-1 doublet of GlcNAc-5', as well as of the H-5 signal of Fuc α(1→2) linked to Gal-6', respectively.

For the major component, 14, the 500-MHz ^{1}H-NMR spectrum affords a significant refinement of all NMR data obtained at 360 MHz (Ref. 6). The introduction of Fuc α(1→3) linked to GlcNAc-5' has no influence on the chemical shift of H-1 of this GlcNAc but the lines of the signal are considerably broadened. Furthermore, the *N*-acetyl singlet of GlcNAc-5' undergoes a significant upfield shift (see Table 5). The anomeric signals of the adjacent residues Man-4' and Gal-6' also shift upfield.

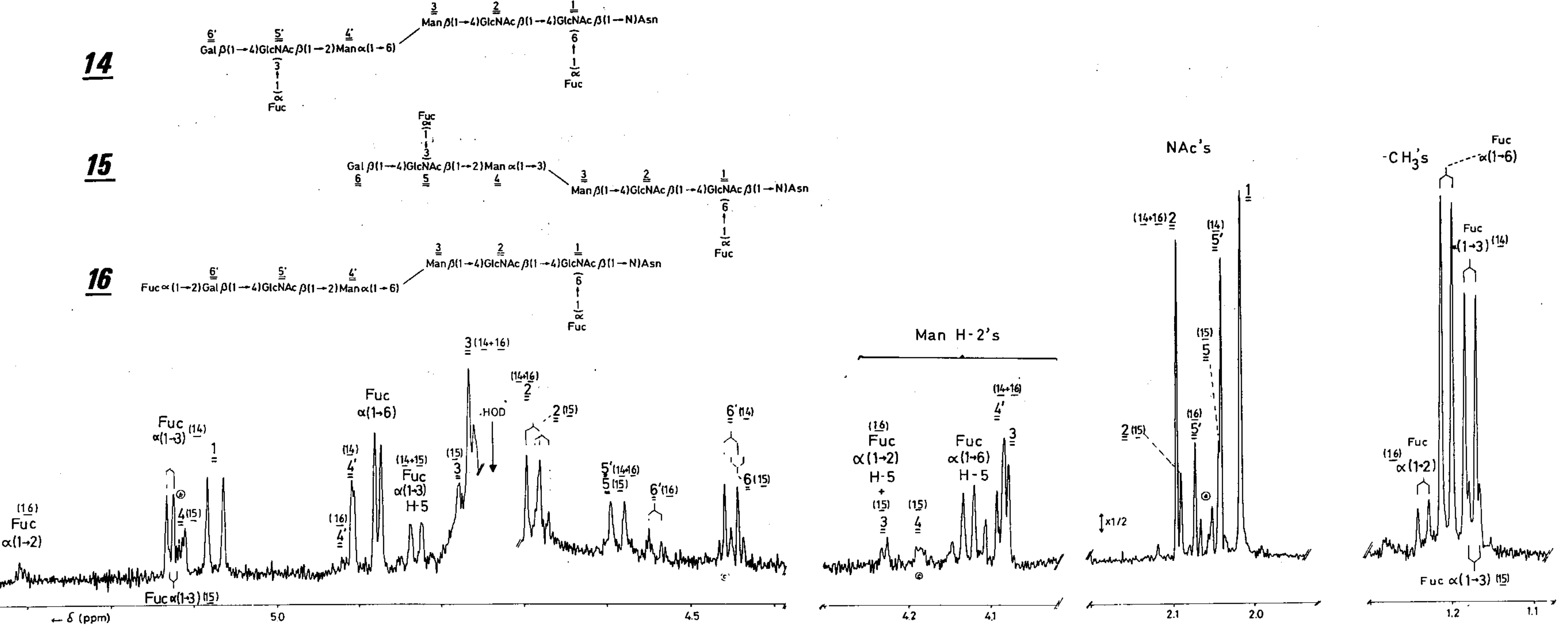


Fig. 14. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of a mixture of *i.a.* the bifuco mono-antennary glycopeptides 14, 15 and 16, in ratio 6:1:1, together with their structures. The double-underlined numbers in the spectrum refer to the corresponding residues in the structures. Signals stemming from corresponding protons of compounds 14, 15 and 16 coincide, unless otherwise indicated (see Table 5). The relative intensity of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow.

As is obvious from the presence of not yet assigned resonances in the spectrum (*e.g.* ⓐ, ⓑ and ⓒ's), other minor components are present in this glycopeptide mixture. Mainly from the Gal H-1's pattern (δ H-1's of 6' and 6'' = 4.442 and 4.456 ppm) as well as from the H-4 signal of Gal-6' at δ ≈ 4.18 ppm evidence can be provided for the presence of a mono-antennary lower branch glycopeptide, containing a so-called additional *N*-acetyllactosamine unit (5''-6'') with Fuc α(1→3) linked to GlcNAc-5'', in the mixture. The structure of the latter one is shown here:

Gal β(1→4)GlcNAc β(1→3)Gal β(1→4)GlcNAc β(1→2)Man α(1→6) Man β(1→4)GlcNAc β(1→4)GlcNAc β(1→N)Asn, with Fuc α(1→3) on GlcNAc-5'' and Fuc α(1→6) on GlcNAc-1 (residues numbered 6'', 5'', 6', 5', 4', 3, 2, 1)

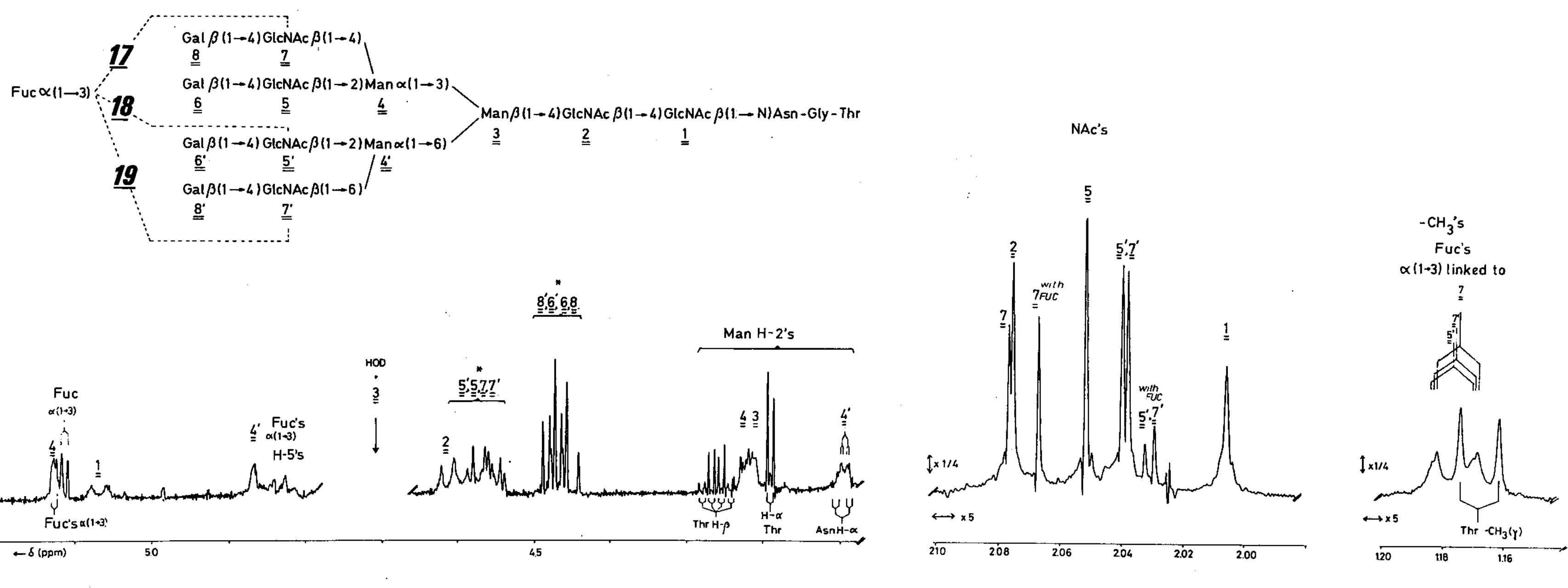


Fig. 15. Structural reporter group regions of the resolution enhanced 500 MHz ^{1}H-NMR spectrum of a mixture of monofuco tetra-antennary glycopeptides, *i.a.* 17, 18 and 19, differing in the location of the α(1→3) linked Fuc residue, together with the structures of these components. The numbers in the spectrum refer to the corresponding residues in the structures. The relative intensity scale, as well as the chemical shift scale of the *N*-acetyl proton region and the Fuc CH_3-region differ from those of the other parts of the spectrum as indicated. The HOD-resonance and the H-1 signal of Man-3 are left out from the spectrum; their position is indicated by an arrow. A detailed interpretation of the H-1 signals of the peripheral GlcNAc residues as well as of the Gal residues (indicated by asterisks), which occur in branches with and without Fuc attached to the respective GlcNAc-residues, in this mixture of components, has not yet been achieved.

One of the minor constituents of the mixture, compound 15, is also a mono-antennary glycopeptide containing the upper branch residues 4, 5 and 6 besides the core residues 1, 2 and 3 as well as a Fuc residue α(1→6) linked to GlcNAc-1. This conclusion is based on the presence of H-1 and H-2 resonances of Man-4. In accordance with this assignment the resonance positions of the structural reporter groups of GlcNAc-2 and Man-3 (in particular its δ H-2, being 4.228 ppm) deviate considerably from those for compound 14. GlcNAc-5 bears a Fuc residue in α(1→3) linkage, which is characterized by the following set of chemical shifts of Fuc structural reporter groups: δ H-1 = 5.126 ppm, δ H-5 = 4.830 ppm and δ CH_3 = 1.172 ppm. The structural reporter group signals of Man-4, GlcNAc-5 and Gal-6 occupy resonance positions which are typical for Fuc α(1→3) linked to GlcNAc-5 (*cf.* compound 14, Table 5).
The presence of a fourth set of Fuc structural reporter groups: δ H-1 = 5.309 ppm, δ H-5 ≈ 4.22 ppm and δ CH_3 = 1.234 ppm, is indicative of another minor component in the mixture containing a Fucα(1→2)Galβ(1→•) moiety (Ref. 1): compound 16. This mono-antennary compound contains the lower branch on the basis of the relative intensities of several structural reporter group signals being markers for the type of mono-substitution of Man-3, *e.g.*, the *N*-acetyl singlets of GlcNAc-2 occurring at δ = 2.095 ppm and δ = 2.090 ppm in a ratio 7:1. This type of Fuc gives rise to downfield shifts of H-1 of Gal-6' and the *N*-acetyl methyl protons of GlcNAc-5' (see Table 5).
The 500-MHz ^{1}H-NMR analysis of this glycopeptide mixture has provided *i.a.* the NMR parameters for Fuc α(1→2) linked to Gal in an *N*-glycosidically linked carbohydrate chain (see also legend to Fig. 14).
Also in higher antennary *N*-acetyllactosamine type structures Fuc may occur in α(1→3) linkage to one or more peripheral GlcNAc residues. The 360-MHz ^{1}H-NMR features of tri- and tetra-antennary glycopeptides bearing a Fuc α(1→3) linked to GlcNAc-7 (classes BF and CF, see Fig. 1) are well documented (Ref. 8 & 12). The reinvestigation of such a monofuco tetra-antennary glycopeptide sample at 500 MHz demonstrated that this sample in fact consists of a mixture of at least three isomers, the originally described component being the main constituent. The structures of the three isomers: compounds 17, 18 and 19, and the 500-MHz ^{1}H-NMR spectrum of the mixture are presented in Fig. 15. Close inspection of the *N*-acetyl region of the spectrum learns that besides the earlier (Ref. 8) described shift decrement of the *N*-acetyl singlet of GlcNAc-7 in comparison to the afuco tetra-antennary compound (Δδ = -0.01 ppm), also similar shifts are observed for the *N*-acetyl signals of GlcNAc-5' and -7' (from δ = 2.040 ppm and δ = 2.038 ppm to δ = 2.033 ppm and δ = 2.029 ppm, respectively) (see Fig. 15). This indicates that in addition to the main component having Fuc attached to GlcNAc-7 (compound 17), also two minor constituents occur in the mixture which have Fuc attached in α(1→3) linkage to GlcNAc-5' (compound 18) or GlcNAc-7' (compound 19), respectively. It should be noted that in the latter two compounds the chemical shift of the *N*-acetyl protons of GlcNAc-7 has the same value as for an afuco tetra-antennary glycopeptide (see Ref. 18), thereby explaining the intensity of the signal at δ = 2.077 ppm.

In general the substitution of Man-4 and/or Man-4' with various *N*-acetyllactosamine branches in *N*-glycosidically linked carbohydrate chains and in related reducing oligosaccharides comes to expression in the resonance positions and patterns of the structural reporter group signals of the Man residues 3, 4 and 4' (Ref. 9 & 18). In 500-MHz ^{1}H-NMR spectra the set of H-2 resonance patterns is already that characteristic that the type of antennary structure can be recognized. This is illustrated both for glycopeptides and for oligosaccharides ending on GlcNAc-2, in Fig. 16.

Oligomannoside type glycopeptides and oligosaccharides derived from *N*-glycoproteins

500-MHz ^{1}H-NMR spectroscopy gave a real breakthrough in the NMR analysis of glycopeptides and oligosaccharides of the oligomannoside type of *N*-glycosidically linked carbohydrate chains. A complete assignment of all structural reporter groups for structures containing up to nine Man residues could be achieved, as will be published elsewhere. The NMR study of such compounds is hampered by the similarity of the constituting units. Only the H-1 and H-2 resonances of each of the Man residues serve as markers for the primary structure. Obviously this demands for a very high spectral resolution. Crucial features for the spectral assignments for this type of carbohydrate structures are the parameters of the structural reporter groups and also the influences of α(1→2) linked Man residues, as well as the manifestation of the second branching point which is a Man residue substituted in the same way as Man-3.
The 360-MHz ^{1}H-NMR spectral features of the trisaccharide

Manα(1→3)Manβ(1→4)GlcNAc

have been published earlier (Ref. 3). The NMR data of this oligosaccharide, compound 20, refined by 500-MHz ^{1}H-NMR spectroscopy, are compiled in Table 6.
Extension of compound 20 with a Man residue α(1→2) linked to Man-4 affords compound 21. The structure of this tetrasaccharide and its 500-MHz ^{1}H-NMR spectrum are presented in Fig. 17. The NMR parameters of compound 21 are summarized in Table 6. The structural reporter groups, *i.e.*, H-1 and H-2, of the α(1→2) linked Man, refered to as Man-C, give rise to relatively sharp lined signals at δ = 5.050 ppm and δ = 4.069 ppm, respectively. This set of chemical shift values is typical for a terminal α(1→2) linked Man residue. The structural reporter

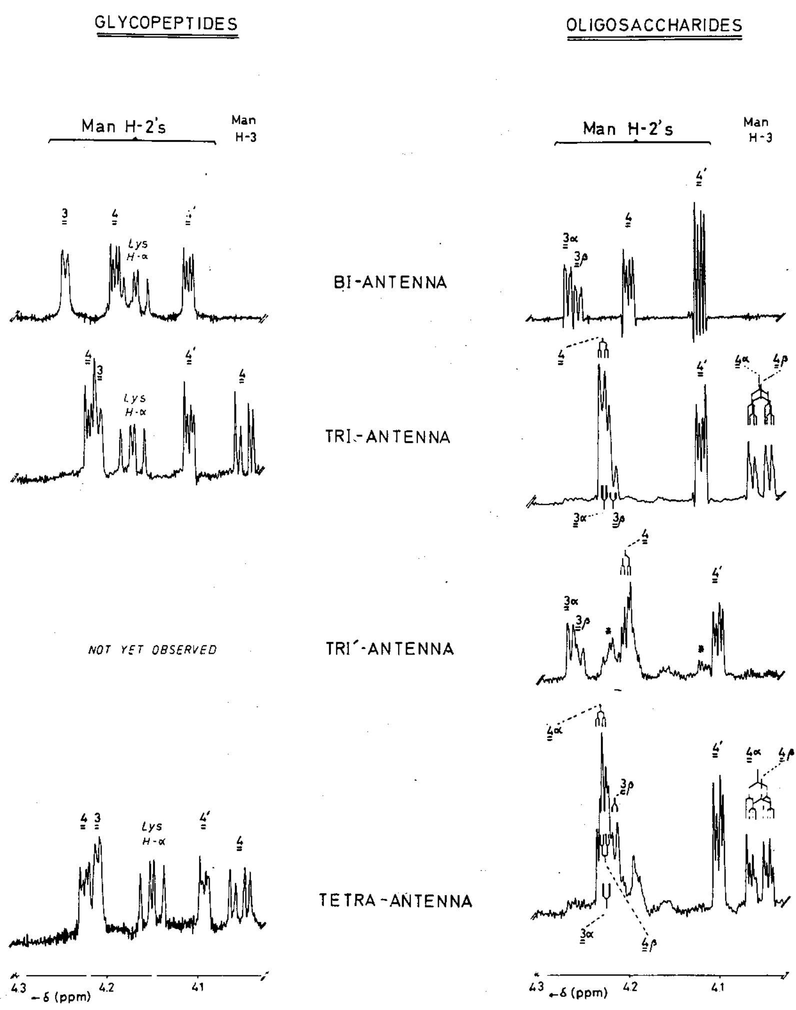


Fig. 16. Characteristic resonance patterns of the Man H-2's in the resolution enhanced 500-MHz ^{1}H-NMR spectra for bi-, tri-, tri'- and tetra-antennary *N*-acetyllactosamine type carbohydrate chains. Arising of the H-3 signal of Man-4 in this region of the spectrum points to the presence of the *N*-acetyllactosamine unit β(1→4) linked to this Man residue (see text). The tri'-antennary oligosaccharide sample is contaminated with a small amount of normal tri-antennary structure; the Man H-2 resonances of the latter one are indicated by asterisks.

group signals of Man-4 undergo downfield shifts upon the attachment of Man-C to Man-4 at C-2 (see Table 6). Furthermore, they are considerably broadened. It should be emphasized that the chemical shifts of the *N*-acetyl singlets of GlcNAc-2 in α- and β-form of this tetrasaccharide are characteristic for oligosaccharides containing a mono-α(1→3) substituted Man-3 (*cf.* the difference in chemical shift between the *N*-acetyl signals of GlcNAc-2 for upper and lower branch mono-antennary glycopeptides: compounds 14 and 16 *vs.* 15, see Table 5).
The structure of compound 22, being a glyco-asparagine possessing the second branching point, together with its 500-MHz ^{1}H-NMR spectrum is presented in Fig. 18. Its NMR data are given in Table 6. For the interpretation of the spectrum the NMR data of compound 13 (Table 5) have been utilized as a reference. The structural reporter group signals of the terminal α(1→6) linked Man-B occupy the same positions as those stemming from Man-4' in compound 13 (Fig. 13). Owing to the attachment of (Man-A in α(1→3)) and Man-B in α(1→6) linkage to Man-4', H-1 of the latter residue undergoes an upfield shift. A similar shift is observed going from a tri- to a tetra-antennary *N*-acetyllactosamine structure upon introduction of the fourth branch at Man-4' (*cf.* Fig. 1)(Ref. 18). Taken into consideration that an α(1→3) substitution of Man-4'

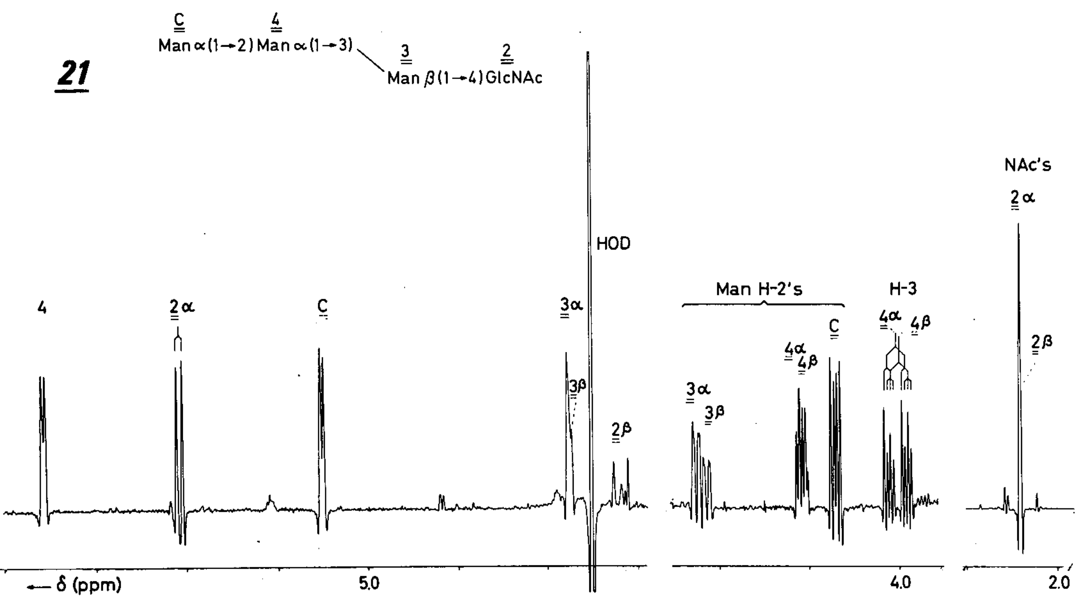


Fig. 17. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 21, together with its structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

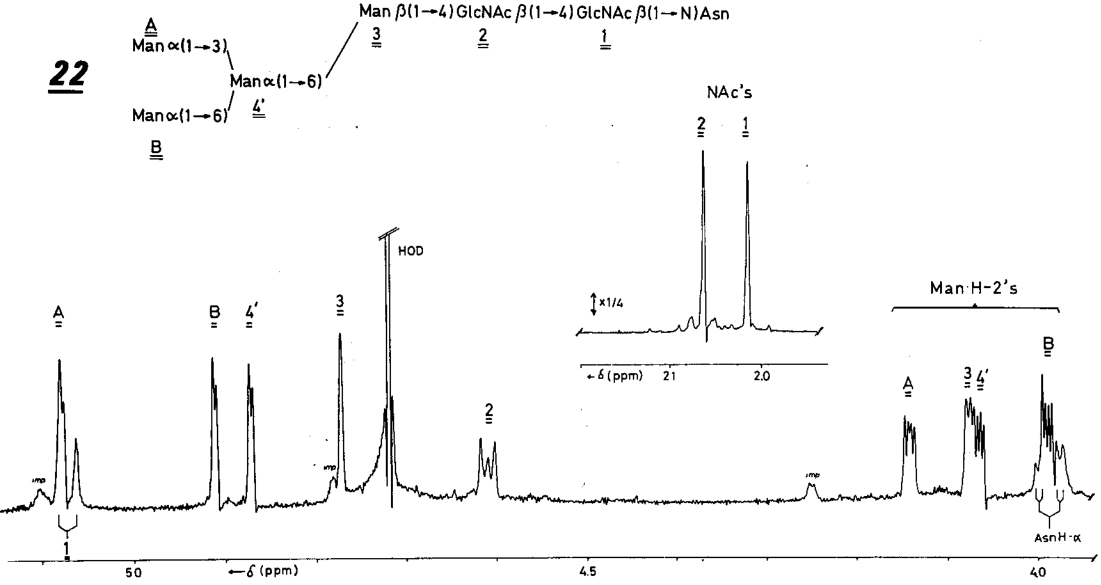


Fig. 18. Structural reporter group regions of the resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 22, together with its structure. The relative intensity scale of the insertion with the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

TABLE 6. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for mono-antennary oligomannoside type oligosaccharides (compounds 20 and 21) and glycopeptide (compound 22) [a]

reporter group	residue	oligosaccharide's anomer	compound 20	compound 21	residue	compound 22
H-1 of					1	5.069
	2	α	5.209	5.206	2	4.608
		β	≈4.72 [b]	4.718		
	3	α	4.787	4.776	3	4.770
		β	4.783	4.772		
	4	α,β	5.111	5.356		
	C	α,β	-	5.050		
					4'	4.870
					A	5.076
					B	4.909
H-2 of	3	α	4.244	4.224	3	4.076
		β	4.233	4.213		
	4	α	4.075	4.108		
		β	4.071	4.105		
	C	α,β	-	4.069		
					4'	4.140
					A	4.064
					B	3.988
NAc of					1	2.013
	2	α	2.043	2.042	2	2.061
		β	2.041	2.041		

[a] For complete structures and coding of monosaccharide residues see Figs. 17 and 18.
[b] Value can not be determined more accurately (± 0.01 ppm) due to interference of the HOD-line in the spectrum at 300 K.

by Man-A hardly influences the chemical shift of H-1 of Man-4' (*cf.* compounds 13 and 20, Tables 5 and 6), the abovementioned shift seems to be typical for a substitution of Man-4' at C-6. The H-2 of Man-4' resonates at δ = 4.140 ppm due to disubstitution of this residue. The remaining new signals in the spectrum, in comparison to that of compound 13, at δ = 5.076 ppm and δ = 4.064 ppm, belong to H-1 and H-2 of the terminal α(1→3) linked Man-A, respectively.
The results obtained so far demonstrate that high-resolution ^{1}H-NMR spectroscopy is excellently suited for the recognition of *N*-glycosidically linked carbohydrate chains of the *N*-acetyllactosamine type as well as of the oligomannoside type. Therefore it can be expected that also chains of the hybrid type can be completely identified by this technique.

Mucin type oligosaccharide-alditols derived from *O*-glycoproteins

The family of *O*-glycosidically linked carbohydrate chains of glycoproteins comprises a large variety of structures. Recently we had the opportunity to investigate several representatives of this type of compounds by 500-MHz ^{1}H-NMR spectroscopy. A few examples of oligosaccharide--alditols derived from the mucin type of *O*-glycoproteins will be given.
A simple element often occurring in these carbohydrate chains is compound 23. Its structure and resolution enhanced 500-MHz ^{1}H-NMR spectrum are shown in Fig. 19. The chemical shifts of the structural reporter groups of compound 23 are summarized in Table 7. It should be noted

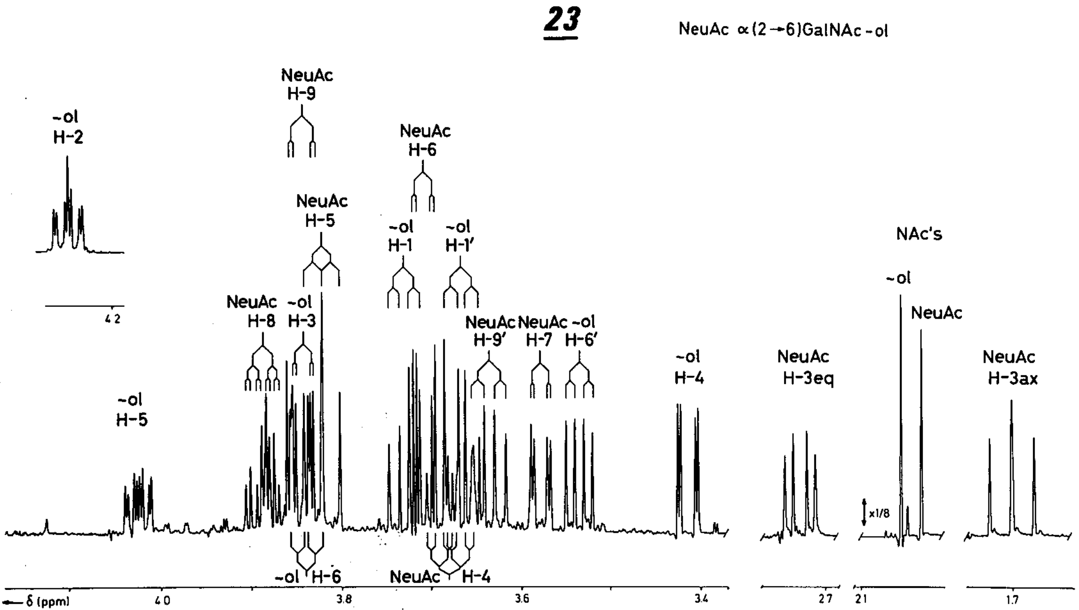


Fig. 19. Structure and in full detail assigned resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 23. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

that the chemical shift values of the H-3's and the *N*-acetyl methyl protons of NeuAc together represent an own set, typical for the α(2→6) linkage of the residue to GalNAc-ol, bearing no other substituents. The chemical shifts of the structural reporter groups of GalNAc-ol, *i.e.*, its H-2, H-3, H-4, H-5 and H-6', reflect the substitution pattern of this reduced monosaccharide (Ref. 1 & 17). The strong magnetic field of a 500-MHz NMR spectrometer gives rise to a nearly first order spectrum of the disaccharide-alditol 23, that could be easily interpreted in full detail (see Fig. 19).

Neuraminic acid bearing an *N*-glycolyl group in stead of an *N*-acetyl group frequently occurs in the carbohydrate chains of mucin glycoproteins. The 360-MHz ^{1}H-NMR spectrum of a trisaccharide-alditol containing NeuGl (compound 24) has been published previously (Ref. 26). The structure and 500-MHz ^{1}H-NMR spectrum of compound 24 are presented in Fig. 20. Relevant NMR parameters of this oligosaccharide-alditol are given in Table 7. An impressive improvement in spectral resolution could be obtained; *e.g.*, the patterns of the H-2 and H-5 resonances of GalNAc-ol were similar at 360 MHz, but are now strikingly different. The H-1 doublet of the terminal β-linked Gal residue is found at δ = 4.477 ppm, which is in accordance with the resonance positions of the corresponding protons in the NMR spectrum of compound 1 (Table 1). However, this is not necessarily the case for all β-linked terminal Gal residues (unpublished results).

Obviously substitution at Gal affects more or less the parameters of its H-1 doublet as illustrated in the 500-MHz ^{1}H-NMR spectrum of compound 25 (Fig. 21). The chemical shifts of the structural reporter groups of the bisialo tetrasaccharide-alditol 25 are compiled in Table 7. Recently the 360-MHz ^{1}H-NMR features of this compound have been described (Ref. 17). The 500-MHz ^{1}H-NMR spectrum shows *i.a.* a more detailed picture of the different H-3 resonances of the two NeuAc residues. The attachment of NeuAc in α(2→3) linkage to Gal causes a downfield shift for the H-1 and also for the H-3 signal of Gal (*cf.* Ref. 5). In contrast to the situation in case of a β(1→3) substitution at Gal (Ref. 1), the resonance position of its H-4 signal is only slightly altered in the step from 24 to 25 (see Table 7).

In *O*-glycosidically linked carbohydrate chains Gal may also be substituted at C-4, like in compound 26. The structure and 500-MHz ^{1}H-NMR spectrum of the trisaccharide-alditol 26 are presented in Fig 22. The NMR parameters of this compound are given in Table 7. The detailed NMR analysis of compound 26 and of other α-linked GlcNAc residues containing

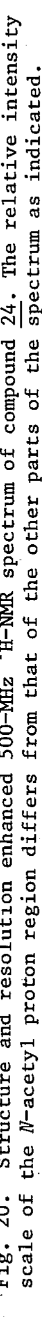


Fig. 20. Structure and resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 24. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

TABLE 7. ^{1}H Chemical shifts of structural reporter groups of constituent monosaccharides for mucin type oligosaccharide-alditols (compounds 23 - 26) [a]

reporter group	residue	compound 23	compound 24	compound 25	compound 26
H-2	of GalNAc-ol	4.245	4.380	4.378	4.404
H-3		3.842	4.061	4.067	4.081
H-4		3.413	3.541	3.524	3.541
H-5		4.024	4.249	4.240	4.193
H-6'		3.528	3.497	3.475	3.663
NAc		2.056	2.049	2.042	2.058
H-1	of Gal	-	4.477	4.541	4.525
H-3		-	3.669	4.117	3.749
H-4		-	3.899	3.927	3.971
H-1	of GlcNAc	-	-	-	4.869
H-4		-	-	-	3.542
H-5		-	-	-	4.183
NAc		-	-	-	2.089
H-3ax	of Neu$\frac{Ac}{Gl}\alpha(2\rightarrow6)$	1.700	1.711	1.692	-
H-3eq		2.728	2.746	2.723	-
NAc		2.033	-	2.032	-
NGl		-	4.123	-	-
H-3ax	of NeuAc$\alpha(2\rightarrow3)$	-	-	1.800	-
H-3eq		-	-	2.774	-
NAc		-	-	2.032	-

[a] For complete structures see Figs. 19, 20, 21 and 22.

oligosaccharide-alditols will be published elsewhere. A new structural reporter group is H-1 of GlcNAc which resonates at δ= 4.869 ppm, its coupling constant $J_{1,2}$ (= 4.0 Hz) being indicative of an α-glycosidic linkage. The H-5 signal of this GlcNAc residue can be clearly recognized in the spectrum, although it is interwoven with that of H-5 of GalNAc-ol. The introduction of GlcNAc $\alpha(1\rightarrow4)$ linked to Gal causes typical shift increments of the structural reporter group resonances of Gal in comparison to compound 24. In particular the position of the H-4 virtual doublet of Gal reflects this substitution.

The application of 500-MHz ^{1}H-NMR spectroscopy to relatively simple representatives of the *O*-glycosidically linked type of carbohydrate chains has opened the field for the extension to more complex structures.

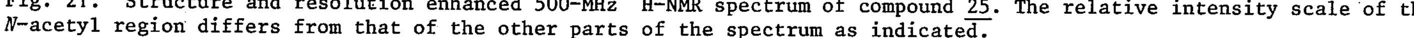


Fig. 21. Structure and resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 25. The relative intensity scale of the *N*-acetyl region differs from that of the other parts of the spectrum as indicated.

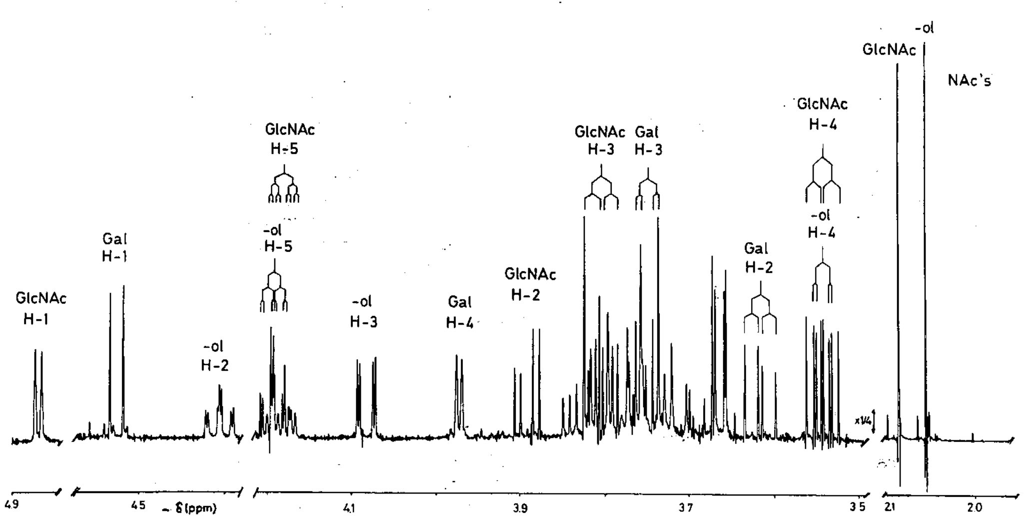


Fig. 22. Structure and resolution enhanced 500-MHz ^{1}H-NMR spectrum of compound 26. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

CONCLUDING REMARKS

1. 500-MHz ^{1}H-NMR spectroscopy of samples of glycopeptides, oligosaccharides and oligosaccharide-alditols derived from glycoproteins turned out to be far superior to earlier NMR investigations at lower magnetic field strengths. Owing to the cooperativity of the stronger magnetic field and the computer resolution enhancement routine, a more precise determination of the chemical shifts of structural reporter groups could be obtained and more details of the splitting patterns of their signals were discovered. The better insight in the spectral dispersion made the assignment of more structural reporter groups possible.
2. The analysis of reducing oligosaccharides showed that the anomeric configuration of the reducing end sugar also exerts its influence on the spectral parameters of residues in its spatial neighbourhood, being sometimes even the non-reducing end sugars. Obviously the NMR spectrum is a superposition of the subspectra of the different anomeric forms of the oligosaccharide.
3. The gain in spectral resolution as well as in sensitivity enabled the structure elucidation of minor constituents in samples, which were earlier considered as unidentifiable contaminants. In other words 500-MHz ^{1}H-NMR spectroscopy can disclose (micro)heterogeneity in samples which can hardly be traced along other routes.
4. A further application can be found in the study of enzyme specificities. This is highly relevant for the unraveling of the sequence of sugar attachment in the biosynthesis of glycoconjugates (*cf.* Ref. 15). On the other hand product analysis of oligosaccharides and glycopeptides, excreted by patients with inborn errors of metabolism, might give a clue to enzyme deficiencies in the catabolism of glycoconjugates of these patients.
5. This study is only the first exploration of the potencies of 500-MHz ^{1}H-NMR spectroscopy. In conjunction with modern pulsed NMR techniques it might be expected that besides new primary structure information, insight can be obtained in the spatial arrangement and conformational freedom of the residues in the carbohydrate chain.

EXPERIMENTAL

Samples of carbohydrate structures, dealt with in this paper, stem from a wide variety of glycoproteins, and were isolated by several research groups (Prof. J. Montreuil, Lille, France; Prof. K. Schmid, Boston, USA; Prof. P. Jollès, Paris, France; Prof. V. Derevitskaya, Moscow, USSR and Prof. D. Aminoff, Ann Arbor, USA) and generously supplied for NMR investigation.
Solutions of glycopeptides, oligosaccharides or oligosaccharide-alditols in D_2O were adjusted to pD ≃ 7, if necessary. Deuterium-exchanged samples were prepared by a five-fold lyophilization of the solution, finally using 99.96 atom% deuterated D_2O (Aldrich). For NMR spectral analysis, in general 0.1 to 3.0 mM solutions of the compounds in 0.4 mL D_2O were used.
The 500-MHz ^{1}H-NMR spectra were recorded on a Bruker WM-500 spectrometer, operating in the pulsed Fourier transform mode, and equipped with a Bruker Aspect 2000 computer with 32k memory capacity. The D-resonance of D_2O was used as field frequency lock signal. The spectra were taken up in 16k memory with an acquisition time of 3.28 sec and a spectral width of 2 times 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 hundreds of acquisitions for each sample were accumulated. The indicated probe temperature was 300 K and was kept constant within 0.1 K. At this temperature the HOD-resonance is found at δ = 4.752 ppm. The chemical shifts (δ) are expressed in ppm downfield from external sodium-2,2-dimethyl-2-silapentane-5-sulfonate (DSS), but were actually measured by reference to internal acetone (δ = 2.225 ppm) with an accuracy of 0.001 ppm.

Abbreviations

NMR - nuclear magnetic resonance
Asn - L-asparagine
Gly - glycine
Lys - L-lysine
Thr - L-threonine
Fuc - L-fucose
Gal - D-galactose
Man - D-mannose
GlcNAc - *N*-acetyl-D-glucosamine
GalNAc - *N*-acetyl-D-galactosamine
GalNAc-ol - *N*-acetyl-D-galactosaminitol
NeuAc - *N*-acetyl-D-neuraminic acid
NeuGl - *N*-glycolyl-D-neuraminic acid

Acknowledgements

The authors are much indebted to Profs. J. Montreuil, G. Strecker, B. Fournet, G. Spik (Lille, France), K. Schmid (Boston, USA), P. Jollès (Paris, France), V. Derevitskaya (Moscow, USSR), D.H. van den Eijnden (Amsterdam, The Netherlands), D. Aminoff (Ann Arbor, USA) and their coworkers for fruitful collaborations. The authors wish to express their gratitude to Bruker Analytische Messtechnik GmbH, Rheinstetten, GFR and Netherlands Bruker Spectrospin NV, Wormer, The Netherlands for making the 500-MHz facility in Rheinstetten, GFR accessible for their studies, and in particular to Dr. W.E. Hull for recording the spectra.
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 the Netherlands Foundation for Cancer Research (KWF), grant UUKC-OC 79-13.

REFERENCES

1. J.F.G. Vliegenthart, H. van Halbeek and L. Dorland, IUPAC 27th International Congress of Pure and Applied Chemistry, Helsinki 1979 (Ed. A. Varmavuori), pp. 253-262, Pergamon Press, Oxford (1980).
2. L. Dorland, B.L. Schut, J.F.G. Vliegenthart, G. Strecker, B. Fournet, G. Spik and J. Montreuil, Eur. J. Biochem. 73, 93-97 (1977).
3. L. Dorland, J. Haverkamp, B.L. Schut, J.F.G. Vliegenthart, G. Spik, G. Strecker, B. Fournet and J. Montreuil, FEBS Lett. 77, 15-22 (1977).
4. 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, 165-171 (1977).
5. L. Dorland, J. Haverkamp, J.F.G. Vliegenthart, G. Strecker, J.-C. Michalski, B. Fournet, G. Spik and J. Montreuil, Eur. J. Biochem. 87, 323-329 (1978).
6. G. Strecker, B. Fournet, J. Montreuil, L. Dorland, J. Haverkamp, J.F.G. Vliegenthart and D. Dubesset, Biochimie 60, 725-734 (1978).
7. D. Leger, V. Tordera, G. Spik, L. Dorland, J. Haverkamp and J.F.G. Vliegenthart, FEBS Lett. 93, 255-260 (1978).
8. B. Fournet, J. Montreuil, G. Strecker, L. Dorland, J. Haverkamp, J.F.G. Vliegenthart, J.P. Binette and K. Schmid, Biochemistry 17, 5206-5214 (1978).
9. 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, 149-152 (1978).
10. B.L. Schut, L. Dorland, J. Haverkamp, J.F.G. Vliegenthart and B. Fournet, Biochem. Biophys. Res. Commun. 82, 1223-1228 (1978).

11. L. Dorland, J. Haverkamp, J.F.G. Vliegenthart, G. Spik, B. Fournet and J. Montreuil, Eur. J. Biochem. 100, 569-574 (1979).
12. K. Schmid, J.P. Binette, L. Dorland, J.F.G. Vliegenthart, B. Fournet and J. Montreuil, Biochim. Biophys. Acta 581, 356-359 (1979).
13. J. Montreuil and J.F.G. Vliegenthart, Glycoconjugate Research, Proc. IVth Int. Symp. Glycoconjugates, Woods Hole 1977 (Eds. J.D. Gregory and R.W. Jeanloz), vol. I, pp. 35-78, Academic Press, New York (1979).
14. J.F.G. Vliegenthart, Structure and Function of Gangliosides (Eds. L. Svennerholm, H. Dreyfus and P.F. Urban), pp. 77-91, Plenum Publishing Corporation, New York (1980).
15. 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, 839-845 (1980).
16. L. Dorland and J.F.G. Vliegenthart, Biophysics of Membranes and Intercellular Communications (Eds. R. Balian et al.), North Holland Publishing Company, Amsterdam (1980), in press.
17. H. van Halbeek, L. Dorland, J.F.G. Vliegenthart, A.-M. Fiat and P. Jollès, Biochim. Biophys. Acta (1980), in press.
18. H. van Halbeek, L. Dorland, J.F.G. Vliegenthart, K. Schmid, J. Montreuil, B. Fournet and W.E. Hull, FEBS Lett. 114, 11-16 (1980).
19. J. Montreuil, Pure Appl. Chem. 42, 431-477 (1975).
20. C.J. Waechter and W.J. Lennarz, Ann. Rev. Biochem. 45, 95-112 (1976).
21. R. Kornfeld and S. Kornfeld, Ann. Rev. Biochem. 45, 217-237 (1976).
22. P.W. Robbins, S.J. Turco, S.C. Hubbard, D. Wirth and T. Liu, Glycoconjugate Research, Proc. IVth Int. Symp. Glycoconjugates, Woods Hole 1977 (Eds. J.D. Gregory and R.W. Jeanloz), vol. I, pp. 441-458, Academic Press, New York (1979).
23. I. Tabas and S. Kornfeld, J. Biol. Chem. 254, 11655-11663 (1979).
24. G. Strecker and J. Montreuil, Biochimie 61, 1199-1246 (1979).
25. J. Haverkamp, L. Dorland, J.F.G. Vliegenthart, J. Montreuil and R. Schauer, Abstracts IXth Int. Symp. Carbohyd. Chem. London, D7, 281-282 (1978).
26. L. Dorland, H. van Halbeek, J. Haverkamp, G.A. Veldink, J.F.G. Vliegenthart, B. Fournet, J. Montreuil and D. Aminoff, Glycoconjugates, Proc. Vth Int. Symp., Kiel 1979, (Eds. R. Schauer et al.), pp. 29-30, Thieme, Stuttgart (1979).