Volume 329, number 1,2, 29–34 FEBS 12853 August 1993

Structure of three acidic O-linked carbohydrate chains of porcine zona pellucida glycoproteins

Cornelis H. Hokke[a], Jan B.L. Damm[b], Johannis P. Kamerling[a] and Johannes F.G. Vliegenthart[a]

[a]*Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, Utrecht, The Netherlands* and [b]*Organon International BV, Oss, The Netherlands*

Received 22 June 1993

Structural analysis by 1D and 2D ^{1}H NMR spectroscopy of three acidic O-linked oligosaccharide alditols, released from porcine zona pellucida glycoproteins by alkaline borohydride treatment, afforded the following structures:

Galβ1-4(6SO$_4^-$)GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol
Neu5Gcα2-3Galβ1-4(6SO$_4^-$)GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol
Neu5Acα2-3Galβ1-4(6SO$_4^-$)GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol

These oligosaccharides are the smallest compounds that contain the structural elements which are present in the acidic, high-molecular mass O-linked carbohydrate chains of porcine zona pellucida glycoproteins.

Porcine zona pellucida glycoprotein; Sulfated *N*-acetyllactosamine; O-Linked carbohydrate chain

1. INTRODUCTION

The zona pellucida (ZP) is the glycoprotein matrix surrounding the mammalian oocyte in the early developmental stage. It is composed of several families of glycoproteins and mediates a number of critical steps in the fertilization process [1,2]. For various species, it has been demonstrated that the carbohydrate chains of the ZP plays an important role in the sperm recognition event [3–6]. So far, most studies dealing with the structure of ZP-derived carbohydrate chains have been carried out on porcine ZP (pZP), since porcine oocytes are relatively easy to obtain in sufficient quantities. The pZP comprises three glycoprotein families, pZP1–pZP3, each of which contain both N- and O-linked oligosaccharides [5,7–9]. It has been shown that the largest family, pZP3, possesses the boar sperm receptor activity [10], but contradictory reports have appeared concerning the question of whether the N- or the O-linked carbohydrate chains of pZP3 are involved in the interaction with the sperm [5,6].

To date, general structural aspects of pZP-linked carbohydrate chains are known, like the presence of sulfated (sialylated) *N*-acetyllactosamine repeats in both N- and O-linked oligosaccharides, resulting in extreme charge and size heterogeneity [5,8,9]. The structures of the neutral N-linked carbohydrate chains of unfractionated pZP glycoproteins [11] and purified pZP3 [6] have been established, and detailed structural features of the acidic *N*-glycans of pZP3 have recently been described [12]. However, no structural analysis of the acidic O-linked carbohydrate chains has been reported. Therefore, a study was initiated, aimed at the determination of the structure of pZP-derived O-linked oligosaccharides. In this paper, three acidic oligosaccharides of low molecular mass are presented, containing a unique combination of structural elements which seems to occur also in the larger pZP *O*-glycans.

Correspondence address: J.F.G. Vliegenthart, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, PO Box 80.075, NL-3508 TB Utrecht, The Netherlands. Fax: (31) (30) 540 980.

Abbreviations: FPLC, fast protein liquid chromatography; HOHAHA, homonuclear Hartmann-Hahn; HPLC, high-pressure liquid chromatography; NAc, *N*-acetyl; Neu5Ac, *N*-acetylneuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid; NGc, *N*-glycolyl; PNGase-F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F; pZP, porcine zona pellucida; ZP, zona pellucida; 1D, one dimensional; 2D, two dimensional.

2. MATERIALS AND METHODS

2.1. *Materials*

Porcine zona pellucida glycoproteins were a gift from Prof. J. Aitken (MRC, Edinburgh, UK). Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase-F) from *Flavobacterium meningosepticum* was obtained from Boehringer-Mannheim.

2.2. *Liberation and isolation of the carbohydrate chains*

The N-linked carbohydrate chains were released from 60 mg pZP glycoproteins and subsequently separated from the N-deglycosylated

glycoproteins as described [13], but using two sequential additions of PNGase-F (1 U per mg pZP) with 4 h of intermediate incubation. The N-deglycosylated glycoprotein was treated with alkaline borohydride as described [14] in order to release the O-linked carbohydrate chains. The reaction mixture was rinsed through an ExtractiGel-D column (8 × 1 cm, Pierce), using 25 mM NH_4HCO_3 as eluent. The run-through fraction was collected and desalted on a Bio-Gel P-2 column (18 × 1 cm, 200–400 mesh, Bio-Rad), eluted with water.

2.3. *Fractionation of the carbohydrate chains*

Charge-based fractionation of the O-linked carbohydrate chains was performed by FPLC on Mono Q [14], using a linear gradient of 0–1 M NaCl. The acidic Mono Q fraction A was further fractionated on a Bio-Gel P-4 column (150 × 1.15 cm, 200–400 mesh, Bio-Rad), eluted with 100 mM NH_4HCO_3 at a flow rate of 7 ml/h. The eluent was monitored at 206 nm and fractions of 1.2 ml were collected. Subfractionation of Bio-Gel P-4 fraction A1 was carried out on a column of Bio-Gel P-6 (135 × 2.2 cm, 200–400 mesh, Bio-Rad), eluted with 100 mM NH_4HCO_3 at a flow rate of 23 ml/h. The eluent was monitored at 206 nm and fractions of 5.8 ml were collected. Bio-Gel P-4 and P-6 subfractions were further purified on a column of Dowex 50W-X8, H^+ form (6 × 0.5 cm, 100–200 mesh, Fluka). The column was eluted with 6 ml 0.01 M formic acid and the eluate was lyophilized. Bio-Gel P-4 fraction A2 and Bio-Gel P-6 fraction A1.7 were further fractionated on Mono Q [14], using a discontinuous gradient of 0–0.5 M NaCl.

2.4. *500 MHz ^{1}H NMR spectroscopy*

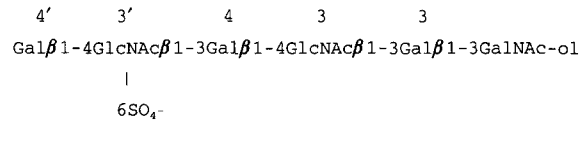
1D 500-MHz ^{1}H NMR spectra were recorded on a Bruker AMX-500 spectrometer as described [15]. For the 2D HOHAHA spectra [16] a MLEV-17 mixing sequence of 120 ms was used. The 90° pulse width was adjusted to about 27 μs and the spectral width was 3,000 Hz in both dimensions for fraction A2.3, and 3,500 Hz in both dimensions for fraction A1.7.4. The HO^2H signal was presaturated for 1 s during the relaxation delay. In total, 434 spectra of 2,048 data points with 80 scans per t_1 value for A2.3, and 430 spectra of 2048 data points with 88 scans per t_1 value for A1.7.4 were recorded. The 2D NMR data were processed as reported [15].

3. RESULTS

The O-linked oligosaccharide alditols released from pZP glycoproteins were subjected to FPLC on Mono Q, giving rise to a neutral fraction N and a non-resolved acidic fraction A. The ^{1}H NMR spectrum of fraction A (not shown) indicated the presence of a complex mixture of oligosaccharide alditols, containing poly-(*N*-acetyllactosamine) sequences, 6-*O*-sulfated GlcNAc residues, and sialic acid residues (Neu5Gc and Neu5Ac). This report is focused on the purification, and identification by 1D and 2D ^{1}H NMR spectroscopy, of three acidic oligosaccharide alditols of low molecular mass, which contain a combination of the above mentioned structural elements from which the larger acidic carbohydrate chains are also constructed. The investigated oligosaccharides were obtained and analyzed as follows.

Fraction A was applied to Bio-Gel P-4, giving rise to one broad unresolved peak. Three carbohydrate-containing subfractions, denoted A1–A3, were collected. The void-volume fraction, A1, containing the bulk of the material, was further fractionated on Bio-Gel P-6, yielding a set of poorly separated peaks. Seven subfractions, denoted A1.1–A1.7, were collected. Fractions A2 and A1.7 were each passed over Dowex H^+ and subsequently subfractionated on Mono Q. Three subfractions were collected for fraction A2 (denoted A2.1–A2.3) and five subfractions for A1.7 (denoted A1.7.1–A1.7.5). The ^{1}H NMR data of fractions A2.3 and A1.7.4, which are discussed below, are compiled in Table I.

The ^{1}H NMR spectrum of fraction A2.3 (Fig. 1A) reveals the presence of the following sulfated hexasaccharide alditol.



Fraction A2.3

The Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol element is readily deduced by comparing the relevant ^{1}H NMR data with those of the tetrasaccharide alditol Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol (cf. compound 12 in [17]). Additional anomeric signals are observed at δ 4.523 (Gal$^{4'}$ H-1) and δ 4.716 (GlcNAc$^{3'}$ H-1) suggesting the presence of an extra *N*-acetyllactosamine unit. 2D HOHAHA spectroscopy (Fig. 2A) demonstrated the interconnection of the GlcNAc$^{3'}$ H-1 signal with the H-6 and H-6′ resonances at δ 4.395 and δ 4.314, respectively. By consequence, GlcNAc$^{3'}$ is 6-O-sulfated [18,19]. Cross peaks are observed at the Gal$^{4'}$ H-1 track at δ 4.523 to the H-2 (δ 3.54), H-3 (δ 3.67) and H-4 (δ 3.926) signals, which show that this Gal residue is in the terminal position (cf. compounds 8 and 9 in [20]). The chemical shift difference ($\Delta\delta$ 0.043) between the Gal$^{4'}$ H-1 signal in A2.3 and the Gal H-1 signal in a non-sulfated terminal Galβ1-4GlcNAc element (cf. compound in 12 in [17]) is similar to that found for the Gal H-1 signals when going

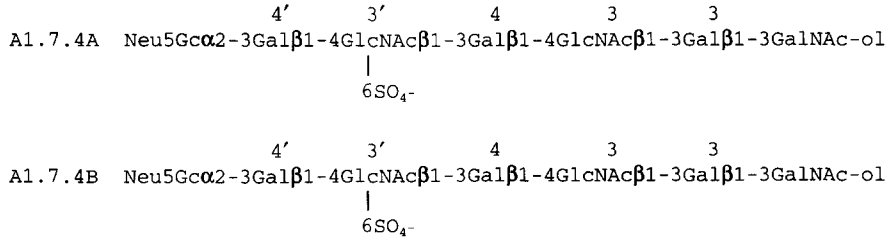
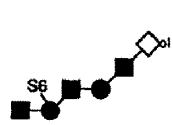
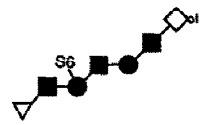
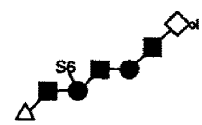


Table I

^{1}H chemical shifts of structural reporter-group protons of the constituent monosaccharides of the oligosaccharides A2.3, A1.7.4A and A1.7.4B, derived from porcine zona pellucida glycoproteins

Residue	Reporter Group	Chemical shift (ppm) in		
				
		A2.3	A1.7.4A	A1.7.4B
GalNAc-ol	H-2	4.401	4.399	4.399
	H-3	4.049	4.048	4.048
	H-4	3.490	3.489	3.489
	H-5	4.187	4.188	4.188
	NAc	2.047	2.047	2.047
Gal^3	H-1	4.460	4.459	4.459
	H-4	4.126	4.126	4.126
GlcNAc^3	H-1	4.679	4.677	4.677
	H-6	3.950	n.d.[a]	n.d.
	NAc	2.038	2.037	2.037
Gal^4	H-1	4.470	4.469	4.469
	H-4	4.188	4.193	4.193
GlcNAc$^{3'}$	H-1	4.716	4.707	4.707
	H-6	4.395	4.403	4.403
	H-6'	4.314	4.314	4.314
	NAc	2.032	2.029	2.029
Gal$^{4'}$	H-1	4.523	4.601	4.601
	H-3	3.67[b]	4.12-13[b]	4.12-13[b]
	H-4	3.926	3.969	3.963
Neu5Gc	H-3a	-	1.819	-
	H-3e	-	2.767	-
	NGc	-	4.116	-
Neu5Ac	H-3a	-	-	1.804
	H-3e	-	-	2.746
	NAc	-	-	2.029

Chemical shifts are given at 22°C and were measured in ^2H_2O relative to internal acetone (δ 2.225 [23]). Compounds are represented by short hand symbolic notation; ◇^{ol}, GalNAc-ol; ■, Gal; ●, GlcNAc; ▽, Neu5Gcα2-3; △, Neu5Acα2-3; S6, 6-*O*-sulfate [17]. For indexing of the monosaccharide residues, see text.

[a]n.d., not determined; [b]value obtained from the 2D HOHAHA spectrum.

from the Neu5Acα2-6Galβ1-4GlcNAc to the Neu5Acα2-6Galβ1-4($6SO_4$-)GlcNAc sequence ($\Delta\delta$ 0.041; cf. compounds 2-2 and S_3-2 in [19]). On the H-1 track of Gal^4 (δ 4.470) in the 2D HOHAHA spectrum, the corresponding H-4 resonance is found at δ 4.188. The downfield shift of the Gal^4 H-4 signal ($\Delta\delta$ 0.03), as compared to Gal H-4 in a non-sulfated GlcNAcβ1-3Gal element as part of a poly-(*N*-acetyllactosamine) sequence (δ 4.16 [20,21]), is also observed when going from the disaccharide GlcNAcβ1-3Gal to ($6SO_4$-)GlcNAcβ1-3Gal (con-

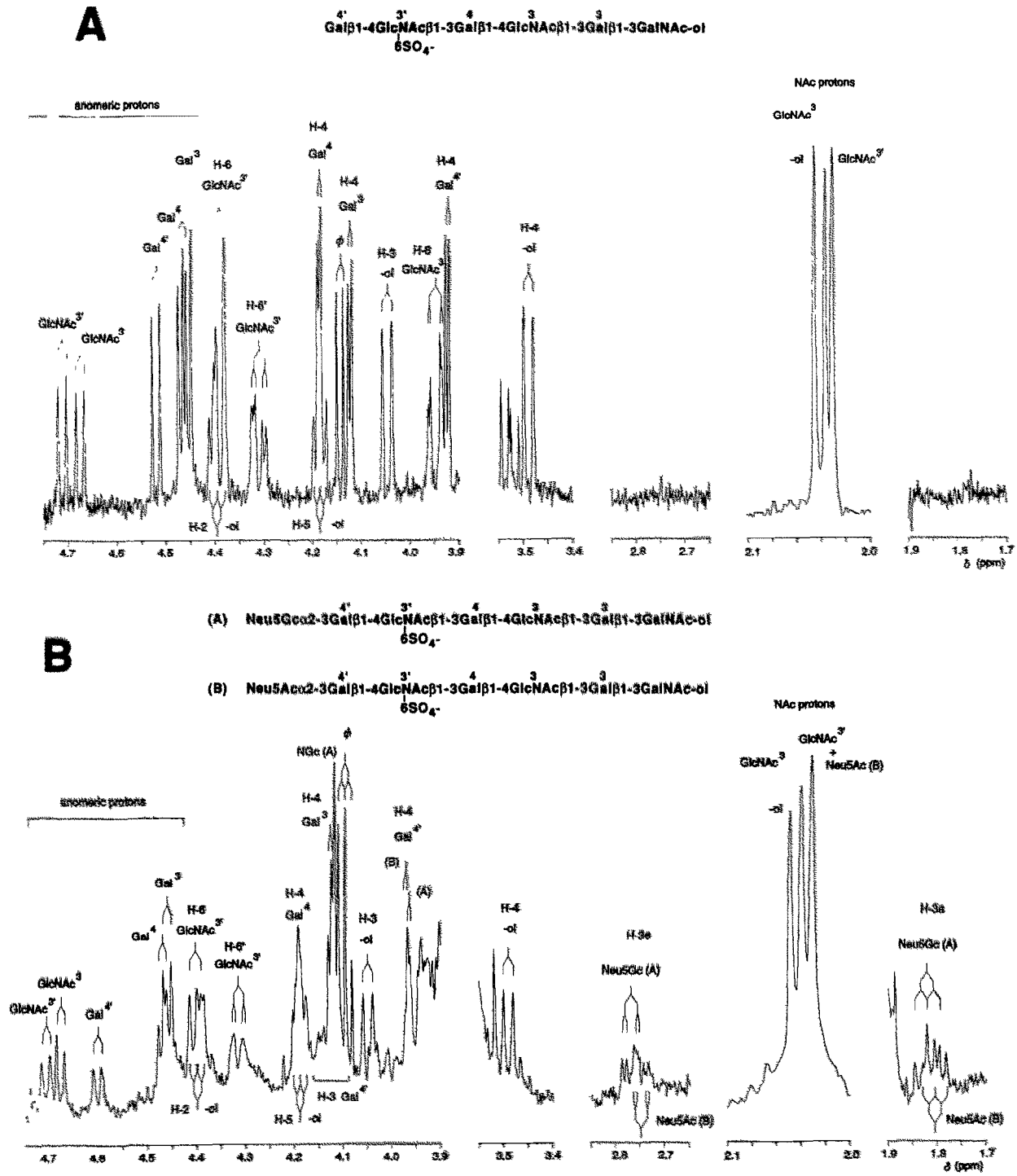


Fig. 1. Structural reporter-group regions of the resolution-enhanced 500-MHz ^1H NMR spectra of O-linked carbohydrate chains derived from porcine zona pellucida glycoproteins, recorded in ^2H_2O at 22°C. The relative scale of the NAc proton region differs from that of the rest of the spectrum. ϕ denotes non-carbohydrate impurity. (A) Fraction A2.3; (B) fraction A1.7.4.

sidering the β-anomers; cf. compounds A and B in [18]). The position of the Gal^4 H-1 signal (δ 4.470) is generally observed for an internal Gal residue in a poly-(*N*-acetyllactosamine) sequence [18,20,21].

The ^1H NMR spectrum of fraction A1.7.4 (Fig. 1B) indicates the occurrence of a mixture of two differently sialylated analogues of A2.3, denoted A1.7.4A and A1.7.4B, in a ratio of about 2:1.

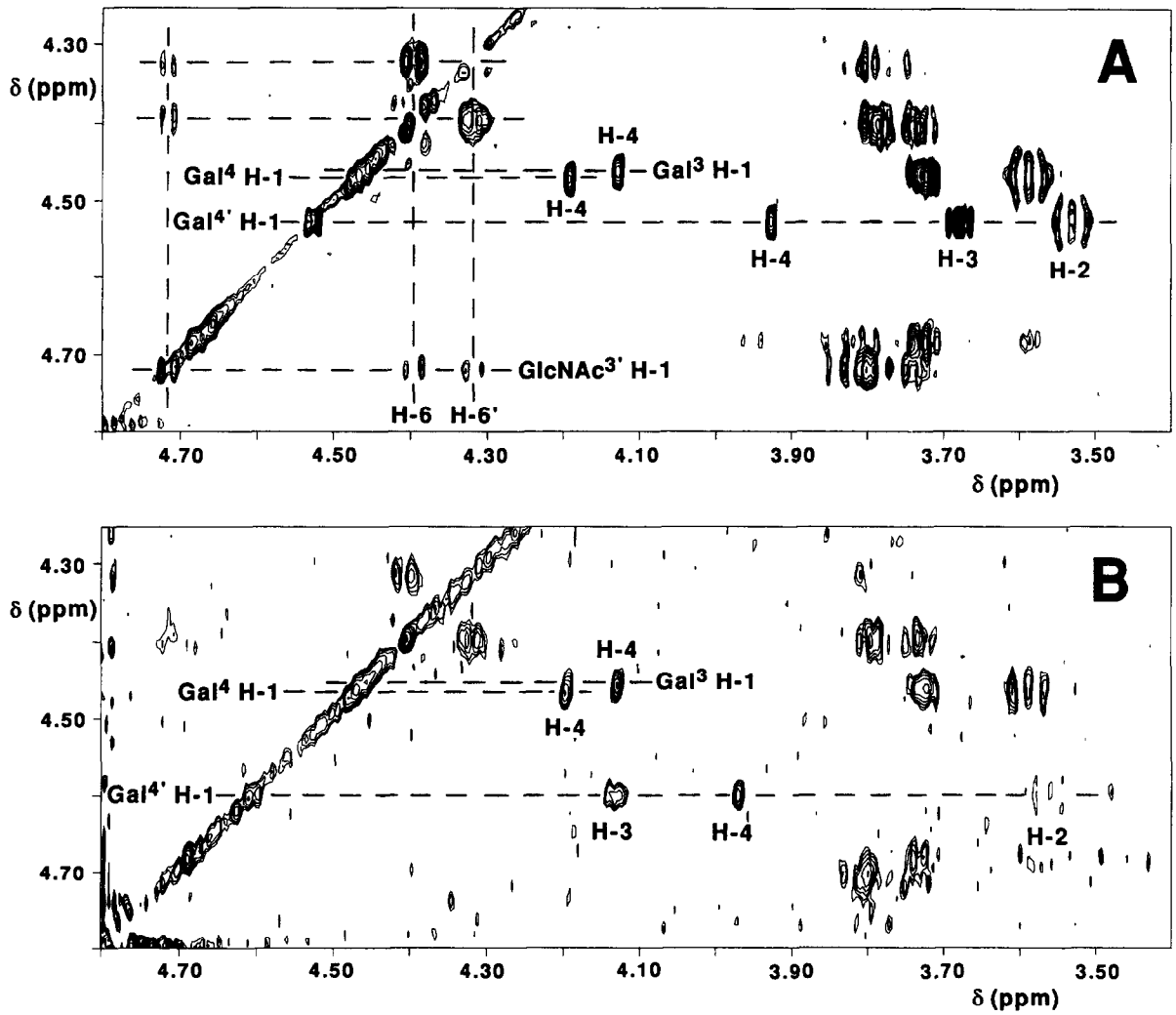


Fig. 2. Relevant regions of the 2D HOHAHA spectra at 500 MHz of O-linked carbohydrate chains derived from porcine zona pellucida glycoproteins, recorded in ^2H_2O at 22°C. Lines are drawn to show the interconnection between the protons of one residue. (A) Fraction A2.3; (B) fraction A1.7.4.

The structural reporter-group signals of the (6SO_4-)GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-3GalNAc-ol element are similar to the corresponding signals in the spectrum of fraction A2.3. The presence of the two different types of α2-3-linked sialic acid is deduced from the characteristic sets of the Neu5Gc (H-3a, δ 1.819; H-3e, δ 2.767; NGc, δ 4.116) and Neu5Ac (H-3a, δ 1.804; H-3e, δ 2.746; NAc, δ 2.029) structural reporter-group signals [22,23]. Furthermore, the Gal$^{4'}$ H-1 signal at δ 4.601 has shifted downfield ($\Delta\delta$ 0.078) when compared to that of Gal$^{4'}$ in A2.3. A similar shift is observed for α2-3-sialylation of a Gal residue in β1-4-linkage to a non-sulfated GlcNAc residue (cf. compounds 12 and 81 in [17]). The presence of the Neu5Gc/Acα2-3Gal element is corroborated by the positions of the Gal$^{4'}$ H-3 and H-4 resonances (δ 4.12-13 and δ 3.969, respectively, for A1.7.4A; δ 4.12-13 and δ 3.963, respectively, for A1.7.4B). The 2D HOHAHA spectrum demonstrates the interconnection of each of the Gal$^{4'}$, Gal^4 and Gal^3 H-1 signals with their respective H-2, H-3 and H-4 resonances (cf. compound A2.3).

4. DISCUSSION

In this paper, the structural analysis by ^{1}H NMR spectroscopy is described of three acidic O-linked oligosaccharide alditols of low molecular mass, derived from porcine zona pellucida glycoproteins. In these oligosaccharide alditols, the structural elements Galβ1-3GalNAc-ol, Galβ1-4GlcNAc, Galβ1-4(6SO_4-)GlcNAc and α2-3-linked Neu5Gc or Neu5Ac are present in a unique combination. On the basis of preliminary chromatographic and ^{1}H NMR spectroscopic data of the other fractions of the O-linked oligosaccharide alditols derived from pZP, it is concluded that the high molecular mass *O*-glycans contain the same structural ele-

ments. It has been suggested previously that both the O-and N-linked carbohydrate chains of pZP glycoproteins contain (sialylated) sulfated poly-(*N*-acetyllactosamine) chains [5,8,9,12]. In fact, it has been proposed that the O-linked carbohydrate chains of pZP3 consist of extensions of a mucin-type tetrasaccharide containing GlcNAc, Gal and GalNAc in the molar ratio 1:2:1 [5]. This is in agreement with the results shown in this study.

The presented structures indicate that the acidic *O*-glycans from pZP can be terminated with a β1-4-linked Gal residue, and that this Gal residue can be substituted with α2-3-linked Neu5Gc or Neu5Ac. This contrasts with the findings of Noguchi and Nakano [12] in that they identified the fragments Neu5Gc/Acα2-3Galβ1-4($6SO_4$-)GlcNAcβ1-3Gal and ($6SO_4$-)GlcNAcβ1-3Gal, but not Galβ1-4($6SO_4$-)GlcNAcβ1-3Gal, in the endo-β-galactosidase digest of pZP3-derived O-linked carbohydrate chains. Interestingly, the peripheral GlcNAc residue is sulfated in each compound, whereas the GlcNAc residue attached directly to the core is not. The preliminary NMR data of the fractions of the high-molecular mass oligosaccharides, which are currently being studied in more detail, suggest that in the longer poly-(*N*-acetyllactosamine) chains almost all peripheral GlcNAc residues are sulfated. Recently, it has been shown that the branch elongations of the acidic N-linked carbohydrate chains of pZP3 are also constructed from the Galβ1-4($6SO_4$-)GlcNAc and Neu5Gc/Acα2-3Galβ1-4($6SO_4$-)GlcNAc elements [6,12]. Together, the aforementioned data show that the peripheral extensions of the acidic N- and O-linked carbohydrate chains have identical structural characteristics. Here it is shown that in the O-linked oligosaccharides, the poly-(*N*-acetyllactosamine) chains are attached to a Galβ1-3GalNAc core, whereas the N-linked carbohydrate chains are formed from poly-(*N*-acetyllactosamine) sequences linked to the branches of di-, tri-, or tetraantennary oligosaccharides [12].

Finally, it should be noted that the isolation and structure identification of the *O*-glycans presented in this paper allows further investigations concerning the correlation of the structure and biological activity of pZP carbohydrate chains, or fragments thereof.

Acknowledgements: This investigation was supported by the Netherlands Program for Innovation Oriented Carbohydrate Research (IOP-k) with financial aid from the Ministry of Economic Affairs and the Ministry of Agriculture, Nature Management and Fisheries, and Organon International BV (Oss, The Netherlands). The authors thank Prof. J. Aitken for providing porcine zona pellucidae.

REFERENCES

[1] Wassarman, P.M. (1987) Annu. Rev. Cell. Biol. 3, 109–142.
[2] Wassarman, P.M. (1988) Annu. Rev. Biochem. 57, 415–442.
[3] Florman, H.M. and Wassarman, P.M. (1985) Cell 41, 313–324.
[4] Bleil, J.D. and Wassarman, P.M. (1988) Proc. Natl. Acad. Sci. USA 85, 6778–6782.
[5] Yurewicz, E.C., Pack, B.A. and Sacco, A.G. (1991) Mol. Repr. Dev. 30, 126–134.
[6] Noguchi, S., Hatanaka, Y., Tobita, T. and Nakano, M. (1992) Eur. J. Biochem. 204, 1089–1100.
[7] Nakano, M., Hatanake, Y., Sawai, T., Kobayashi, N. and Tobita, T. (1987) Biochem. Int. 14, 417–423.
[8] Nakano, M., Hatanake, Y., Kobayashi, N., Noguchi, S., Ishikawa, S. and Tobita, T. (1990) J. Biochem. 107, 144–150.
[9] Yurewicz, E.C., Sacco, A.G. and Subramanian, M.G. (1987) J. Biol. Chem. 262, 564–571.
[10] Sacco, A.G., Yurewicz, E.C., Subramanian, M.G. and Matzat, P.D. (1989) Biol. Reprod. 41, 523–532.
[11] Mori, E., Takasaki, S., Hedrick, J.L., Wardrip, N.J., Mori, T. and Kobata, A. (1991) Biochemistry 30, 2078–2087.
[12] Noguchi, S. and Nakano, M. (1992) Eur. J. Biochem. 209, 883–894.
[13] Damm, J.B.L., Bergwerff, A.A., Hård, K., Kamerling, J.P. and Vliegenthart, J.F.G. (1989) Recl. Trav. Chim. Pays-Bas 108, 351–359.
[14] Damm, J.B.L., Kamerling, J.P., Van Dedem, G.W.K. and Vliegenthart, J.F.G. (1987) Glycoconjugate J. 4, 129–144.
[15] Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J.P. and Vliegenthart, J.F.G. (1992) Eur. J. Biochem. 209, 895–915.
[16] Bax, A. and Davis, D.G. (1985) J. Magn. Reson. 65, 355–360.
[17] Kamerling, J.P. and Vliegenthart, J.F.G. (1992) Biol. Magn. Reson. 10, 1–194.
[18] Hounsell, E.F., Feeney, J., Scudder, P., Tang, P.W. and Feizi, T. (1986) Eur. J. Biochem. 157, 375–384.
[19] De Waard, P., Koorevaar, A., Kamerling, J.P. and Vliegenthart, J.F.G. (1991) J. Biol. Chem. 266, 4237–4243.
[20] Feizi, T., Hounsell, E.F., Alais, J., Veyrieres, A. and David, S. (1992) Carbohydrate Res. 228, 289–297.
[21] Hokke, C.H., Kamerling, J.P., Van Dedem, G.W.K. and Vliegenthart, J.F.G. (1991) FEBS Lett. 286, 18–24.
[22] Hokke, C.H., Bergwerff, A.A., Van Dedem, G.W.K., Van Oostrum, J., Kamerling, J.P. and Vliegenthart, J.F.G. (1990) FEBS Lett. 275, 9–14.
[23] Vliegenthart, J.F.G., Dorland, L. and Van Halbeek, H. (1983) Adv. Carbohydr. Chem. Biochem. 41, 209–374.