Primary structure of a novel *N*-glycosidic carbohydrate unit, derived from hen ovomucoid

A 500-MHz ^{1}H-NMR study

J. Paz-Parente, G. Strecker, Y. Leroy, J. Montreuil, B. Fournet*, H. van Halbeek, L. Dorland and J.F.G. Vliegenthart

Laboratoire de Chimie Biologique, Université des Sciences et Techniques de Lille I and Laboratoire Associé au CNRS no. 217, F-59655 Villeneuve d'Ascq Cédex, France and *Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, NL-3522 AD Utrecht, The Netherlands*

Received 30 December 1982

The *N*-glycosidic carbohydrate chains of hen *β*-ovomucoid were released from the protein by hydrazinolysis, and separated by HPLC. Primary structural analysis of 3 major fractions was conducted by applying 500-MHz ^{1}H-NMR spectroscopy in combination with methylation analysis. One of the fractions investigated appeared to consist of an intersected penta-antennary structure extended with one Gal residue. The location of the latter in a certain branch could be established unambiguously by NMR. This structure is a novel member of the family of *N*-glycosidic carbohydrates of glycoproteins.

Ovomucoid *Carbohydrate structure* *(Micro)heterogeneity* *Hydrazinolysis* *HPLC* *^{1}H-NMR*

1. INTRODUCTION

In [1] we have described the separation by HPLC of 17 oligosaccharides which were released from hen ovomucoid by hydrazinolysis. Recently, the primary-structural analysis of oligosaccharide *11*, the most abundant one, was carried out by methylation analysis, partial acid hydrolysis and 500-MHz ^{1}H-NMR spectroscopy [2]. Combination of these techniques enabled to connect unambiguously the primary structure of oligosaccharide *11* to the novel type of pentaantennary *N*-glycosidic glycans discovered in turtle-dove ovomucoid [3,4].

Here, we report on the primary structure of oligosaccharides *1*, *7* and *14*, determined by application of 500-MHz ^{1}H-NMR spectroscopy in combination with methylation analysis.

2. MATERIALS AND METHODS

Hen ovomucoid was prepared as in [5]. After pronase digestion, the asialo glycopeptide designated '*β*-glycopeptide' was isolated as in [6]. Oligosaccharides were released from the *β*-glycopeptide by hydrazinolysis [7]. Subsequently they were re-*N*-acetylated [8] and reduced with $NaBH_4$. The resulting mixture of oligosaccharide alditols was fractionated by semi-preparative HPLC; 17 fractions were obtained [1]. Here, the oligosaccharides *1*, *7* and *14* are further investigated.

Quantitative carbohydrate analysis of the oligosaccharides was carried out by GLC after methanolysis and pertrifluoroacetylation [9].

* To whom correspondence should be addressed

Abbreviations: HPLC, high-pressure liquid chromatography; GLC, gas-liquid chromatography; GLC-MS, gas-liquid chromatography combined with mass spectrometer; NMR, nuclear magnetic resonance

Published by Elsevier Biomedical Press
00145793/83/0000–0000/$3.00

Permethylation of the oligosaccharides was performed according to [10] and the partially methylated monosaccharides were identifed by GLC–MS according to [11].

For ^{1}H-NMR analysis, the oligosaccharides were repeatedly exchanged in D_2O. ^{1}H-NMR spectroscopic analysis was performed on a Bruker WM-500 spectrometer (SON hf-NMR facility, Nijmegen) operating at 500 MHz in the Fourier transform mode at a probe temperature of 300 K [12]. Chemical shifts are given in ppm relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) (indirectly to acetone in D_2O: δ 2.225).

3. RESULTS

3.1. *Carbohydrate composition; permethylation studies*

The molar carbohydrate compositions of oligosaccharides *1*, *7* and *14* are reported in table 1. The number of Man residues for each oligosaccharide being set to 3, one residue of GlcNAc-ol is found in each of the 3 oligosaccharides; besides 1, 5 and 7, GlcNAc residues occur in oligosaccharides *1*, *7* and *14*, respectively. For the latter oligosaccharide, the same relatively high value for the ratio total GlcNAc:Man (2.6) is found, as for the glycopeptides derived from the turtle-dove ovomucoid [4] and oligosaccharide *11* from hen ovomucoid [2].

The molar ratios of the monosaccharide methyl ethers derived from the permethylated oligosaccharides *1*, *7* and *14* are compiled in table 2. The results for oligosaccharide *1* are in accord with a trimannosyl-di-*N*-acetylchitobiose core structure, as has been described also for quail ovomucoid [13]. For oligosaccharide *7*, methyl-2-mono-*O*-methyl mannoside is found, which suggests the occurrence of an intersecting GlcNAc residue (1⟶4)-linked to Man 3. (For numbering of sugar residues see fig.1–3). The same methyl ethers as for *7* were found previously for glycopeptides obtained from chicken ovotransferrin [14]. Methylation analysis of the oligosaccharide *14*, like that of oligosaccharide *11* [2] reveals the presence of two different mono-*O*-methyl-mannosides in equal amounts: methyl 2- and methyl 3-mono-*O*-methyl mannosides. The presence of the two mono-*O*-methyl ethers of mannose had been demonstrated already by methylation analysis of glycopeptides from turtle-dove ovomucoid [4] and of total hen ovomucoid [3,15]. These results indicate that one residue of mannose is substituted at positions 2 and 4, a second mannose at positions 3, 4 and 6 and the third mannose at positions 2, 4 and 6. The difference between oligosaccharides *11* [2] and *14* is the presence of one Gal residue in terminal non-reducing position in *14*, witness the presence of methyl-2,3,4,6-tetra-*O*-methyl galactoside (1⟶4)-linked to one of the GlcNAc residues. No indications useful for localization of the Gal residue in a certain branch can be derived from sugar and methylation analysis.

3.2. *500-MHz ^{1}H-NMR spectroscopy*

To elucidate the primary structures of the hen ovomucoid oligosaccharides *1*, *7* and *14*, 500-MHz ^{1}H-NMR spectra of the compounds in D_2O were recorded. Relevant NMR parameters for the 3 fractions are listed in table 3; for reference purposes, those for fraction *11* have been included (*cf.* [2]). The structural-reporter-group regions of the

Table 1

Carbohydrate composition of oligosaccharide *1*, *7* and *14* obtained by semi-preparative HPLC of glycans liberated from hen ovomucoid neutral-glycopeptide by hydrazinolysis

Oligosaccharide	Molar ratio[a] of				GlcNAc + GlcNAc-ol ratio
	Gal	Man	GlcNAc	GlcNAc-ol	Man
1	0	3	1.30	0.94	0.74
7	0	3	4.64	0.94	1.86
14	1.07	3	6.80	1.07	2.62

[a] Man taken as 3

Table 2

Molar ratios of monosaccharide methyl ethers present in the methanolysates of the permethylated oligosaccharides *1*, *7* and *14*

Oligosaccharide	Partially methylated monosaccharide (mol/mol) of oligosaccharides)									
	2,3,4,6-tetra-OMe-Gal[a]	2,3,4,6-tetra-OMe-Man	3,4,6 tri-OMe-Man	2,4-di-OMe-Man	3,6-di-OMe-Man	2-Mono-OMe-Man	3-Mono-OMe-Man	3,4,6-tri-OMe-Glc NAcNMe[b]	3,6-di OMe-Glc NAcNMe	1,3,5,6-tetra-OMeGlc-ol NAcNMe
1	–	1.90	–	1	–	–	–	–	0.91	0.95
7	–	–	1.1	–	0.87	1	–	3.88	0.9	0.98
14	0.9	–	–	–	0.92	1.1	1	4.82	1.9	0.91

[a] *O*-Methyl is abbreviated as OMe

[b] *N*-acetyl-*N*-methyl is abbreviated as NAcNMe

500-MHz ^{1}H-NMR spectrum of *14*, as a typical example, are depicted in fig.1.

Comparison of the ^{1}H-NMR data for compounds *1*, *7*, *11* and *14* reveals that all 4 of them are reduced oligosaccharides having in common the GlcNAc $\beta(1\longrightarrow4)$-GlcNAc-ol structural element. The spectral features that are characteristic for the presence of this unit are the H-2 signal of GlcNAc-1-ol at $\delta = 4.25$ (seemingly, a broadlined quartet in the Man H-2 region of the spectrum, $4.0 < \delta < 4.3$, see fig.1), the H-1 doublet of GlcNAc-2 at $\delta \simeq 4.63$ (this chemical shift value, in combination with $J_{1,2}$ being 7.85 Hz, points to the $\beta(1\longrightarrow4)$-linkage of GlcNAc-2 to reduced GlcNAc-1; see also [2]), and two *N*-acetyl methyl singlets at $\delta \simeq 2.055$ and $\simeq 2.08$ for GlcNAc-1-ol and GlcNAc-2, respectively. Furthermore, all 4 compounds contain the mannotriose branching core that is usually found in *N*-glycosidic carbohydrate chains. Evidence for this stems from the occurrence of 3 Man H-1 signals, and also of 3 Man H-2 signals, in each of the spectra (see, for example, fig.1). The characteristic shapes of these signals point to a β-glycosidic linkage for one of the mannoses (designated Man-3) and to α-glycosidic linkages for the other two (designated 4 and 4′) [16].

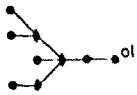
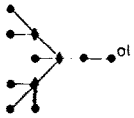
Compound *1* is an incomplete diantennary structure that possesses both Man-4 and Man-4′ in terminal, non-reducing position. This can be concluded from comparison of the H-1 and H-2 chemical shifts of these residues in *1* (see table 3) with those for terminal Man-4 and Man-4′ residues in the oligomannoside-type glyco-asparagines M_2GP and M_5GP [17], or in oligosaccharides like M_2G from mannosidosis urine [18], Ala and Alb from human meconium [19], and *2a* and *2b* from urine of patients with Morquio syndrome type B [20]. Therefore, the structure of compound *1* appears to be Man $\alpha(1\longrightarrow3)$[Man $\alpha(1\longrightarrow6)$]Man $\beta(1\longrightarrow4)$-GlcNAc$\beta(1\longrightarrow4)$GlcNAc-ol.

The resonance positions of the Man-3, -4 and -4′ H-1 and H-2 atoms in the spectrum of compound *7* indicate the branching pattern to be intersected triantennary. This can be readily inferred from comparison of these shift data (see table 3) with those described for a glycopeptide derived from chicken ovotransferrin [14,16]. Moreover, the peripheral part of the structure of the ovotransferrin glycopeptide just mentioned and that of compound *7* from hen ovomucoid are identical. The correctness of this conclusion is proved by the perfect accord between the H-1 and also the NAc signals of the 4 terminal GlcNAc residues for these compounds, both as to their number, as well as to their chemical shifts.

The ^{1}H-NMR spectral features of the intersected pentaantennary structure *11* have been described in detail [2]. It should be mentioned that such a type of branching (i.e., a 2,4-disubstitution of Man-4 and a 2,4,6-trisubstitution of Man-4′ besides the presence of the intersecting GlcNAc-2) gives rise to a unique set of Man H-1 and H-2 chemical shifts. In addition, the Man-4 H-3 and Man-4′ H-4 signals occupy typical positions. Comparison of the corresponding Man H-1, H-2, H-3 and H-4

Table 3

^{1}H chemical shifts of structural-reporter groups of constituent monosaccharides for some major carbohydrates from hen β-ovomucoid[a]

Reporter group	Residue[b]	Compound and schematic structure[c]			
		1	*7*	*11*	*14*
					
H-1 of	2	4.637	4.632	4.626	4.626
	3	~4.78	4.696	4.712	4.712
	4	5.103	5.057	5.067	5.066
	4′	4.915	4.999	4.889	4.890
	5	–	4.540	4.539	4.538
	5′	–	4.543	4.545	4.545
	7	–	4.516	*4.517*	*4.538*
	7′	–	–	4.517[d]	4.518[d]
	8	–	–	–	4.471
	9	–	4.464	4.443	4.443
	10′	–	–	4.583[d]	4.586[d]
H-2 of	1-ol	4.244	4.246	4.255	4.254
	3	4.259	4.146	4.145	4.145
	4	4.067	4.284	4.276	4.277
	4′	3.974	4.146	4.161	4.158
H-3 of	4	n.d.	4.048	4.043	4.044
H-4 of	4′	n.d.	n.d.	4.194	4.194
NAc of	1-ol	2.055	2.055	2.054[e]	2.053[e]
	2	2.076	2.079	2.084	2.084
	5	–	2.055	2.065[e]	2.064[e]
	5′	–	2.048	2.045	2.045
	7	–	2.083	*2.084*	*2.079*
	7′	–	–	2.093[f]	2.093[f]
	9	–	2.064	2.065	2.064
	10′	–	–	2.122[f]	2.121[f]

[a] Chemical shifts are given at 300 K, in ppm downfield from internal DSS in D_2O

[b] For numbering of monosaccharide residues and complete structures, see fig.3

[c] Compounds are represented by schematic structures (*cf.* [17]); (●—) GlcNAc; (◆—) Man; (■—) Gal

[d,e,f] Assignments may have to be interchanged

n.d., value could not be determined

chemical shifts for compounds *7* and *11* (see table 3) reveals that those for Man-4′ deviate considerably, while those for Man-4 are essentially the same for both compounds. This led us to the conclusion [2] that Man-4′ in *11* is the trisubstituted Man residue, the occurrence of which was established by methylation analysis. Apart from the core GlcNAc signals, another seven GlcNAc

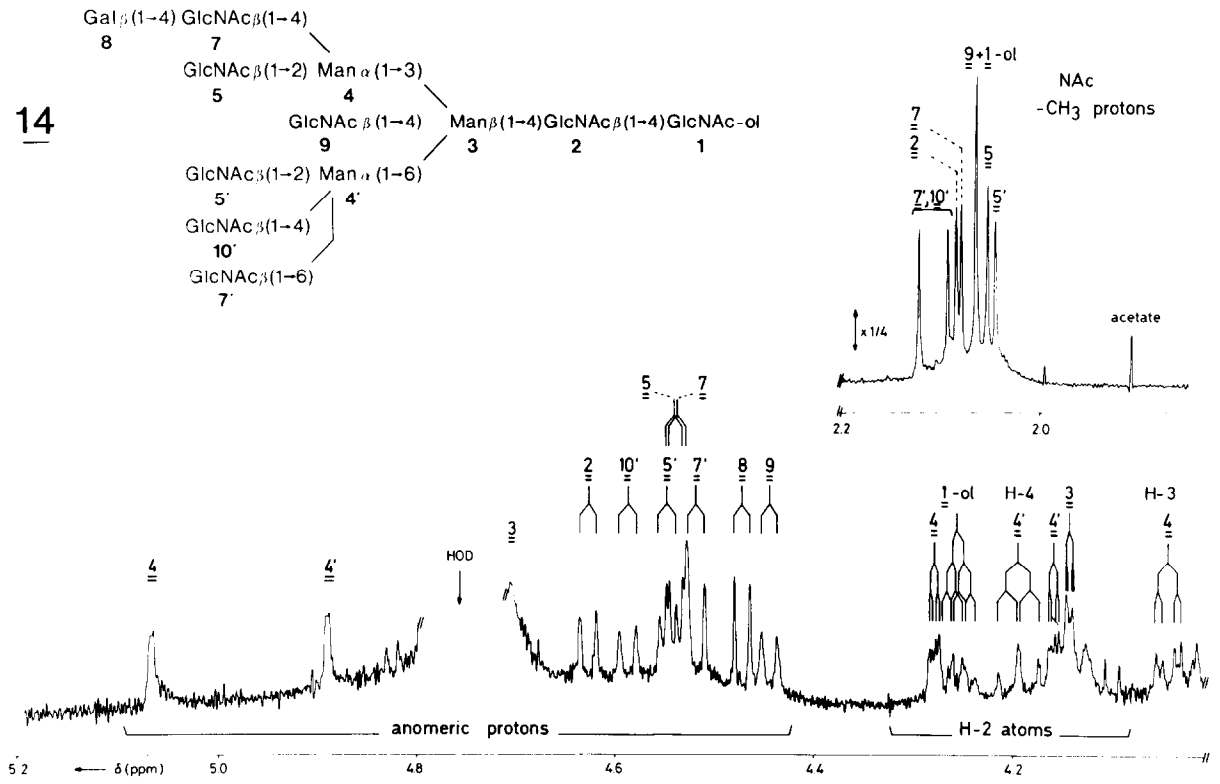


Fig.1. Structural reporter-group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectrum of oligosaccharide-alditol *14*, derived from hen ovomucoid, in D_2O at 300 K. The numbers in the spectrum refer to the corresponding residues in the structure. The relative-intensity scale of the *N*-acetyl proton region differs from that of the other part of the spectrum, as indicated.

H-1 doublets and NAc singlets are observed.

Methylation analysis of compound *14* (see table 2) suggested that the latter might differ from structure *11* only in the presence of a terminal Gal residue, (1⟶4)-linked to a GlcNAc. Once this sugestion is verified, the position of Gal in a certain branch remains to be established. In order to facilitate comparison between the 500-MHz ^{1}H-NMR spectra of fractions *11* and *14*, pertinent parts of both are presented in fig.2.

The branching pattern of the mannotriose core in *14* is indeed the same as in *11*, as can be readily inferred from the perfect accordance of the Man H-1 and also the H-2 chemical shifts (see also table 3). The resonance position of Man-4 H-3 (δ 4.044) corroborates the 2,4-disubstitution of this residue, that of Man-4′ H-4 (δ 4.194) confirms the 2,4,6-trisubstitution of the latter residue. From comparison of the β-anomeric region ($4.4 < \delta <$ 4.7) of both spectra (see fig.2), it is obvious that in the spectrum of *14* one additional doublet is present, at δ 4.471 ($J_{1,2}$ = 7.8 Hz). This combination of δ- and *J*-value is known to be characteristic for a terminal Gal residue β(1⟶4)-linked to GlcNAc [10,16]. Thus, compound *14* appears to be an extension of compound *11* with one Gal residue.

The attachment of this Gal residue to structure *11* causes two significant chemical-shift alterations of GlcNAc structural-reporter groups. First, one of the two H-1 doublets at δ 4.517 for *11* is shifted towards δ 4.538 for *14*, while the positions of all other anomeric doublets are unchanged, going from *11* to *14*. Secondly, one of the two *N*-acetyl singlets, coinciding at δ 2.084 for *11*, is shifted towards δ 2.079. The *N*-acetyl signals at δ 2.084 for *11* were unambiguously assigned to GlcNAc-2 and -7 (*cf.* compound *7*, table 3) [2]. The doublets coinciding at δ 4.517 for *11* belong to the H-1 atoms of

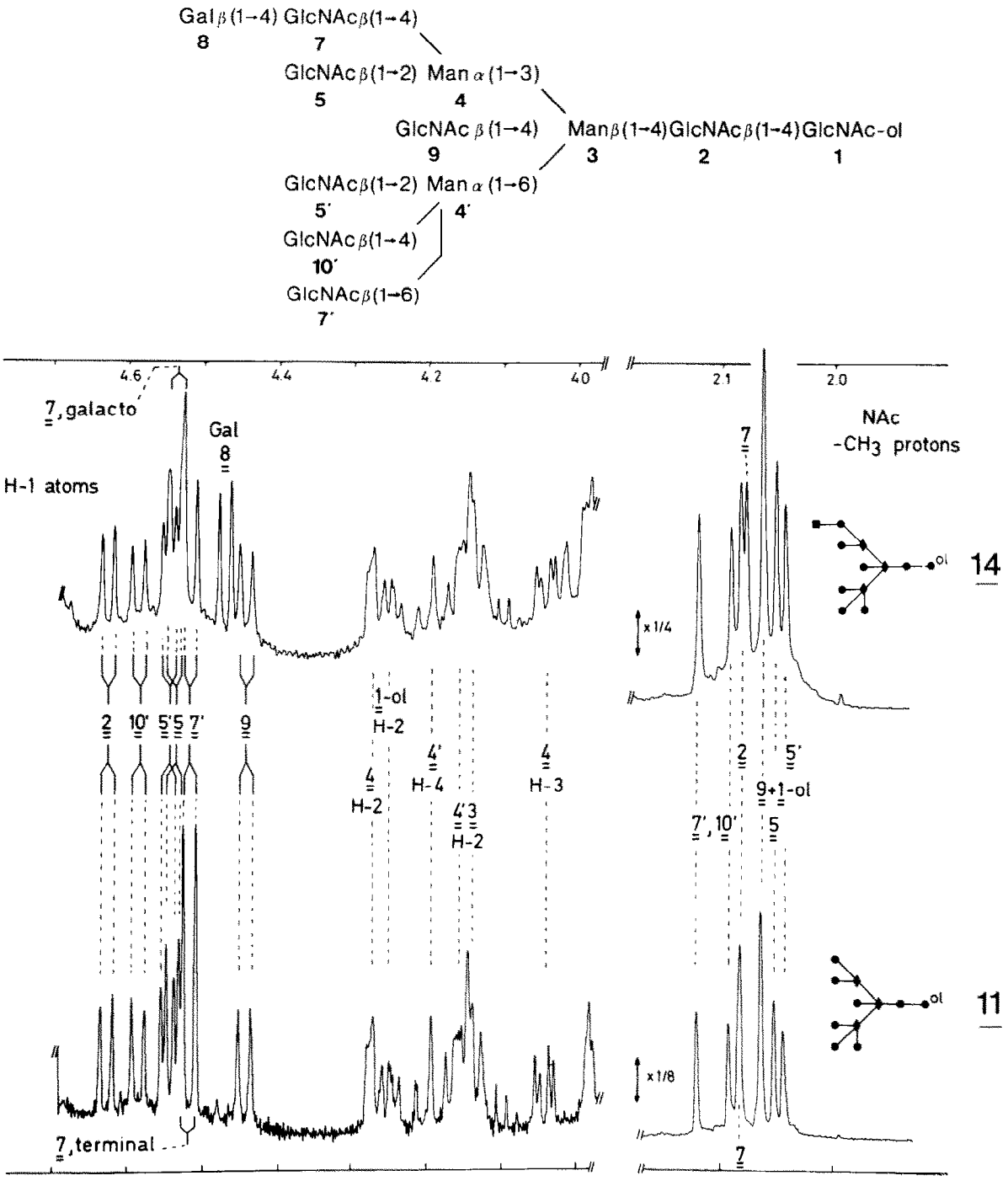


Fig.2. Pertinent structural reporter-group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectra of oligosaccharide-alditols *11* (lower trace) and *14* (upper trace) from hen ovomucoid. Compounds are represented by schematic structures; (●——) GlcNAc; (◆——) Man; (■——) Gal. The numbers in the spectra refer to corresponding residues in the structures (see fig.1 and 3). The relative-intensity scales of the *N*-acetyl proton regions differ from those of the other parts of the spectra, as indicated.

GlcNAc-7 and either of GlcNAc-7′ or of -10′. The observed shift increments (for H-1, $\Delta\delta \simeq 0.02$ ppm, and for NAc, $\Delta\delta \simeq -0.005$ ppm), if occurring simultaneously, are very typical for the attachment of Gal in $\beta(1\longrightarrow4)$-linkage to GlcNAc [16,21]. Since GlcNAc-2 is already substituted by Man-3 in $\beta(1\longrightarrow4)$-linkage, only GlcNAc-7 can be considered to be the GlcNAc residue that bears Gal in oligosaccharide *14*. Therefore, Gal in fraction *14* has been identified to be Gal-8 (see fig.1).

4. CONCLUSIONS

The structures which can be proposed on the basis of the results of 500-MHz ^{1}H-NMR spectroscopy in conjunction with methylation analysis for the oligosaccharides *1*, *7* and *14* from hen ovomucoid are summarized in fig.3. Oligosaccharide *1* corresponds to the core structure of *N*-glycosidic glycans of glycoprotein; the same primary structure has been isolated from quail

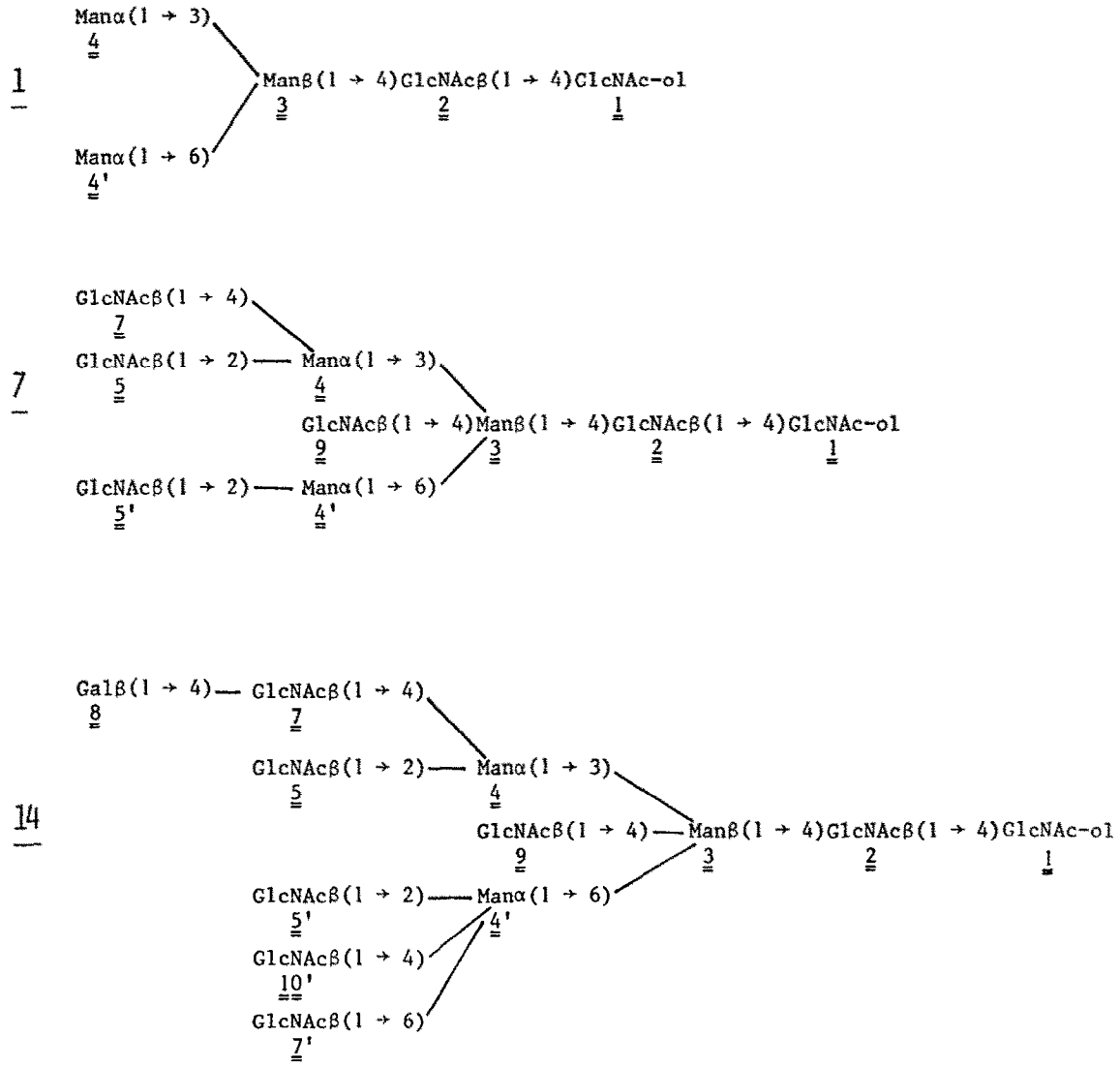


Fig.3. Structures of oligosaccharide-alditols *1*, *7* and *14*, isolated from hen ovomucoid.

ovomucoid [13]. Oligosaccharide *7* is essentially identical to the carbohydrate moiety of chicken ovotransferrin [14]. Oligosaccharide *14* possesses a pentaantennary structure with an intersecting GlcNAc residue and one Gal residue $\beta(1\longrightarrow4)$-linked to GlcNAc-$\underline{\underline{7}}$. It is an extension of oligosaccharide *11*, the NMR parameters of which were described in [2]. The subtle differences between the NMR parameters of the structural-reporter groups of GlcNAc $\underline{\underline{7}}$ for oligosaccharide *14* as compared to *11*, enabled to establish that Gal in *14* is linked to GlcNAc-$\underline{\underline{7}}$. It is remarkable that in hybrid type, *N*-glycosidic oligosaccharides containing a Gal residue, this Gal has been found to occur also in $\beta(1\longrightarrow4)$-linkage to GlcNAc-$\underline{\underline{7}}$ [22]. This novel pentaantennary structure was also demonstrated in turtle-dove ovomucoid [3,4].

ACKNOWLEDGEMENTS

This investigation was supported by the Centre National de la Recherche Scientifique (Laboratoire Associé no. 217); the Délégation Générale à la Recherche Scientifique et Technique (79.7.0669), 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 (Grant UUKC-OC 79-13). Thanks are due to M.G. Ricart for GLC-MS analysis and to B. Mahieu, G. Tinel and J. Celen for skillful technical assistance.

REFERENCES

[1] Paz-Parente, J., Strecker, G., Leroy, Y., Montreuil, J. and Fournet, B. (1982) J. Chromatogr. 249, 199–204.

[2] Paz-Parente, J., Wieruszeski, J.M., Strecker, G., Montreuil, J., Fournet, B., Van Halbeek, H., Dorland, L. and Vliegenthart, J.F.G. (1982) J. Biol. Chem. 257, 13173–13176.

[3] Montreuil, J., Fournet, B., Strecker, G., François-Gérard, Ch., Pierce-Crétel, A. and Spik, G. (1980) Proc. 13th FEBS-Meeting, Jerusalem, p.194.

[4] François-Gérard, Ch., Brocteur, J., André, A., Gerday, C., Pierce-Crétel, A., Montreuil, J. and Spik, G. (1980) Blood Transf. Immunohaemat. 23, 579–587.

[5] Fredericq, E. and Deutsch, H.F. (1949) J. Biol. Chem. 181, 499.

[6] Monsigny, M., Adam-Chosson, A. and Montreuil, J. (1968) Bull. Soc. Chim. Biol. 50, 857–874.

[7] Bayard, B. and Fournet, B. (1975) Carbohydr. Res. 46, 75–86.

[8] Reading, G.L., Penhoet, E. and Ballou, C. (1978) J. Biol. Chem. 253, 5600–5612.

[9] Zanetta, J.P., Breckenridge, W.C. and Vincendon, G. (1972) J. Chromatogr. 69, 291–304.

[10] Finne, J., Krusius, T. and Rauvala, H. (1980) Carbohyd. Res. 80, 336–339.

[11] Fournet, B., Strecker, G., Leroy, Y. and Montreuil, J. (1981) Anal. Biochem. 116, 486–502.

[12] Vliegenthart, J.F.G., Van Halbeek, H. and Dorland, L. (1981) Pure Appl. Chem. 53, 45–77.

[13] Hase, S., Okawa, K. and Ikenaka, T. (1982) J. Biochem. (Tokyo) 91, 735–737.

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

[15] Conchie, J., Hay, A.J. and Lomax, J.A. (1982) Carbohydr. Res. 103, 129–132.

[16] Vliegenthart, J.F.G., Dorland, L. and Van Halbeek, H. (1983) Adv. Carbohydr. Chem. Biochem. 41, in press.

[17] Van Halbeek, H., Dorland, L., Veldink, G.A., Vliegenthart, J.F.G., Michalski, J.-G., Montreuil, J., Strecker, G. and Hull, W.E. (1980) FEBS Lett. 121, 65–70.

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

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

[20] Michalski, J.-C., Strecker, G., Van Halbeek, H., Dorland, L. and Vliegenthart, J.F.G. (1982) Carbohydr. Res. 100, 351–363.

[21] Carver, J.P. and Grey, A.A. (1981) Biochemistry 20, 6607–6616.

[22] Yamashita, K., Tachibana, Y. and Kobata, A. (1978) J. Biol. Chem. 253, 3862–3869.