Glycoconj J (2006) 23:377–399
DOI 10.1007/s10719-006-6126-4

Structure elucidation of NeuAc, NeuGc and Kdn-containing *O*-glycans released from *Triturus alpestris* oviductal mucins

Characterization of the poly LacdiNAc sequence: HSO3(4)(GalNAcβ1-4GlcNAcβ1-3)$_{1-3}$GalNAcβ1-4(GlcNAcβ1-3)$_{0-1}$GlcNAcβ1-6GalNAc-ol

Doina Florea · Emmanuel Maes · Yann Guérardel · Adeline Page · Jean-Pierre Zanetta · Dan Cogalniceanu · Gérard Strecker

Received: 16 September 2005 / Revised: 4 November 2005 / Accepted: 9 November 2005

Abstract Eggs from Amphibia are surrounded by several layers of jelly that are needed for proper fertilization. Jelly coat is composed of highly glycosylated mucin-type glycoproteins containing up to 60% of carbohydrates, which display a remarkable species-specificity. This material obtained from *Triturus alpestris* was submitted to reductive β-elimination and the released oligosaccharide-alditols were further fractionated by HPLC. Structural characterization was performed through a combination of two dimensional ^{1}H-^{1}H and ^{1}H-^{13}C NMR and ESI-MS/MS analysis. Numerous carbohydrate chains are characterized by the presence of the Cad (Sd^{a}) determinant, including respectively NeuAc, NeuGc or Kdn as a sialic acid. But the most significant *O*-glycan sequences which mark the difference between the jelly of *T. alpestris* and other studies amphibian jellies are polymers of GalNAc(β1-4)GlcNAc (LacdiNAc) which form part of the following sequence:

HSO_3(4)(GalNAcβ1-4GlcNAcβ1-3)$_{1-3}$GalNAcβ1-4(GlcNAcβ1-3)$_{0-1}$GlcNAcβ1-6GalNAc-ol

This work is dedicated to the memory of Prof. Yasuo Inoue, for his remarkable contribution to Sialobiology.

D. Florea · E. Maes · Y. Guérardel · J.-P. Zanetta · G. Strecker (✉)
Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576, USTL, 59655 Villeneuve d'Ascq Cedex France
e-mail: gerard.strecker@univ-lille1.fr
Tel: 33-(0)3-20 33 72 42
Fax: 33-(0)3-20 43 65 55

A. Page
Centre Commun de Mesure de Spectrométrie de Masse, Université des Sciences et Technologies de Lille 1, Bâtiment C4, 59655 Villeneuve d'Ascq Cedex, France

D. Cogalniceanu
Departamentul de Ecologie, Facultatea de Biologie, Universitatea din Bucuresti, România

Keywords Amphibia · *Triturus alpestris* · Jelly coat · Mucin · O-glycan · Poly LacdiNAc · NMR · ESI-MS/MS

Abbreviations

ESI-MS	electrospray mass spectrometry
ESI-MS/MS	electrospray mass-spectrometry-mass spectrometry
COSY	COrrelation SpectroscopY
ROESY	Rotating-frame nuclear Overhauser Enhancement SpectroscopY
HMQC	Heteronuclear Multiple-Quantum Coherence
Gal	galactose
GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
Fuc	fucose
GlcA	glucuronic acid
NeuAc	*N*-acetylneuraminic acid
NeuGc	*N*-glycolylneuraminic acid
Kdn	2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid or deaminoneuraminic acid

Introduction

Previous reports on fish [1], bird [2] and amphibia [3] mucins have shown that these materials are easily accessible sources of carbohydrates and have confirmed that *O*-glycans are highly species-specific. The number of all possible linear and branched isomers of a hexasaccharide has been calculated and found to be higher than 10^{12} [4]. This number appears higher than the number of living species on earth. Indeed, whereas scientists have successfully classified over 1.5 million species, as little as 2 million to as many as 50 million more species have not yet been found, according

to conflicting estimations of specialists. Considering the role ascribed to carbohydrates as an evolutionary potential of information, their species-specificity could be correlated with a great number of recognition systems, such as, *inter allia*, parasitism or microbial pathogenesis [5]. Another consequence of the diversity of oligosaccharide structures was to define the *Isomer barrier* [4], that means a persistent technological barrier to the development of a single method for the absolute characterisation of carbohydrate chains. Nevertheless, strategies combining NMR experiments and various mass spectrometry procedures easily overcome nowadays most of the encountered difficulties. Here, we report the structures of the main carbohydrate chains which appear to be specific for *Triturus alpestris*.

Materials and methods

Sampling of jelly coat mucus

Spawns from *Triturus alpestris* were induced by injections of 500 IU of human chorionic gonadotrophin (ORGANON, France).

Isolation of oligosaccharide-alditols

O-linked oligosaccharides were released from the crude material (1.2 g) by reductive β-elimination in a 1M $NaBH_4$/0.05M NaOH solution (pH 12.6) at 30°C for 48 h. The reaction was stopped by addition of acetone, followed by Dowex 50 × 8 (25–50 Mesh; H^+ form). After filtration the solution was adjusted to pH 6.5 with ammonia and concentrated under vacuum. Borate salts were removed by repeated evaporations with methanol and residual peptides retained by cation exchange chromatography using a Dowex 50 × 2 column (200–400 Mesh; H^+ form). The pH of the effluent was adjusted to neutrality and lyophilized. Oligosaccharide-alditols were fractionated according their size by gel permeation on a Bio-Gel P-6 column (125 × 2cm) and then isolated on a primary amine-bonded silica column (Supelcosyl^{TM}, LC-NH_2; 4.6 mm × 25 cm, Supelco Inc., Bellefonte, USA) by high performance liquid chromatography (HPLC) using gradient of acetonitrile (A)/30 mM potassium phosphate buffer pH 5.2 (B)/H_2O (C): time 0: A 80%, C 20%; time 60 min: A 50%, B 50%; time 90min: A 50%, B 50%. Oligosaccharides were detected by UV spectroscopy at 206 nm.

Analytical procedures

The molar ratios of component monosaccharides and the diversity of sialic acids were determined by GLC as heptafluorobutyrate derivatives [6,7].

Mild periodate oxidation

For mild oxidation, the reduced oligosaccharides were treated with 200 nmol of sodium meta-periodate in 40 mM imidazole–HCl (pH 6.5) on ice in the dark for 45 min. Excess oxidant was reacted with glycerol (2 μmol) under the same conditions for 40 min. The reaction product was reduced with 10 mg/ml sodium borohydride in 2M NH_4OH for 2 h at room temperature. The reaction was terminated by adding glacial acetic acid and desalted on a Dowex (50 × 8, 50–100 Mesh, H^+ form) column in water. Borates were removed by repeated codistillation in methanol and samples were finally desalted by gel filtration on Bio-Gel P-2 column (Biorad).

Electrospray mass spectrometry (nano-ESI-MS/MS)

All analyses were performed using a Q-STAR Pulsar quadrupole time-flight (Q-TOF) mass spectrometer (Applied Biosystems/MDS Sciex, Toronto, Canada) fitted with a nanoelectrospray ion source (Protana, Odense, Denmark). Glycans dissolved in a solution of 50% methanol and 1% formic acid (0.5 pmol/μL) were sprayed from gold-coated "medium length" borosilicate capillaries (Protana). A potential of 800 V was applied to the capillary tip. For generation of MS/MS data, the precursor ion was selected by the quadrupole, and was subsequently fragmented in the collision cell using nitrogen at a pressure of about 5.3×10^{-5} Torr and an appropriate collision energy. The CID spectra were recorded by the orthogonal TOF analyzer over the range *m/z* 50–2000.

Results

Characterization of sialic acids

GC/MS analyses of the diversity of sialic acids showed the presence of Neu5Ac, Neu5Gc and Kdn derivatives (in the ratio 75:8:17, respectively) and a high degree of *O*-acetylation of Neu5Ac and Kdn derivatives. Indeed, Neu5,9Ac_2 represented 56% of Neu5Ac derivatives (traces of Neu4,5Ac_2 were also observed), whereas, Kdn was present as such (42% of total Kdn derivatives), as mono-*O*-acetylated sialic acids (Kdn9Ac: 39%, Kdn7Ac: 6%, Kdn5Ac: 2%) and as di-*O*-acetylated Kdn (Kdn8,9Ac_2: 7% and Kdn7,9Ac_2: 5%). These *O*-acetylated derivatives of sialic acids were not observed in the analysis of oligosaccharide-alditols since these ester bonds were cleaved during the reductive β-elimination procedure.

Isolation of oligosaccharide-alditols

The jelly-coat material of *Triturus alpestris* was submitted to reductive β-elimination, and the released oligosaccharides were desalted by gel filtration on a Bio-Gel P-2 column. The pool of *O*-linked carbohydrate chains was then submitted to a cationic exchange chromatography procedure to remove residual peptides or glycopeptides. A fractionation by gel filtration on a Bio-Gel P-6 column of the carbohydrate material resulted in seven fractions, namely **I** to **VII**. These fractions contained oligosaccharides presenting distinct chromatographic properties as observed in TLC analysis, indicating the presence of different compounds in each fraction (data not shown). Fractions **II** to **VII** were further fractionated by HPLC on a primary-amine-bonded silica column to obtain either individual glycans or mixtures of two or three glycans (Fig. 1) that were analysed as such by NMR spectroscopy and mass spectrometry. Fraction **I** was identified as a complex mixture containing very high molecular *O*-linked glycans that were analysed as such by NMR and mass spectrometry. Fifty nine different oligosaccharides, labelled from **1** to **59** ordered by increasing molecular mass, were finally observed and fully described. Proposed structures of all labelled 59 *O*-glycans are depicted in Scheme 1, with indication of the origin of each component.

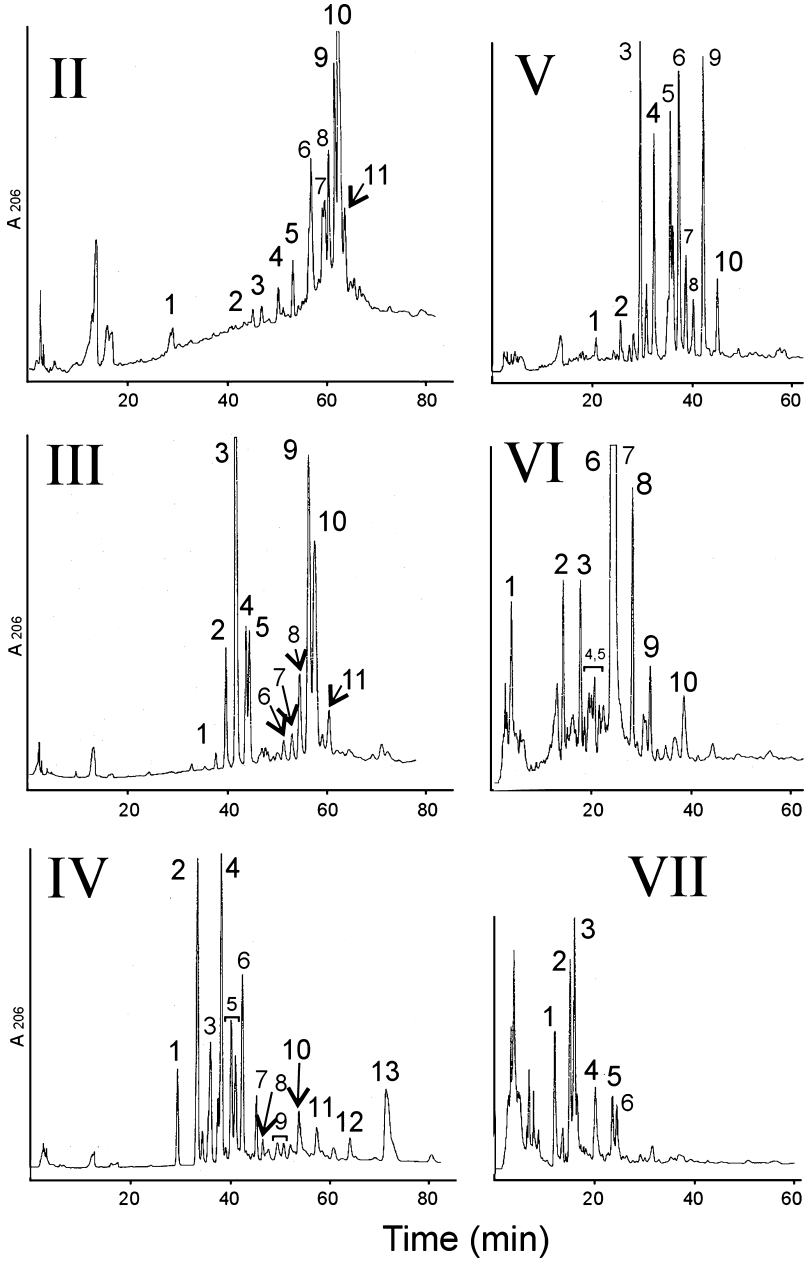
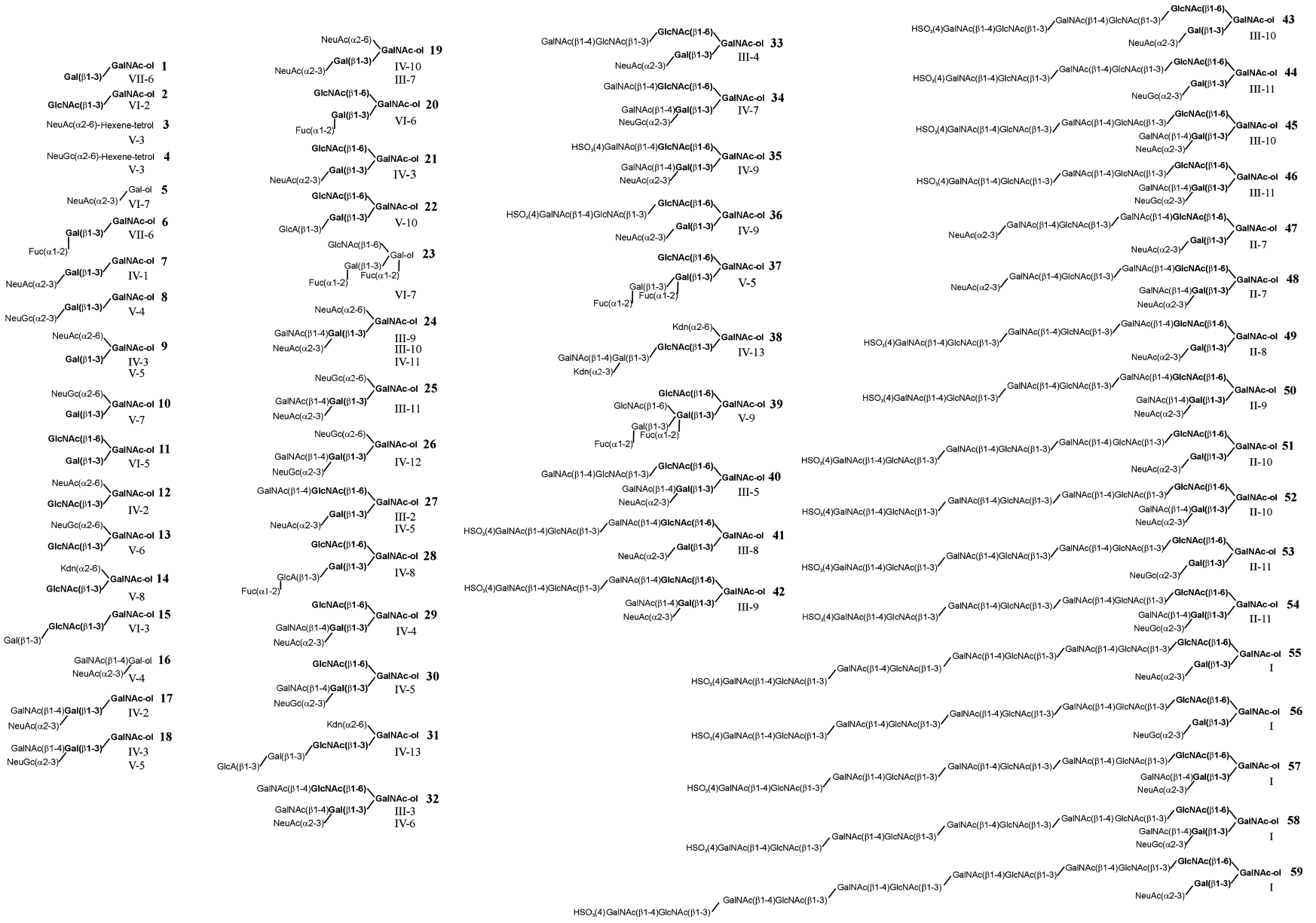


Fig. 1 HPLC profiles on a primary amine-bonded silica column of fractions II to VII obtained by gel filtration on a Bio-Gel P6 column



Scheme 1 Structures of oligosaccharide-alditols released from *T. alpestris* egg jelly coats. (Key: *ie* VII-6 means HPLC peak 6 of Bio-Gel P-6 fraction VII, as shown in Fig. 1)

Structural determination of oligosaccharide-alditols

The primary structures of all oligosaccharide-alditols were established by a combination of NMR and ESI-MS/MS experiments. Chosen experimental data are presented (Fig. 2 to 9) and NMR parameters of all compounds are shown in Tables 1–8. Systematic analyses revealed that 28 components presented well known structures, which have been already described and will not be further discussed [8]. In contrast, 31 compounds (**22**, **25**, **27** and **31** to **58**) exhibited original sequence structures that appear as characteristic of the *Triturus alpestris* species.

Characterization of the core structures

The nature of the core structure is usually inferred from the resonance parameters of GalNAc-ol H-2 and H-5, easily observable in one dimension ^{1}H-NMR experiments, despite the fact that the substitution of Gal residue of the Gal(β1-3)GalNAc-ol unit may strongly influence the GalNAc-ol H-2 parameters. In the present report, we took advantage of the 2D NMR techniques to directly identify GalNAc-ol H-1,1′, and H-6,6′ signals from the bulk region. As observed in Table 1, these data give more reliable information about the nature of the core structure and permit to unambiguously discriminate between cores 1, 2 and 3, as well as their eventual sialylated counterparts present in the analysed components. These attributions were in good agreement with structural reporter groups already published for GalNAc-ol substituted in C-3 position by β-Gal or β-GlcNAc and substituted in C-6 by β-GlcNAc or a-NeuAc/Gc residues [8].

In a second step the structures of backbone and/or peripheral elements in C-3 and C-6 positions of GalNAc-ol were determined owing to the delineation of structural-reporter-groups deduced from the complete NMR analyses (COSY, TOCSY, ROESY and HMQC) of the most representative sequences described in Scheme 1. Structural analyses will be presented according to the features of C-3 and C-6 branches independently.

Sia(α2-3)Gal(β1-3) and Sia(α2-3)[GalNAc(β1-4)] Gal(β1-3) sequences in C-3 position of GalNAc-ol

These two sequences directly attached to C-3 position of GalNAc-ol are present in the vast majority of *T. alpestris* *O*-glycans and characterise 36 compounds out of the 59 depicted in Scheme 1. Sialic acid is either NeuAc or NeuGc, most *O*-glycans being simultaneously observed as their acetylated and glycolylated counterparts. Kdn was also observed in compound **38** as a "Sd^{a}/Cad-type" epitope linked to a type 3 core instead of type 1 and 2 cores. The chemical shift values of characteristic structural-reporter-groups for these sequences are summarized in Table 2. The combined set of the sialic acid H-3$_{ax}$ and H-3$_{eq}$ atom resonances gives direct information about the type of glycosidic linkage (α2-3 or 2-6), as well as the nature of the sugar unit (NeuAc, NeuGc or Kdn) and the molecular environment in which it occurs. β-Gal H-3 and H-4 chemical shift increments are independent of the nature of the sialic acid but are largely influenced by the presence of the β1-4-linked GalNAc unit (blood-group Cad or Sd^{a} determinant) [8]. Data summarized in Scheme 1 permitted to attribute the sequences NeuAc(α2-3)Gal, NeuGc(α2-3)Gal, NeuAc(α2-3)[GalNAc(β1-4)]Gal and NeuGc(α2-3)[GalNAc(β1-4)]Gal in C-3 positions of respectively 13, 4, 12 and 7 compounds out of 59.

Neutral and acidic LacdiNAc and poly LacdiNAc sequences (GalNAcβ1-4GlcNAcβ1-3)$_{n}$GalNAcβ1-4GlcNAcβ1-6 and (GalNAcβ1-4GlcNAcβ1-3)$_{n=1\text{-}5}$GlcNAcβ1-6 in C-6 position of GalNAc-ol

Presence of poly LacdiNAc substituted *O*-glycans appeared as the most characteristic feature of *T. alpestris* mucin glycosylation. Compounds presenting from 1 to 5 LacdiNAc repeat units were characterized in the present study. Compounds **27**, **32** and **34** presented a single LacdiNAc unit and were fully analyzed in ^{1}H-^{1}H COSY and ^{13}C-^{1}H HMQC NMR experiments in order to define the structural-reporter-groups specific for these sequences. COSY spec-

Table 1
Structural-reporter-group data of GalNAc-ol, specific for the type of core structures 1 to 3

	GalNAc-ol		
Core structure	H1,1′	H-6	H-6′
Core 1			
Gal(β1-3)GalNAc-ol	3.73–3.83	3.64–3.74	3.64–3.74
Gal(β1-3)[Sia^{a}(α2-6)]GalNAc-ol	3.73–3.83	3.85–3.88	3.44–3.51
Core 2			
Gal(β1-3)[GlcNAc(β1-6)]GalNAc-ol	3.73–3.83	3.88–3.95	3.64–3.71
Core 3			
GlcNAc(β1-3)GalNAc-ol	3.55–3.63	3.63–3.65	3.63–3.65
GlcNAc(β1-3)[Sia^{a}(α2-6)]GalNAc-ol	3.55–3.63	3.81–3.85	3.46–3.50

[a]Sia: NeuAc, NeuGc or Kdn.

Table 2 Structural-reporter-group data of sialic acid (NeuAc, NeuGc or Kdn), Gal and GalNAc for Sia(α 2-6)GalNAc-ol, Sial(α 2-3)Gal(β1-3)GalNAc-ol and Sia(α 2-3)[GalNAc(β1-4)]Gal(β1-3) sequences[8, and this study]

	Sialic acid				Gal		GalNAc
Structural element	NAc	NGc	H-3ax	H-3eq	H-3	H-4	H-1
NeuAc(α2-6)GalNAc-ol	2.030–2.033	–	1.69–1.71	2.72–2.74	–	–	–
NeuGc(α2-6)GalNAc-ol	–	4.122	1.71–1.72	2.74–2.75	–	–	–
Kdn(α2-6)GalNAc-ol	–	–	1.66	2.68	–	–	–
NeuAc(α2-6)Gal/GalNAc	2.030–2.033	–	1.72	2.67	–	–	–
NeuAc(α2-3)Gal(β1-3)GalNAc-ol	2.030–2.033	–	1.80	2.77–2.78	4.11–4.12	3.93	–
NeuGc(α2-3)Gal(β1-3)GalNAc-ol	–	4.122	1.808–1.818	2.78–2.79	4.11–4.13	3.93	–
Kdn(α2-3)Gal(β1-3)GalNAc-ol	–	–	1.750	2.720	4.09	3.92	–
NeuAc(α2-3)[GalNAc(β1-4)]Gal(β1-3)GalNAc-ol	2.030–2.034	–	1.93–1.96	2.68–2.71	4.16–4.18	4.08–4.11	4.68–4.73
NeuGc(α2-3)[GalNAc(β1-4)]Gal(β1-3)GalNAc-ol	–	4.122	1.95–1.96	2.70–2.71	4.11–4.13	4.09–4.10	–
Kdn(α2-3)[GalNAc(β1-4)]Gal(β1-3)GlcNAcl	–	–	1.862	2.634	4.10	4.09	4.73

trum of compound **32** (Fig. 2) enabled a direct recognition of the blood group Cad determinant, owing to the presence of its representative signals defined in the previous paragraph. Chemical shifts of the GalNAc-ol H-6 and H-6′ signals typified the presence of a type 2 core, whereas the sugar units II′ and III′ were unambiguously identified as β-GlcNAc and β-GalNAc, respectively, on the basis of their first four ^3J coupling constants and of their J connectivities. Finally, the substitution of β-GlcNAc II′ in C-4 position by the β-GalNAc unit III′ was deduced from the downfield shift of the β-GlcNAc II′ C-4 atom resonance at δ∼80.8 ppm (Tables 6 and 8) and permitted to establish the structure of compound **32** as depicted on Scheme 1.

On these bases, compound **34** was identified as a glycolylated equivalent of compound **32** and presented almost identical NMR parameters but for the substitution of the *N*-acetyl signal at δ2.029 ppm for a *N*-glycolyl signal at δ4.122 ppm, whereas compound **27** was identified as its equivalent with a NeuAc(α2-3)Gal lower branch, according to Tables 2 and 6.

The two-step relayed COSY and HMQC spectra of the compound **33** are shown in Fig. 3. As for compound **32**,

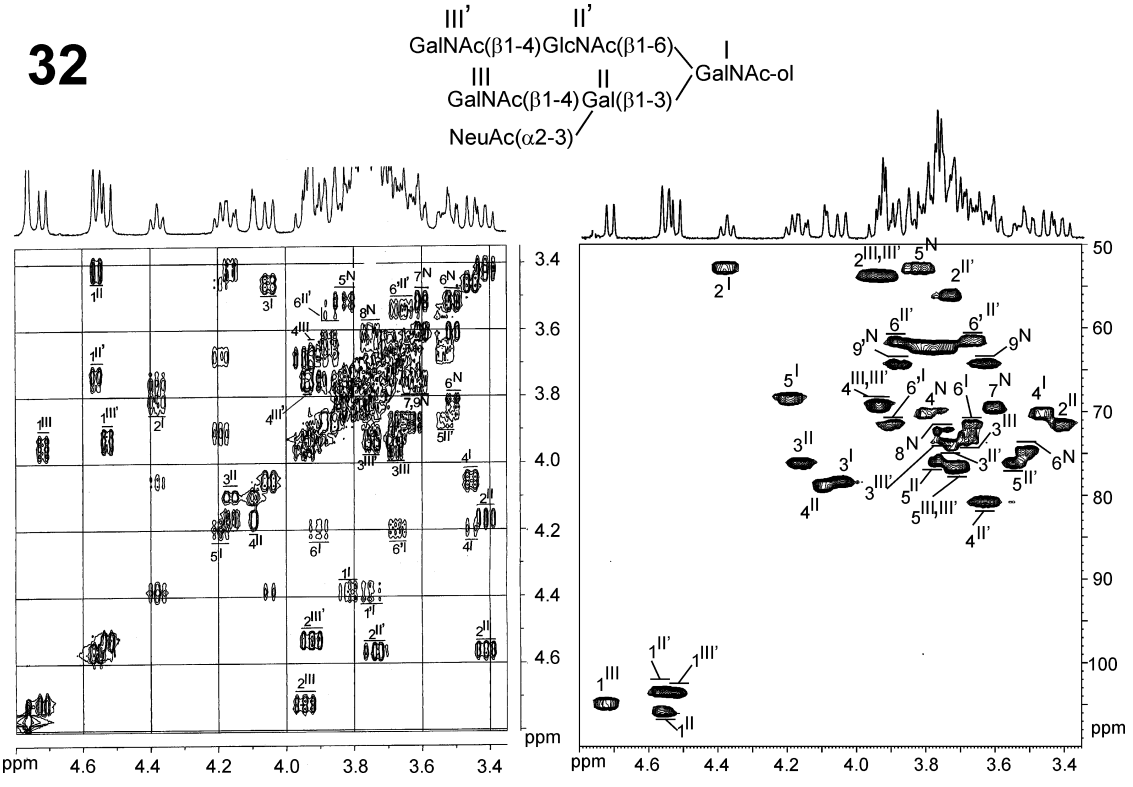


Fig. 2 COSY (left) and HMQC (right) spectra of compound **32**

Table 3
Structural-reporter-group data of HexNAc units specific for the sequence **I** : HSO_3(4) GalNAc(β1-4)[GlcNAc (β1-3)GalNAc(β1-4)]$_n$ GlcNAc$^{II'}$(β1-6) and **II** : HSO_3(4)GalNAc(β1-4) [GlcNAc(β1-3)GalNAc (β1-4)]$_n$GlcNAc$^{II''}$(β1-3) GlcNAc$^{II'}$ (β1-6)

[a] No Cad and Cad are relative to the presence of the sequences Sia(α2-3)Gal(β1-3) and Sia(α2-3)[GalNAc(β1-4)]Gal(β1-3), at the lower branches, respectively.

		Sequence I		Sequence II	
Sugar unit	Lower branch	no Cad[a]	Cad[a]	no Cad[a]	Cad[a]
GlcNAc(β1-6) **II′**	H-1	**4.53**	**4.55**	**4.48**	**4.506**
	H-2	**3.72**	**3.74**	**3.77**	**3.77**
GlcNAc(β1-3) **II″**	H-1	–	–	4.56–4.57	4.59–4.60
	H-2	–	–	3.71–3.72	3.72
3)GalNAc(β1-4)	H-1	4.51	4.51	4.51	4.51
	H-2	3.97	3.97	3.97	3.97
4)GlcNAc(β1-3)	H-1	4.56–4.57	4.56–4.57	4.56–4.57	4.56–4.57
	H-2	3.71–3.73	3.71–3.73	3.71–3.73	3.71–3.73
HSO_3(4)GalNAc(β1-4)	H-1	4.57–4.58	4.57–4.58	4.57–4.58	4.57–4.58
	H-2	3.93	3.93	3.93	3.93
	H-4	4.69	4.69	4.69	4.69

the occurrence of the Gal(β1-3)[GlcNAc(β1-6)]GalNAc-ol core 2 structure was clearly established owing to GalNAc-ol H-1,1′ and H-6,6′ signals whereas the lower branch was identified as NeuAc(α2–3)Gal(β1-3) on the basis of Gal(β1-3) H-3, H-4 parameters and of the presence of the *N*-acetyl signal (Tables 2 and 6). In addition to the terminal βGalNAc III′, two other monosaccharide units were observed and identified as GlcNAc units (II′, II″) on the basis of the values of their first four vicinal coupling constants $^3J_{1,2}$ to $^3J_{4,5}$. On ^{1}H-^{13}C heteronuclear spectrum (Fig. 3), GlcNAc II′ residue exhibited a C-3 signal at δ 82.6 ppm, establishing that it was substituted in C-3 position whereas chemical shift of GlcNAc II″ C-4 at δ 81.1 ppm established that this residue was substituted in C-4 position. Taken together NMR data established the structure of **33** as depicted on Scheme 1.

Then, an equivalent of compound **33** presenting a Cad type epitope instead of the NeuAc(α2–3)Gal(β1-3) unit was observed for compound **40** owing to the differential parameters of their respective Gal II H-3 and H-4 (Tables 2 and 6). It is noteworthy that the chemical shift value of the GlcNAc(β1-6) II′ H-1 is greatly affected by the presence of the β-1,4-linked GalNAc unit at the lower branch. Systematic comparison of couples of compounds only differing by the presence of the GalNAc(β1-4) III unit (**33** *vs* **40**, **41** *vs* **42**,

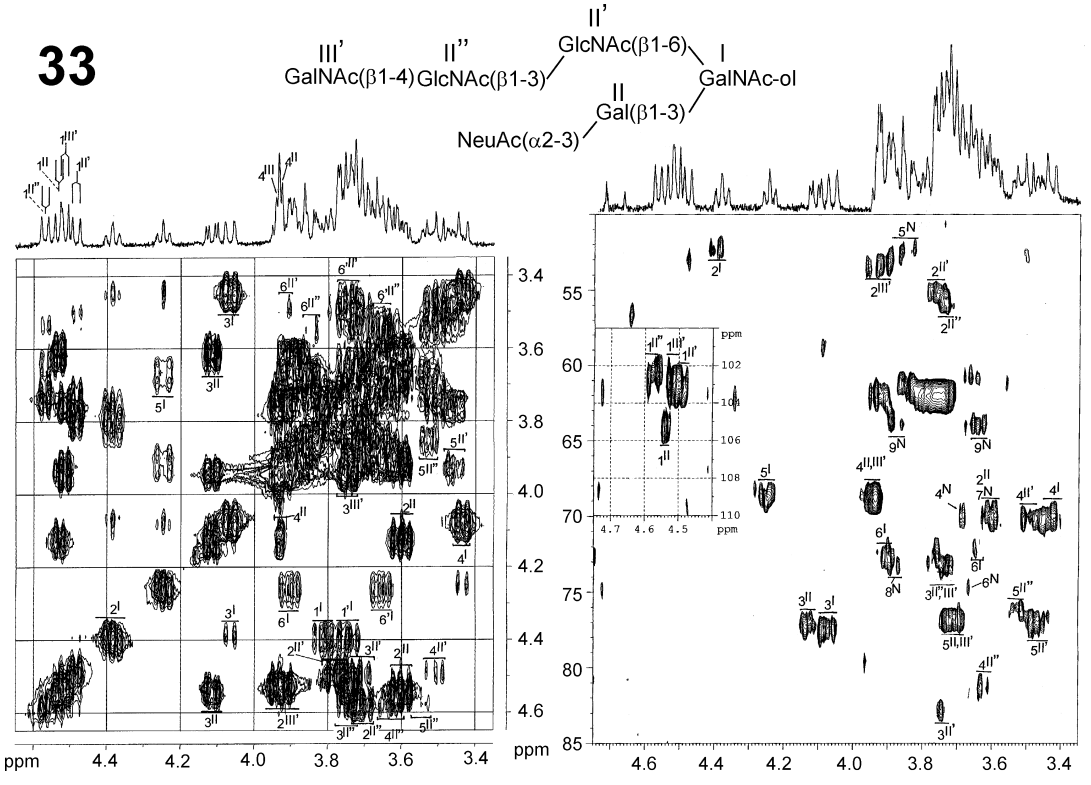


Fig. 3 Relayed COSY (left) and HMQC (right) spectra of compound **33**

Table 4 Monosaccharide composition data for oligosaccharide-alditols released from the mucin of the jelly coat. The molecular mass of the compounds was deduced from negative ions detected by ESI-MS

		Molecular mass		Molar ratio							
Compound	Ion detected	Measured	Theorical	GalNAc-ol	HexNAc	Gal	NeuAc	NeuGc	Kdn	GlcA	Sulphate
24	1190.40 (M-2H+Na)	1170.42	1170.43	1	1	1	2	0	0	0	0
	584.29 (M-2H)/2	1170.58									
25	1207.64 (M-2H+Na)	1186.64	1186.42	1	1	1	1	1	0	0	0
	592.30 (M-2H)/2	1186.60									
31	506.19 (M-2H)/2	1014.38	1014.33	1	1	1	0	0	1	1	0
35,36	1364.51 (M-H)	1365.51	1365.43	1	3	1	1	0	0	0	1
	1386.56 (M-2H+Na)	1365.58									
38	644.80 (M-2H)/2	1291.60	1291.45	1	2	1	0	0	2	0	0
42	884.90 (M-2H)/2	1771.80	1771.07	1	5	1	1	0	0	0	1
43	884.92 (M-2H)/2	1771.84	1771.07	1	5	1	1	0	0	0	1
	589.71 (M-3H)/3	1772.13									
44	892.97 (M-2H)/2	1787.94	1787.06	1	5	1	0	1	0	0	1
	594.70 (M-3H)/3	1787.10									
45	986.46 (M-2H)/2	1974.92	1974.68	1	6	1	1	0	0	0	1
	657.35 (M-3H)/3	1975.05									
46	994.45 (M-2H)/2	1990.90	1990.67	1	6	1	0	1	0	0	1
	662.77 (M-3H)/3	1991.31									
47	889.0 (M-2H)/2	1780.0	1779.7	1	4	1	2	0	0	0	0
48	990.5 (M-2H)/2	1983.0	1982.6	1	5	1	2	0	0	0	0
49	986.7 (M-2H)/2	1975.4	1974.7	1	6	1	1	0	0	0	1
	657.6 (M-3H)/3	1975.8									
50	1088.1 (M-2H)/2	2178.2	2177.2	1	7	1	1	0	0	0	1
51	1088.1 (M-2H)/2	2178.2	2177.2	1	7	1	1	0	0	0	1
52	1189.6 (M-2H)/2	2381.2	2380.8	1	8	1	1	0	0	0	1
53	730.7 (M-3H)/3	2195.1	2193.7	1	7	1	0	1	0	0	1
54	1197.8 (M-2H)/2	2397.6	2396.8	1	8	1	0	1	0	0	1
	798.4 (M-3H)/3	2398.2									
55	1291.0 (M-2H)/2	2584.0	2585.44	1	9	1	1	0	0	0	1
56	1299.3 (M-2H)/2	2600.6	2601.44	1	9	1	0	1	0	0	1
57	1392.6 (M-2H)/2	2787.2	2788.64	1	10	1	1	0	0	0	1
58	1400.9 (M-2H)/2	2803.8	2804.64	1	10	1	0	1	0	0	1
59	1494.0 (M-2H)/2	2990.0	2991.84	1	11	1	1	0	0	0	1

Table 5 ^{1}H-NMR chemical shifts of constituent monosaccharides for oligosaccharide-alditols with the Gal(β1-3)GalNAc core structure

Residue	Reporter groups	**1**	**6**	**7**	**8**	**17**	**18**	**9**	**10**	**19**	**24**	**26**
GalNAc-ol I	H-1	N.D.	N.D.	3.820	3.81	N.D.	3.81	3.81	N.D.	3.808	3.786	N.D.
	H-1′	N.D.	N.D.	3.753	3.76	N.D.	3.75	3.75	N.D.	3.739	3.729	N.D.
	H-2	4.395	4.399	4.384	4.389	4.376	4.388	4.388	4.381	4.374	4.363	4.363
	H-3	4.066	4.091	4.070	4.076	4.059	4.063	4.070	4.061	4.066	4.049	4.050
	H-4	3.507	3.522	3.496	3.498	3.531	3.535	3.440	N.D.	3.522	3.571	N.D.
	H-5	4.195	4.162	4.124	4.191	4.145	4.149	4.248	4.249	4.238	4.185	4.189
	H-6	N.D.	N.D.	3.670	3.65	3.73	3.64	3.857	N.D.	3.855	3.858	N.D.
	H-6′	3.63	N.D.	3.623	3.65	3.73	3.64	3.491	N.D.	3.481	3.446	N.D.
	Nac	2.050	2.046	2,046	2.045	2.048	2,048	2,047	2.049	2.041	2.038	2.042
Gal(β1-3) II	H-1	4.478	4.584	4.544	4.549	4.562	4.569	4.476	4.478	4.540	4.561	4.568
	H-2	3.564	N.D.	3.601	3.613	3.409	3.417	3.572	N.D.	3.605	3.415	N.D.
	H-3	N.D.	N.D.	4.118	4.135	4.166	4.179	N.D.	N.D.	4.115	4.163	4.176
	H-4	3.901	3.924	3.929	3.937	4.089	4.100	N.D.	N.D.	3.927	4.092	4.099
	H-5	N.D.	N.D.	N.D.	N.D.	3.768	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
GalNAc(β1-4) III	H-1	–	–	–	–	4.729	4.738	–	–	–	4.704	4.734
	H-2	–	(-	–	–	3.927	3.911	–	–	–	3.944	N.D.
	H-3	–	–	–	–	3.684	N.D.	–	–	–	3.679	N.D.
	H-4	–	–	–	–	3.924	N.D.	–	–	–	3.914	3.918
	H-5	–	–	–	–	3.511	N.D.	–	–	–	N.D.	N.D.
	H-6,6′	–	–	–	–	N.D.	N.D.	–	–	–	N.D.	N.D.
	NAc	–	–	–	–	2.028	2,031	–	–	–	2.025	2.031
Fuc(α1-2)	H-1	–	5.257	–	–	–	–	–	–	–	–	–
	H-5	–	4.278	–	–	–	–	–	–	–	–	–
	H-6	–	1.244	–	–	–	–	–	–	–	–	–
NeuAc or NeuGc (α2–3)	H-3ax	–	–	1.798	1.818	1.928	1.947	–	–	1.798	1.933	1.951
	H-3eq	–	–	2.770	2.791	2.683	2.705	–	–	2.771	2.683	2.700
	H-4	–	–	3.690	3.808	3.783	3.867	–	–	3.688	3.764	N.D.
	H-5	–	–	3.860	3.956	3.817	N.D.	–	–	3.858	3.828	N.D.
	NAc	–	–	2.033	–	2.032	–	–	–	2.030	2.030	–
	NGc	–	–	–	4.122	–	4.122	–	–	–	–	4.120
NeuAc or NeuGc (α2–6)	H-3ax	–	–	–	–	–	–	1.693	1.711	1.691	1.708	1.725
	H-3eq	–	–	–	–	–	–	2.731	2.746	2.720	2.734	2.750
	H-4	–	–	–	–	–	–	3.672	N.D.	3.674	3.666	N.D.
	H-5	–	–	–	–	–	–	N.D.	N.D.	3.820	3.837	N.D.
	NAc	–	–	–	–	–	–	2.032	–	2.030	2.030	–
	NGc	–	–	–	–	–	–	–	4.123	–	–	4.120

43 *vs* **45** and **44** *vs* **46**) clearly showed a constant increment of the GlcNAc(β1-6) II′ H-1 value by 0.023-0.026 ppm. This observation was used as an additional mean to distinguish NeuAc(α2-3)Gal(β1-3) from Cad type epitope containing glycans (Table 3).

The structure of compounds **27**, **32**, **33**, **34** and **40** established the presence of LacdiNAc sequences in *T. alpestris* mucins. It also suggested the occurrence of two distinct organizations of *O*-glycans upper branch depending on the insertion or not of a β-1-3 substitued GlcNAc on the GlcNAc(β1-6)GalNAc-ol core structure. As will be demonstrated, both types of upper branches can be further elongated by addition of LacdiNAc units and will be refered to as poly LacdiNAc **type I** for the (GalNAcβ1-4GlcNAcβ1-3)$_n$GalNAcβ1-4GlcNAcβ1-6GalNAc-ol sequence and **type II** for the (GalNAcβ1-4GlcNAcβ1-3)$_n$GlcNAcβ1-6GalNAc-ol sequence. It is noteworthy that elongated poly LacdiNAc sequences were systematically terminally capped either by sulphate group or by sialic acid, whereas no neutral poly LacdiNAc was observed.

Compounds **35** and **36** represented the smallest sulphated LacNAc containing *O*-glycans of **type I** and **type II**, respectively. Compounds **35** and **36** ^{1}H NMR parameters only differed from those of compounds **32** and **33** by the large scale downfield shifts of their H3- and H4-GalNAc III′ signals from δ3.74 and 3.94 ppm to δ3.915 and 4.690 ppm, respectively (Table 6). These parameters clearly typified compounds **35** and **36** as sulphated equivalents of compounds **32** and **33**, respectively in GalNAc III′ C-4 position. Starting from these two compounds a

Table 6 ^{1}H-NMR chemical shifts of constituent monosaccharide for oligosaccharide-alditols with the Gal(β1-3) [GlcNAc(β1-6]GalNAc-ol (core2) and GlcNAc(β1-3)GalNAc-ol (core3) structures

		Core 2									
Residue	Reporter group	**11**	**20**	**21**	**27**	**29**	**30**	**32**	**33**	**34**	**40**
GalNAc-ol **I**	H-1	N.D.	N.D.	3.81	3.81	3.82	3.81	3.815	3.808	N.D.	3.818
	H-1′	N.D.	N.D.	3.75	3.76	3.76	3.76	3.752	3.749	N.D.	3.764
	H-2	4.393	4.402	4.388	4.385	4.387	4.385	4.380	4.384	4.379	4.390
	H-3	4.060	4.083	4.070	4.063	4.056	4.057	4.049	4.067	4.050	4.060
	H-4	3.47	3.051	3.440	3.433	3.450	3.450	3.455	3.437	3.456	3.428
	H-5	4.279	4.252	4.268	4.264	4.187	4.193	4.193	4.247	4.192	4.166
	H-6	3.93	N.D.	3.921	3.909	3.910	3.912	3.906	3.905	N.D.	3.893
	H-6′	N.D.	N.D.	3.68	3.677	3.681	3.681	3.672	3.656	N.D.	3.658
	NAc	2.066	2.055	2.064	2.068	2.072	2.072	2.068	2.068	2.068	2.070
Gal(β1-3)**II**	H-1	4.465	4.570	4.539	4.528	4.558	4.567	4.559	4.532	4.560	4.563
	H-2	N.D.	N.D.	3.607	3.603	3.436	3.419	3.413	3.603	3.422	3.417
	H-3	N.D.	N.D.	4.117	4.113	4.181	4.181	4.164	4.115	4.181	4.169
	H-4	3.90	3.93	3.931	3.930	4.085	4.108	4.096	3.930	4.097	4.107
GalNAc(β1-4) **III**	H-1	–	–	–	–	4.680	4.726	4.718	–	4.712	4.708
	H-2	–	–	–	–	3.936	3.951	3.947	–	3.94	3.937
	H-3	–	–	–	–	3.699	3.686	3.681	–	N.D.	3.678
	H-4	–	–	–	–	3.913	3.926	3.927	–	4.925	3.927
	H-5	–	–	–	–	N.D.	N.D.	3.51	–	N.D.	N.D.
	H-6,6′	–	–	–	–	N.D.	N.D.	3.73	–	N.D.	N.D.
	NAc	–	–	–	–	2.038	2.038	2.034	–	2.036	2.046
GlcNAc(β1-6) **II′**	H-1	4.538	4.549	4.534	4.536	4.558	4.559	4.559	4.483	4.560	4.506
	H-2	N.D.	N.D.	3.701	3.74	3.703	3.701	3.740	3.775	–	3.777
	H-3	N.D.	N.D.	3.516	3.74	3.533	3.535	3.75	3.716	–	3.716
	H-4	N.D.	N.D.	3.45	3.65	3.439	3.441	3.64	3.511	–	3.50
	H-5	N.D.	N.D.	3.745	N.D.	3.469	N.D.	3.532	3.455	–	3.46
	H-6	N.D.	N.D.	3.931	N.D.	3.938	N.D.	3.84	3.924	–	3.92
	H-6′	N.D.	N.D.	3.67	N.D.	3.751	N.D.	3.67	3.746	–	3.75
	NAc	2.066	2.055	2.067	2.062	2.092	2.089	2.068	2.064	2.068	2.111
GlcNAc(β-3) **II″**	H-1	–	–	–	–	–	–	–	4.568	–	4.599
	H-2	–	–	–	–	–	–	–	3.710	–	3.718
	H-3	–	–	–	–	–	–	–	3.751	–	3.72
	H-4	–	–	–	–	–	–	–	3.636	–	N.D.
	H-5	–	–	–	–	–	–	–	3.527	–	3.50
	H-6	–	–	–	–	–	–	–	3.849	–	3.85
	H-6′	–	–	–	–	–	–	–	3.67	–	3.64
	NAc	–	–	–	–	–	–	–	2.007	–	2.009
GalNAc(β1-4) **III′**	H-1	–	–	–	4.516	–	–	4.527	4.514	4.525	4.513
	H-2	–	–	–	3.927	–	–	3.928	3.928	3.94	3.926
	H-3	–	–	–	3.744	–	–	3.747	3.742	N.D.	3.742
	H-4	–	–	–	3.93	–	–	3.945	3.939	3.925	3.938
	H-5	–	–	–	N.D.	–	–	3.51	3.523	N.D.	N.D.
	H-6,6′	–	–	–	N.D.	–	–	3.73	3.73	N.D.	N.D.
	NAc	–	–	–	2.057	–	–	2.073	2.079	2.075	2.075
Fuc(α1-2)	H-1	–	5.221	–	–	–	–	–	–	–	–
	H-5	–	4.274	–	–	–	–	–	–	–	–
	H-6	–	1.243	–	–	–	–	–	–	–	–

Continued

Table 6 (*Continued*)

Residue	Reporter group	Core 2									
		11	**20**	**21**	**27**	**29**	**30**	**32**	**33**	**34**	**40**
NeuAc(α2-3)	H-3ax	–	–	1.801	1.799	1.957	1.960	1.942	1.802	1.959	1.946
NeuGc(α2-3)	H-3eq	–	–	2.775	2.774	2.710	2.705	2.682	2.775	2.706	2.690
	H-4	–	–	3.692	3.692	3.76	3.871	3.767	3.690	N.D.	3.76
	H-5	–	–	N.D.	3.87	3.82	3.64	N.D.	3.866	N.D.	3.82
	H-6	–	–	N.D.	N.D.	3.500	N.D.	N.D.	3.667	N.D.	3.500
	H-7	–	–	N.D.	N.D.	3.605	N.D.	N.D.	N.D.	N.D.	N.D.
	H-8	–	–	N.D.	N.D.	3.86	N.D.	N.D.	N.D.	N.D.	N.D.
	H-9	–	–	N.D.	N.D.	3.858	N.D.	N.D.	N.D.	N.D.	N.D.
	H-9′	–	–	N.D.	N.D.	3.605	N.D.	N.D.	N.D.	N.D.	N.D.
	NAc	–	–	2.033	2.033	2.039	–	2.029	2.034	–	2.035
	NGc	–	–	–	–	–	4.122	–	–	4.122	–

		35	**36**	**41**	**42**	**43**	**44**	**45**	**46**
GalNAc-ol **I**	H-1	3.80	3.80	3.801	3.794	3.81	3.806	3.81	3.806
	H-1′	3.76	3.76	3.752	3.741	3.76	3.761	3.76	3.761
	H-2	4.383	4.383	4.385	4.369	4.386	4.387	4.386	4.387
	H-3	4.046	4.065	4.063	4.047	4.062	4.061	4.062	4.061
	H-4	3.465	3.440	3.432	3.458	3.434	3.434	3.434	3.434
	H-5	4.26	4.194	4.263	4.192	4.249	4.247	4.166	4.177
	H-6	3.80	3.80	3.905	3.900	3.902	3.888	3.899	3.888
	H-6′	3.68	3.68	3.673	3.668	3.655	3.654	3.656	3.654
	NAc	2.07	2.07	2.068	2.068	2.067	2.067	2.067	2.067
Gal (β1-3) **II**	H-1	4.554	4.530	4.526	4.561	4.531	4.532	4.561	4.564
	H-2	3.416	3.604	3.602	3.417	3.600	3.603	3.415	3.421
	H-3	4.165	4.115	4.112	4.163	4.114	4.126	4.167	4.177
	H-4	4.093	3.928	3.926	4.092	3.929	3.932	4.094	4.099
GalNAc (β1-4) **III**	H-1	4.719	–	–	4.728	–	–	4.729	4.712
	H-2	3.947	–	–	3.913	–	–	3.925	3.937
	H-3	N.D.	–	–	3.679	–	–	3.68	3.68
	H-4	N.D.	–	–	3.914	–	–	3.91	3.91
	H-5	N.D.	–	–	3.51	–	–	N.D.	N.D.
	H-6,6′	N.D.	–	–	3.73	–	–	N.D.	N.D.
	NAc	2.044	–	–	2.038	–	–	2.042	2.047
GlcNAc (β1-6) **II**′	H-1	4.541	4.483	4.526	4.550	4.481	4.48	4.506	4.508
	H-2	3.73	3.775	3.717	3.735	3.770	3.771	3.770	3.771
	H-3	N.D.	N.D.	3.72	3.73	3.72	3.72	3.72	3.72
	H-4	N.D.	N.D.	3.638	3.62	3.52	3.52	N.D.	N.D.
	H-5	N.D.	N.D.	3.494	3,501	3.50	3.50	3.50	3.50
	H-6	N.D.	N.D.	3.84	N.D.	N.D.	N.D.	N.D.	N.D.
	H-6′	N.D.	N.D.	3.72	N.D.	N.D.	N.D.	N.D.	N.D.
	NAc	N.D.	2.07	2.061	2.068	2.107	2.107	2.107	2.107
GlcNAc (β1-3) **II**′′	H-1	–	4.556	–	–	4.562	4.562	4.562	4.562
	H-2	–	3.72	–	–	3.72	3.713	3.72	3.713
	H-3	–	N.D.	–	–	3.72	3.72	3.72	3.72
	H-4	–	N.D.	–	–	3.62	3.62	3.62	3.62
	H-5	–	N.D.	–	–	3.52	3.52	3.52	3.52
	NAc	–	2.006	–	–	2.003	2.005	2.003	2.005

Continued

Table 6 (*Continued*)

		Core 2							
Residue	Reporter group	**35**	**36**	**41**	**42**	**43**	**44**	**45**	**46**
GalNAc (β1-4) **III**′	H-1	4.600	4.581	4.507	4.520	4.506	4.505	4.506	4.505
	H-2	3.929	3.929	3.975	3.971	3.973	3.975	3.973	3.975
	H-3	3.915	3.915	3.806	3.811	3.806	3.806	3.806	3.806
	H-4	4.690	4.690	4.158	4.159	4.157	4.155	4.157	4.155
	H-5	N.D.	N.D.	N.D.	N.D.	3.72	3.72	3.72	3.72
	H-6,6′	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
	NAc	2.07	2.07	2.014	2.012	2.013	2.014	2.013	2.014
GlcNAc (β1-3) **IV**′	H-1	–	–	4.566	4.568	4.586	4.587	4.586	4.587
	H-2	–	–	3.731	3.735	3.71	3.712	3.71	3.712
	H-3	–	–	3.73	3.73	3.73	3.73	3.73	3.73
	H-4	–	–	3.638	3.64	3.64	3.64	3.64	3.64
	H-5	–	–	3.494	3.501	3.50	3.50	3.50	3.50
	H-6	–	–	3.84	N.D.	N.D.	N.D.	N.D.	N.D.
	H-6′	–	–	3.72	N.D.	N.D.	N.D.	N.D.	N.D.
	NAc	–	–	2.053	2.068	2.073	2.073	2.073	2.073
GalNAc (β1-4) **V**′	H-1	–	–	4.587	4.586	4.587	4.585	4.587	4.585
	H-2	–	–	3.924	3.933	3.933	3.933	3.914	3.915
	H-3	–	–	3.896	3.899	3.896	3.895	3.896	3.895
	H-4	–	–	4.690	4.690	4.690	4.688	4.690	4.688
	H-5	–	–	3.83	3.83	3.83	3.83	3.83	3.83
	H-6,6′	–	–	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
	NAc	–	–	2.068	2.068	2.067	2.067	2.067	2.067
NeuAc (α2-3)	H-3ax	1.941	1.802	1.798	1.937	1.802	1.808	1.940	1.957
NeuGc (α2-3)	H-3eq	2.685	2.773	2.775	2.683	2.777	2.784	2.685	2.706
	H-4	3.768	3.694	3.688	3.764	3.690	3.785	3.770	3.881
	H-5	N.D.	N.D.	3.862	3.828	3.86	N.D.	3.83	N.D.
	H-6	N.D.	N.D.	3.73	3.49	3.73	N.D.	3.52	N.D.
	NAc	2.034	2.034	2.033	2.033	2.033	–	2.033	–
	NGc	–	–	–		–	4.122	–	4.122

		Core 3						
		2	**12**	**13**	**14**	**15**	**31**	**38**
GalNAc-ol **I**	H-1	N.D.	3.61	N.D.	3.61	N.D.	3.60	3.60
	H-1′	N.D.	3.59	N.D.	3.59	N.D.	3.60	3.60
	H-2	4.288	4.261	4.261	4.262	4.288	4.262	4.262
	H-3	3.988	3.987	3.988	3.982	4.012	3.993	3.993
	H-4	3.55	3.61	3.61	3.599	N.D.	3.625	3.625
	H-5	4.142	4.186	4.186	4.181	4.137	4.174	4.174
	H-6	3.65	3.82	N.D.	3.823	N.D.	3.821	3.821
	H-6′	N.D.	3.48	3.48	3.467	N.D.	3.462	3.462
	NAc	2.037	2.032	2.032	2.030	2.034	N.D.	N.D.
Gal (β1-3) **III**	H-1	–	–	–	–	4.460	4.505	4.528
	H-2	–	–	–	–	N.D.	3.672	3.352
	H-3	–	–	–	–	N.D.	3.793	4.097
	H-4	–	–	–	–	3.93	4.186	4.085
GalNAc (β1-4) **IV**	H-1	–	–	–	–	–	–	4,727
	H-2	–	–	–	–	–	–	3,907
	H-3	–	–	–	–	–	–	3,697
	H-4	–	–	–	–	–	–	3,919
	H-5	–	–	–	–	–	–	N.D.
	H-6,6′	–	–	–	–	–	–	N.D.
	NAc	–	–	–	–	–	–	N.D.

Continued

Table 6 (*Continued*)

Residue	Reporter group	Core 3						
		2	**12**	**13**	**14**	**15**	**31**	**38**
GlcA (β1-6) **IV**	H-1	–	–	–	–	–	4.667	–
	H-2	–	–	–	–	–	3.416	–
	H-3	–	–	–	–	–	3.51	–
	H-4	–	–	–	–	–	3.51	–
	H-5	–	–	–	–	–	3.719	–
GlcNAc (β1-3) **II**	H-1	4.604	4.608	4.608	4.603	4.654	4.655	4.643
	H-2	N.D.	3.75	N.D.	3.745	N.D.	3.912	3.891
	H-3	N.D.	3.58	N.D.	3.578	N.D.	3.791	3.815
	H-4	N.D.	3.46	N.D.	3.46	N.D.	N.D.	3.58
	H-5	N.D.	3.48	N.D.	3.46	N.D.	3.501	3.501
	H-6	3.950	3.94	3.95	3.939	3.95	3.939	3.939
	H-6′	N.D.	3.75	N.D.	3.752	N.D.	3.770	3.770
	NAc	2.085	2.079	2.079	2.077	2.073	N.D.	N.D.
NeuAc NeuGc Kdn (α2–6)	H-3ax	–	1.697	1.715	1.657	–	1.656	1.656
	H-3eq	–	2.733	2.748	2.677	–	2.676	2.676
	H-4	–	3.67	N.D.	3.603	–	3.596	3.596
	H-5	–	–	–	3.522	–	3.523	3.523
	H-6	–	–	–	3.670	–	3.66	3.66
	H-7	–	–	–	3.86	–	N.D.	N.D.
	H-8	–	–	–	3.89	–	N.D.	N.D.
	H-9	–	–	–	3.92	–	N.D.	N.D.
	H-9′	–	–	–	3.698	–	N.D.	N.D.
	NAc	–	2.034	–	–	–	–	–
	NGc	–	–	4.122	–	–	–	–
Kdn (α2-3)	H-3ax	–	–	–	–	–	–	1.862
	H-3eq	–	–	–	–	–	–	2.634
	H-4	–	–	–	–	–	–	3.679
	H-5	–	–	–	–	–	–	3.523
	H-6	–	–	–	–	–	–	3.45

large family of sulphated poly LacdiNAc compounds was further identified. In particular, the sulphated **type I** GalNAc(β1-4)GlcNAc(β1-3)GalNAc(β1-4)GlcNAc(β1-6) sequence was observed in both compounds **41** and **42**. **41** and **42** that differed only by the nature of their C-3 lower branches, NeuAc(α2-3)Gal(β1-3) or Cad type, respectively (Table 2). ESI-MS analysis established the molecular mass of compound **42** to 1771 Da, which indicated it contained 1 HexNAc-ol, 5 HexNAc, 1 Hex, 1 NeuAc and 1 sulphate group (Table 4). In addition to the Cad epitope in C-3 position of the GalNAc-ol unit, two-step relayed COSY indicated the presence of two β-GlcNAc units (II′ and IV′) and two β-GalNAc units (III′ and V′) owing to the position of their respective H-2 atom resonances, and the ^3J coupling constants deduced from the shape of their H-2, H-3 and H-4 signals (Fig. 4). The ^{13}C-^{1}H HMQC spectrum of **42** (Fig. 4) clearly showed that GlcNAc II′ and IV′ were substituted in C-4 position (δ^{13}C-4 at 80.5 and 80.6, respectively) and that GalNAc III′ was substituted in C-3 position (δ^{13}C-4 at 80.3). Then, the observation of the GalNAc V′ H-4 signal strongly shifted towards low-field at δ4.690 indicated the presence of a terminal HSO_3(4)GalNAc unit. Taken together, these observations established the nature of the C-6 branch of compound **42** as **type I** HSO_3(4)GalNAc(β1-4)GlcNAc(β1-3)GalNAc(β1-4)GlcNAc(β1-6). Identical ^{1}H and ^{13}C parameters were observed in compound **41**, which confirmed that both compounds **41** and **42** presented identical upper branches.

The sulphated **type II** GalNAc(β1-4)GlcNAc(β1-3)GalNAc(β1-4)GlcNAc(β1-3)GlcNAc(β1-6) sequence was observed in all compounds **43**, **44**, **45** and **46**. ESI-MS analyses permitted to calculate molecular mass of compounds **43** and **45** to 1771 and 1974, respectively (Table 4). Combining the results of mass spectrometry and NMR analyses, compounds **43** and **45** appeared to be octa- and nonasaccharides containing Gal, GlcNAc, GalNAc, NeuAc, GalNAc-ol and sulphate in the ratios 1:3:2:1:1:1 and 1:3:3:1:1:1, respectively. From MS and NMR data, it was established that compounds **43** and **45** only differ from the nature of their C-3 lower branches, NeuAc(α2-3)Gal(β1-3)

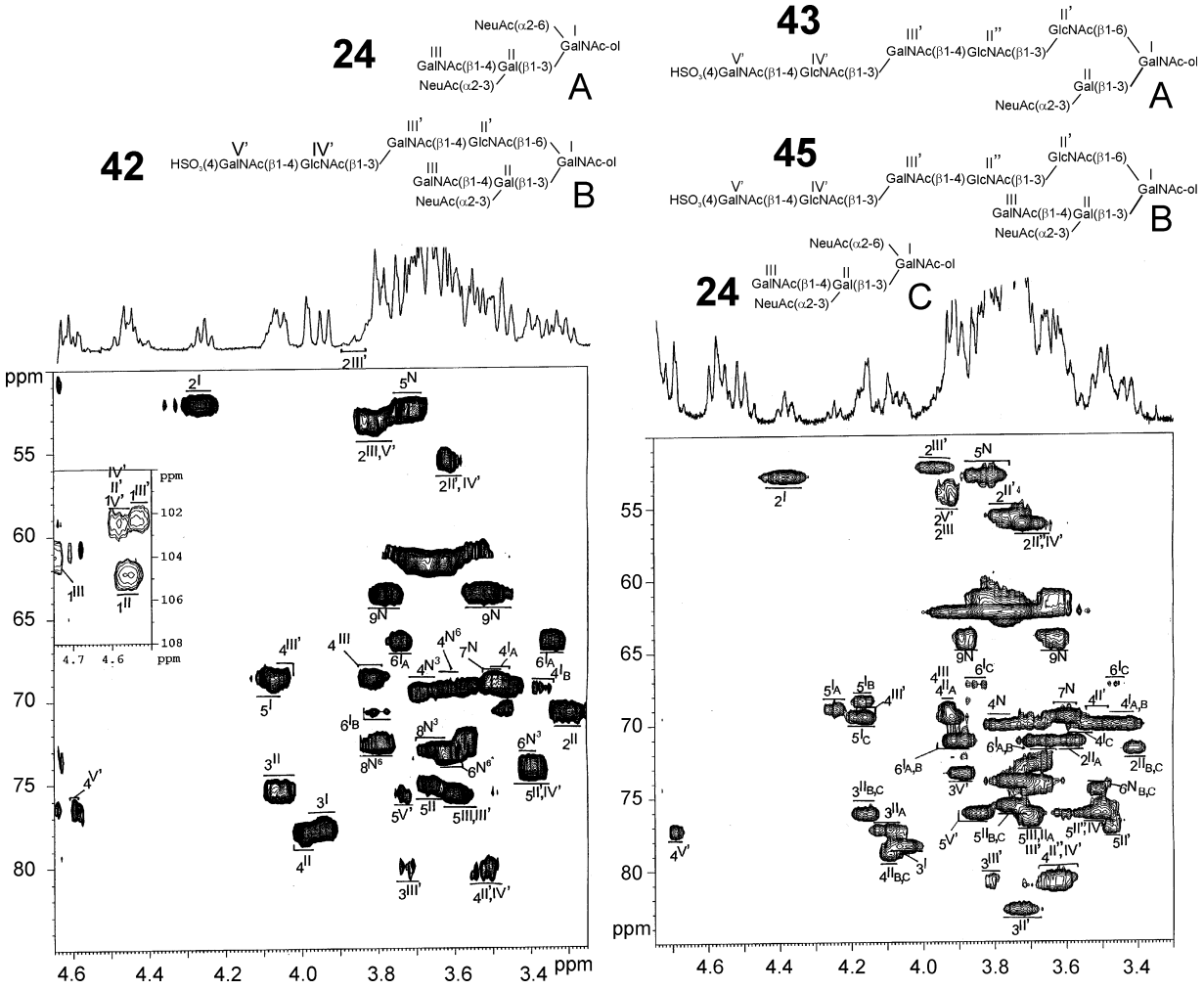


Fig. 4 HMQC spectra of mixtures **24**, **42** (left) and **24**, **43**, **45** (right)

or Cad type, respectively. Comparison of NMR data from **41/42** and **43/45** permitted to identify the additional GlcNAc signals present in **43/45** as a GlcNAc(β1,3) II″ unit, accordingly to parameters already defined for compounds **33**, **36** and **40** (Table 6). In order to verify the above proposed sequence, the ESI-MS/MS spectra of compounds **43** and **45** were compared (Fig. 5). In negative mode, only negatively charged fragments ions, including sulphated fragments, were observed as $[M\text{-}H]^-$ adducts. In agreement with NMR data, ESI-MS/MS spectra of both compounds **43** and **45** showed Y1β and Z1β ions at *m/z* 1317 and 1299 that demonstrated the presence of a sulphated $HexNAc_5$ sequence linked to the GalNAc-ol unit. As shown Fig. 5, the HSO_3-$HexNAc_5$ sequence and its recurrent degradation products from reducing end were also directly observed owing to the B/C5α to B/C2α ions series. In order to unambiguously confirm the existence of the elongated **type II** poly LacdiNAc sequence, compound **43** was oxidised in very mild conditions that exclusively cleaves C4-C5 linkage of the reduced GalNAc unit. ESI-MS/MS analysis in negative mode of the oxidation products showed an intense signal at *m/z* 1156 tentatively attributed to the $[M\text{-}H]^-$ adduct ion of HSO_3-$HexNAc_5$-CHOH-CH_2OH resulting from the liberation of the C-6 branch with GalNAc-ol C-5 and C-6 carbons (Fig. 6). The observation of this ion definitively established the presence of a HSO_3-$HexNAc_5$ sequence in C-6 position of GalNAc-ol. Fragmentation pattern of this sequence in ESI-MS/MS confirmed its nature with a CHOH-CH_2OH aglycon in reducing end and a sulphate group in non reducing end (Fig. 6). By comparison, identical experiments conducted on compound **41** showed an $[M\text{-}H]^-$ ion at *m/z* 953 (Fig. 6), indicative of the generation of a HSO_3-$HexNAc_4$-CHOH-CH_2OH product. Its fragmentation in ESI-MS/MS established the nature of the upper branch of **41** as a sulphated di LacdiNAc structure with an even number of HexNAc residues. These experiments confirmed the existence of two different organizations, **type I** and **type II**, of poly LacdiNAc sequences present in C-6 position of *O*-glycans. Taken together, data from NMR and mass

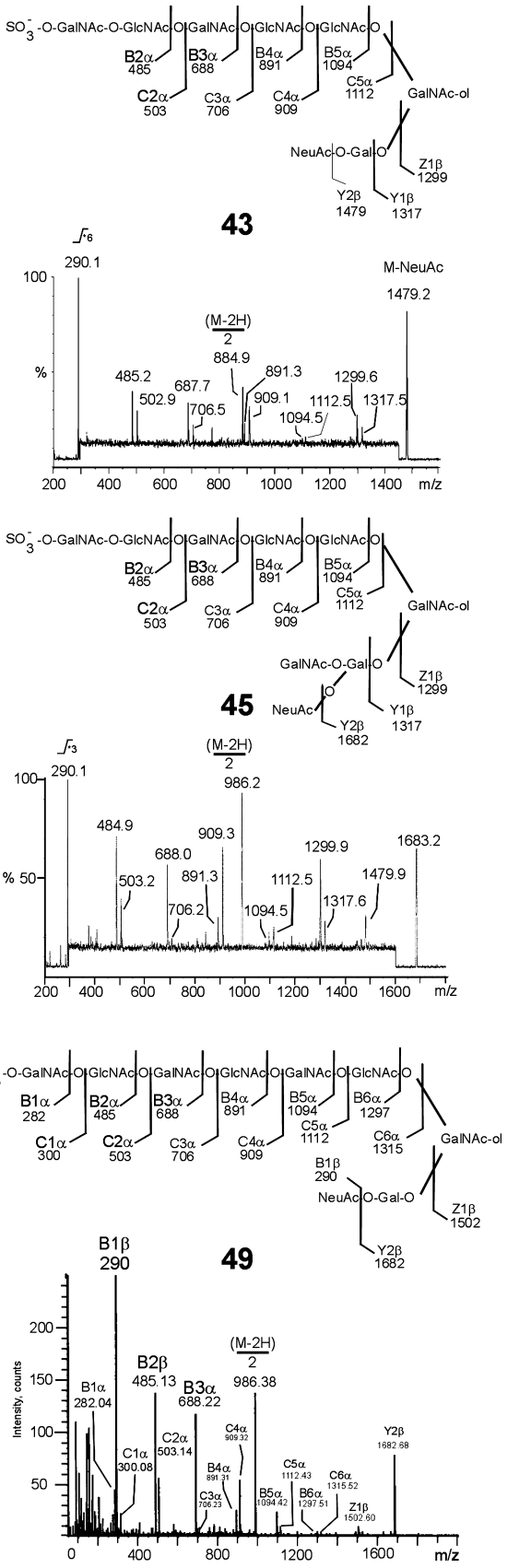


Fig. 5 Fragmentation patterns in negative mode by ESI-MS/MS of compounds **43**, **45** and **49**

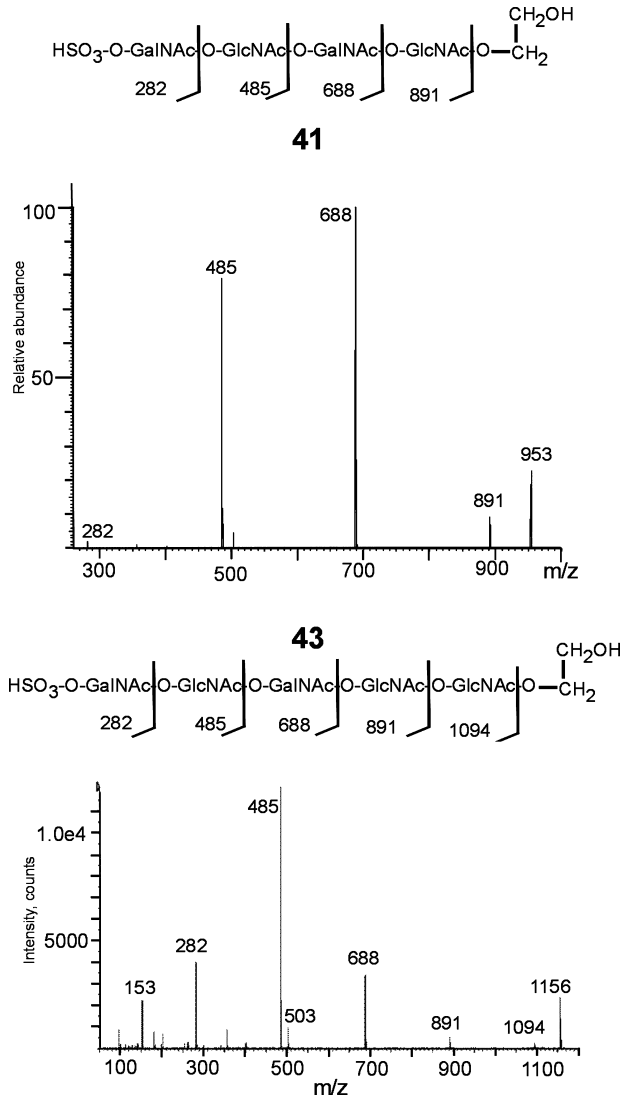


Fig. 6 Fragmentation patterns in negative mode by ESI-MS/MS of C-6 upper branch of compounds **41** and **43** after mild periodate oxidation and reduction

spectrometry permitted to establish structures of compounds **43** and **45,** as shown in Scheme 1. Then, comparison of the ^{1}H NMR and mass spectrometry data of **44** and **46** established that these compounds possessed similar structures as **43** and **46**, with NeuGc instead of NeuAc attached at the terminal position of the lower branch, as proved by the presence of the –CO-CH_2OD ^{1}H signal at δ4.122 ppm in ^{1}H-NMR and by the mass increment of 16 *a.m.u.* (Tables 4 and 6).

The exhaustive structural deciphering of above poly LacdiNAc containing *O*-glycans by NMR spectroscopy enabled us to establish specific sets of structural-reporter-groups that are reported in Table 3. As already noticed, some of these parameters are influenced by the nature of the lower branch the at GalNAc-ol C-3 position. These reporter groups were used to analyse compounds with structure of higher complexity. Particularly, the starting point of

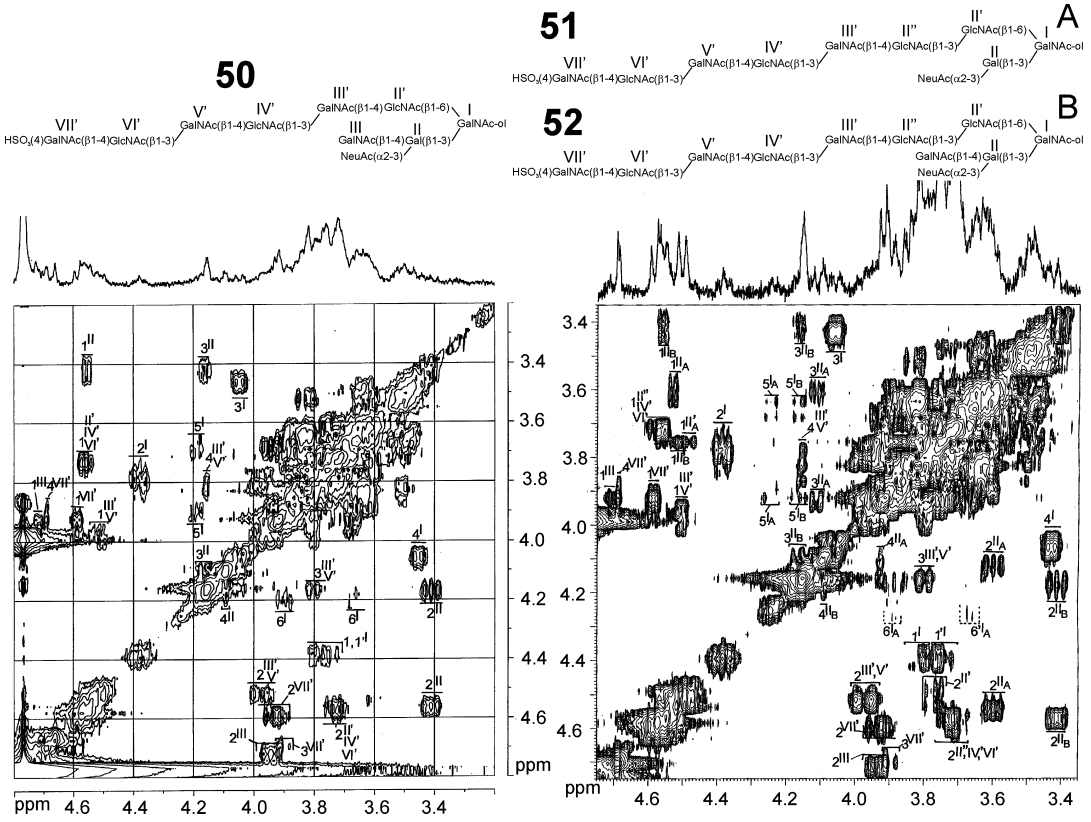


Fig. 7 COSY spectra of compounds **50** (left) and **51**, **52** (right)

sequences **type I** GalNAc(β1-4)**GlcNAc**(β1-6) and **type II** GalNAc(β1-4)**GlcNAc**(β1-3)**GlcNAc**(β1-6) can be easily recognised owing to the characteristic GlcNAc(β1-6) II′ H-1 and H-2 signals indicated in bold in Table 3. The **type II** poly LacdiNAc sequences can be distinguished from the **type I** by the fact that the 3-substituted GlcNAc II′ H-1 atom resonates at a more upfield position ($\Delta\delta$ − 0.05) than its 4-substituted homologue, whereas the other ^{1}H resonances, relative to the repeating unit HSO_3(4)(GalNAcβ1-4GlcNAcβ1-3)$_n$GalNAcβ1-4GlcNAc, are identical for both sequences **type I** and **II**. On this basis, upper branches of compounds **49** to **54** were characterised by NMR (Fig. 7) as sulphated **type I** (**49** and **50**) or **type II** (**51** to **54**) tri LacdiNAc sequences. Attributions were also controlled by ESI-MS/MS fragmentation, as exemplified for compound **49** (Fig. 6). As for compounds **43** and **45**, fragment ions Z1β at *m/z* 1502 and B/C6α at *m/z* 1297/1315 permit to directly observe the HSO_3-$HexNAc_6$ upper branch. Its recurrent degradation is then followed owing to fragment ions B/C5α to B/C1α.

Then, a minor family of even higher molecular mass (compounds **54** to **59**) was observed by ESI-MS from Bio-Gel P-6 fraction I. Low quantities of individual components coupled to high mass prevented their separation by HPLC and their further analysis. The observed molecular mass of compounds **54** to **59** permitted to calculate their individual compositions (Table 4). Their very high content in HexNAc (from 9 to 11) strongly suggested that they were substituted by poly LacdiNAc sequences. Based on the structure of the twenty one LacdiNAc containing *O*-glycans described above, we tentatively attributed the structures of compounds **54** to **59** to sulphated tetra and penta LacdiNAc containing *O*-glycans, as depicted in Scheme 1.

Beside sulphate groups, poly LacdiNAc sequences may also be terminally capped by sialic acid residues as observed in the two minor *O*-glycans **47** and **48**. These products were shown by MS analysis to have the following compositions: $NeuAc_2HexNAc_4Gal_1HexNAc$-ol and $NeuAc_2HexNAc_5Gal_1HexNAc$-ol, respectively (Table 4). ESI-MS/MS fragmentation of these compounds showed the presence of an upper NeuAc-$HexNAc_4$ branch owing to the characteristic Z1β ion at *m/z* 1307 (Fig. 8). Characteristic NeuAc H-3ax and H-3eq signals (Table 2), permitted to distinguish **47** from **48** according to the nature of the C-3 branch; NeuAc(α2-3)Gal or Cad type respectively. Surprisingly, no trace of α-2-6-linked NeuAc (the H-3ax and H-3eq signals of which being normally present at δ1.72 and 2.67) was observed, whereas the signals at δ1.80 and 2.77 suggested the α-2-3 sialylation of the terminal β-GalNAc unit.

Sia(α2-6)GalNAc-ol containing O-glycans

Sialylation of reducing GalNAc residue was observed in two distinct families of *O*-glycans. The first one was based on sialylated type 1 core Sia(α2-6)[Gal(β1-3)]GalNAc-ol, the simplest members observed being the ubiquitously found compounds **9** and **10**, NeuAc(α2-6)[Gal(β1-3)]GalNAc-ol and NeuGc(α2-6)[Gal(β1-3)]GalNAc-ol [8]. These two components might be further elongated in C-3 position by addition of GalNAc and sialic acid residues to form Sia(α2-3)Gal or Cad type sequences in disialylated compounds **19**, **24**, **25** and **26**. Compounds **19**, **24** and **26** were already described elsewhere [8] and were easily identified on the basis of their structural-reporter-group data reported in Table 5.

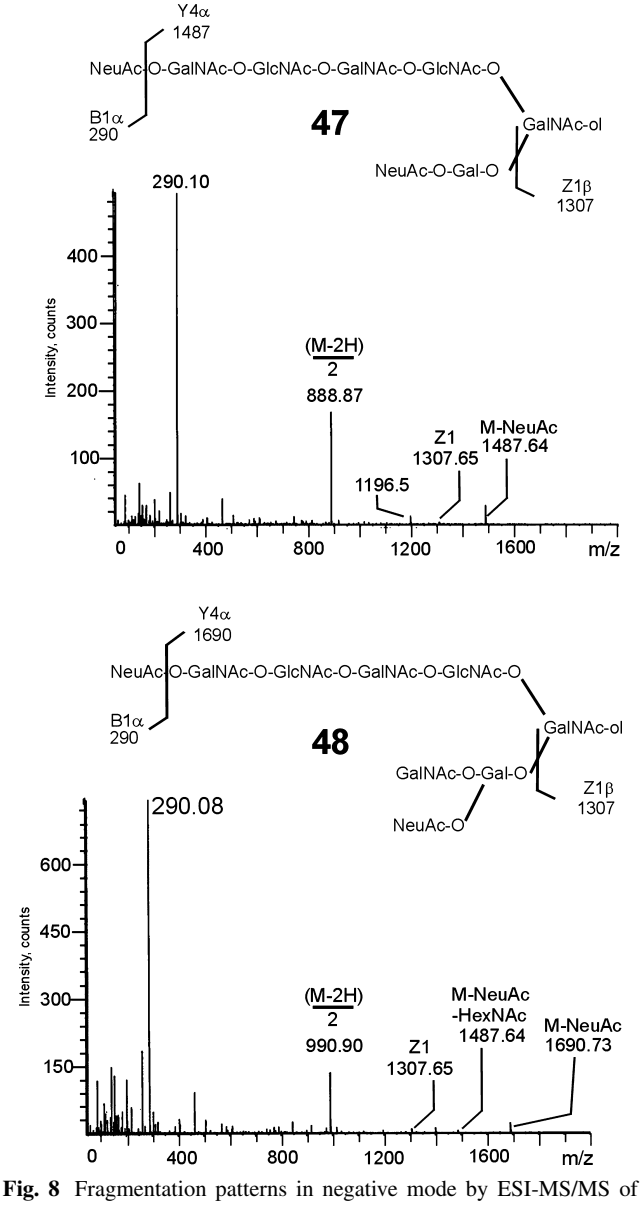


Fig. 8 Fragmentation patterns in negative mode by ESI-MS/MS of compounds **47** and **48**

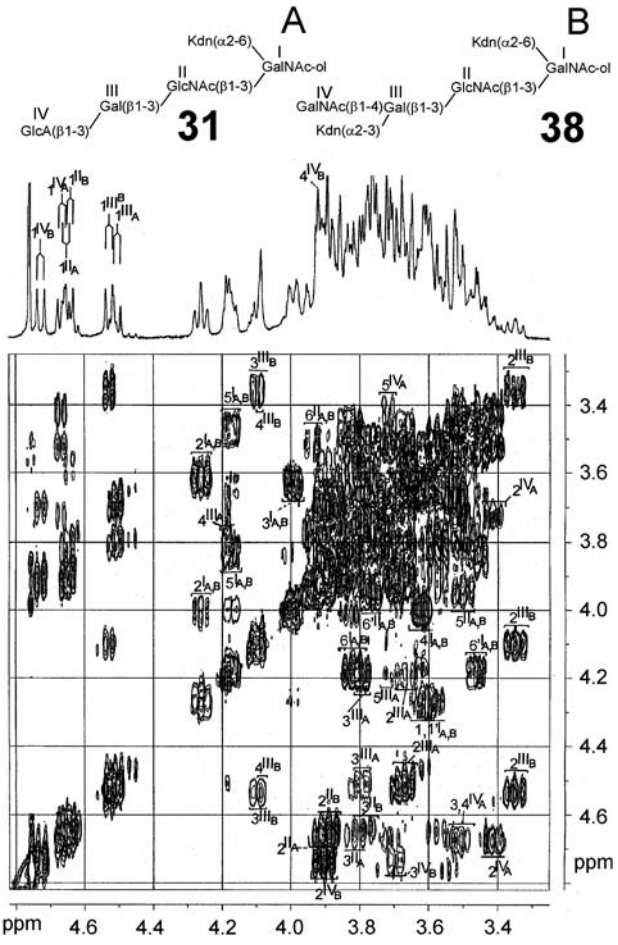


Fig. 9 Relayed COSY spectrum of the mixture **31**, **38**

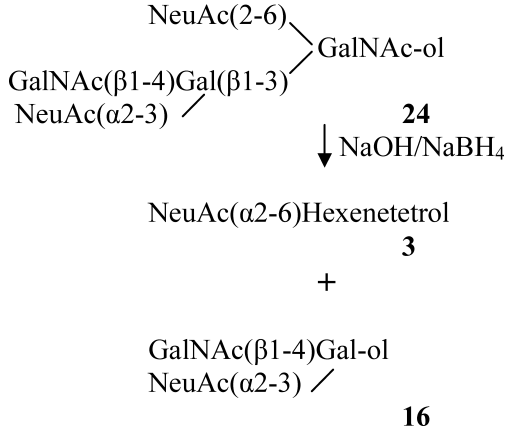
They were substituted either by two NeuAc or two NeuGc residues. On the other hand, molecular mass of compound **25** suggested the presence of both NeuAc and NeuGc in the same *O*-glycan (Table 4). The presence of H-3ax and H-3eq signals at δ1.72 and 2.74 was in favour of a NeuGc unit α-2-6-linked to GalNAc-ol, whereas the H-3ax signals of the second sialic acid unit were not distinguished from those of compounds **44** and **46** at $\delta \sim$1.95 (H-3ax NeuAc δ1.93-1.96; NeuGc δ1.95-1.96). These observations, coupled to comparison with structural reporter groups of Table 2, supported the original sequence proposed for the compound **25** in Scheme 1.

The second family of compounds was based on type 3 core and characterized by the presence of a Kdn residue, as exemplified by simple compound **14** Kdn(α2-6)[GlcNAc(β1-3)]GalNAc-ol. As previously observed [9–12], the presence of Kdn α-2-6-linked to GalNAc-ol, instead of NeuAc, is reflected by the absence of NAc signal at $\delta \approx 2.03$, whereas the H-3ax and H-3eq multiplets show upfield shifts of $\Delta\delta \approx -0.04$, as compared to the NeuAc analog. This structure might be elongated from C-3 branch to form either compound **31** or compound **38**. Compound **31** had a calculated composition of $HexA_1Hex_1HexNAc_1Kdn_1$

$HexNAc\text{-}ol_1$ (Table 4). The nature of the core 3 was deduced from the GalNAc-ol H-1,1′ signals at δ3.60 (Table 1) and the attachment of β-Gal at *O*-3 of the GlcNAc by the shifts of the GlcNAc H-2 and H-3 atom resonances from δ3.75 and 3.58 to 3.912 and 3.791. *Gluco* conformation of HexA residue was established by NMR whereas its β1-3 linkage to the β-Gal unit was determined owing to the β-Gal H-3 and H-4 parameters at δ3.793 and 4.186 [13]. Altogether, ^{1}H-NMR permitted to establish the full structure of compound **31** as depicted on Scheme 1. Compound **38** presented a distinct arrangement as suggested by its calculated composition of $Hex_1HexNAc_2Kdn_2HexNAc\text{-}ol_1$ (Table 4). In compound **38**, GlcA(β1-3)Gal(β1-3) motif of **31** was replaced by a Sd^a-antigen-like motif as evidenced by the occurrence of structural-reporter-group signals for Kdn H-3ax, Gal H-3, H-4 and GalNAc H-1, which are reported in Tables 1 and 6.

Oligosaccharides terminating in galactitol or 1,2,5,6-tetrahydroxy hex-3-ene (hexenetetrol) (Table 7)

As previously reported peeling reactions often proceeded during the reductive β-elimination and produce oligosaccharides terminating in either galactitol or hexenetetrol [14,15], and compounds **3**, **4**, **5**, **16** and **23** resulted from such reactions. From an earlier observation [16] we postulated that hexenetetrol was derived from GalNAc-ol itself, instead of Gal, as shown below:



Miscellaneous O-glycans (Table 7)

O-glycans containing GlcA or Fuc(α1–2)GlcA (compounds **22**, **28**) were already described in different amphibian species [16–18]. Another class of *O*-glycans was represented by compounds **37** and **39**, which contained the ubiquitous sequence Fuc(α1-2)Gal(β1-3)[Fuc(α1-2)]Gal(β1-3) also largely represented among different amphibian species [16–19].

Discussion

The complete structures of 59 *O*-glycans, ranging in size from disaccharides to tetradecasaccharides, were determined using NMR and ESI-MS/MS analyses (Scheme 1). Some of them, already described in numerous human and animal mucins, are relatively simple oligosaccharides which consist mainly of neutral and sialylated core 1, 2 and 3. Interestingly, the presence of three types of *O*-acetylated sialic acids -NeuAc, NeuGc and Kdn- was observed within the jelly coat. It is noteworthy that $Neu5,9Ac_2$ represented 56% of NeuAc derivatives, whereas Kdn was present in six forms, Kdn, Kdn9Ac, Kdn7Ac, Kdn5Ac, $Kdn8,9Ac_2$ and $Kdn7,9Ac_2$, in the ratio 42:39:6:2:7:5. Such a diversity has already been observed among other amphibian species. For example, the oviductal mucins of *Rana temporaria* and *Rana arvalis*, which are poorly sialylated (about 1% of the total sugar components) presented Kdn, Kdn9Ac, Kdn8Ac, Kdn7Ac, Kdn5Ac, $Kdn8,9Ac_2$, $Kdn7,9Ac_2$ and $Kdn4,5Ac_2$ in the ratios 45:23:1.5:7:4.5:4:1:2 and 37:38:0:3.5:0:0:3.5:1, respectively [unpublished results]. However, no systematic study of sialic acid *O*-acetylation in amphibians has been conducted so far.

Though Kdn is here exclusively observed in oligosaccharide sequences based on the core 3 structure (compounds **14**, **31** and **38**), we can not speculate that a Kdn transferase would specifically act on this type of acceptor. Indeed, the jellies of *T. alpestris* were studied as a whole, even if the number of jelly coat layers, secreted by different part of the oviduct, varies among species, from as many as six to as few as three. For these reasons, we rather consider that Kdn-containing components were probably synthetized in specific cells of the genital tract characterized by a defect in NeuAc and/or core 1 biosynthesis.

Kdn was first isolated from rainbow trout eggs [20] and then obtained from the capsular polysaccharide of *Klebsiella ozaenae* serotype K4 [21]. Kdn is now acknowledged to be an ubiquitous component of vertebrate glycoconjugates, owing to the remarkable contribution of Dr Yasuo Inoue in the field of Sialobiology [1,22,23]. It was later found as the sole sialic acid present in jelly coats of *Pleurodeles* and *Ambystoma* species [9–12]. In these cases, the Kdn unit formed a part of *O*-glycans based on type 1, 2 and 3 core structures. Among seven different species of the genus *Rana*, traces of Kdn were only observed in *Rana temporaria* [13]. Then, Kdn was also characterized as the sole sialic acid present in the oviducal mucins from the toad *Bombina bombina*, whereas only NeuGc was observed in the case of *Bombina variegata* [24]. From an evolutionary point of view, some hypotheses claim

Table 7 ^{1}H-NMR chemical shifts of miscellaneous oligosaccharide-alditols

Residues	Reporter groups	**3**	**4**	**16**	**5**	**23**	**37**	**39**	**22**	**28**
Hexitol Hexene-tetrol	H-1	N.D.	N.D.	–	–	–	3.790	3.78	3.785	3.774
GalNAc-ol **I**	H-1′	N.D.	N.D.	–	–	–	3.745	3.78	3.735	3.723
	H-2	4.24	4.24	–	–	–	4.236	4.221	4.394	4.413
	H-3	5.80	5.80	–	–	–	4.047	3.979	4.075	4.069
	H-4	5.80	5.80	–	–	–	3.623	3.620	3.457	3.446
	H-5	4.34	4.34	–	–	–	4.191	4.154	4.275	4.277
	H-6/H6′	N.D./N.D.	N.D./N.D.	–	–	–	3.925/3.698	3.925/3.700	3.927/3.684	3.934/3.690
	NAc	–	-	–	–	–	2.050	2.045	2.068	2.06
Gal(β1-3) **II**	H-1	–	-	–	–	–	4.687	4.688	4.520	4.450
	H-2	–	-	–	–	–	3.813	3.792	3.721	3.635
	H-3	–	-	–	–	–	4.178	4.170	3.815	3.783
	H-4	–	-	–	–	–	4.017	4.001	4.179	4.187
	H-5	–	-	–	–	–	N.D.	3.93	N.D.	N.D.
	H-6/H-6′	–	-	–	–	–	N.D./N.D.	3.97/3.88	N.D./N.D.	N.D./N.D.
Gal-ol **II**	H-1	–	–	3.702	N.D.	3.75	–	–	–	–
	H-1′	–	–	3.608	N.D.	3.71	–	–	–	–
	H-2	–	–	3.974	N.D.	4.050	–	–	–	–
	H-3	–	–	4.229	4.136	4.085	–	–	–	–
	H-4	–	–	4.157	3.921	3.751	–	–	–	–
	H-5	–	–	3.788	N.D.	4.361	–	–	–	–
	H-6/H-6′	–	–	3.61/3.61	N.D./N.D.	3.964/3.728	–	–	–	–
Gal(β1-3) **III**	H-1	–	–	–	–	4.665	4.899	4.901	–	–
	H-2	–	–	–	–	3.742	3.691	3.686	–	–
	H-3	–	–	–	–	3.89	3.89	3.890	–	–
	H-4	–	–	–	–	3.91	3.91	3.911	–	–
	H-5	–	–	–	–	N.D.	N.D.	3.67	–	–
	H-6,6′	–	–	–	–	N.D.	N.D.	3.75	–	–
GalNAc(β1-4) **III**	H-1	–	–	4.713	–	–	–	–	–	–
	H-2	–	–	3.927	–	–	–	–	–	–
	H-3	–	–	3.714	–	–	–	–	–	–
	H-4	–	–	3.930	–	–	–	–	–	–
	NAc	–	–	2.063	–	–	–	–	–	–
GlcA(β1-3) **III** or **IV**	H-1	–	–	–	–	–	–	–	4.670	4.717
	H-2	–	–	–	–	–	–	–	3.424	3.576
	H-3	–	–	–	–	–	–	–	3.52	3.731
	H-4	–	–	–	–	–	–	–	3.52	3.527
	H-5	–	–	–	–	–	–	–	3.725	N.D.

Continued

Table 7 (*Continued*)

Residues	Reporter groups	**3**	**4**	**16**	**5**	**23**	**37**	**39**	**22**	**28**
GlcNAc(β1-6) **IV**	H-1	–	–	–	–	4.577	–	4.577	–	–
	H-2	–	–	–	–	3.730	–	3.709	–	–
	H-3	–	–	–	–	3.547	–	3.531	–	–
	H-4	–	–	–	–	3.46	–	3.44	–	–
	H-5	–	–	–	–	3.46	–	3.44	–	–
	H-6/H-6′	–	–	–	–	3.960/3.756	–	3.935/3,724	–	–
	NAc	–	–	–	–	2.058	–	2.058	–	–
GlcNAc(β1-6) **II**′	H-1	–	–	–	–	–	4.559	4.558	4.539	4.542
	H-2	–	–	–	–	–	3.715	3.712	3.704	3.704
	H-3	–	–	–	–	–	3.534	3.531	3.520	3.522
	H-4	–	–	–	–	–	3.46	3.44	3.45	N.D.
	H-5	–	–	–	–	–	3.46	3.44	3.45	N.D.
	H-6/H-6′	–	–	–	–	–	3.937/3.743	3.935/3.724	3.932/3.744	N.D./N.D.
	NAc	–	–	–	–	–	2.050	2.065	2.068	2.068
Fuc(α1-2) **II**^{d}	H-1	–	–	–	–	5.036	5.414	5.394	–	–
	H-2	–	–	–	–	3.805	3.766	3.762	–	–
	H-3	–	–	–	–	3.80	N.D.	3.850	–	–
	H-4	–	–	–	–	3.80	N.D.	3.798	–	–
	H-5	–	–	–	–	4.053	4.277	4.276	–	–
	H-6	–	–	–	–	1.224	1.197	1.185	–	–
Fuc(α1-2) **II**^{e}	H-1	–	–	–	–	5.478	5.332	5.331	–	5.353
	H-2	–	–	–	–	3.800	3.800	3.797	–	3.792
	H-3	–	–	–	–	3.82	N.D.	3.733	–	3.860
	H-4	–	–	–	–	N.D.	N.D.	3.814	–	3.80
	H-5	–	–	–	–	4.292	4.333	4.325	–	4.613
	H-6	–	–	–	–	1.212	1.234	1.227	–	1.223
NeuAc or Kdn (α2-3)	H-3ax	–	–	1.805	1.757	–	–	–	–	–
or NeuAc, NeuGc, Kdn	H-3eq	–	–	2.76	2.762	–	–	–	–	–
(α2-6)	NAc	–	–	2.032	2.032	–	–	–	–	–
	H-3ax	1.704	1.754	–	–	–	–	–	–	–
	H-3eq	2.731	2.756	–	–	–	–	–	–	–
	NAc	2.031	–	–	–	–	–	–	–	–
	NGc	–	4.122	–	–	–	–	–	–	–

Table 8 ^{13}C-NMR chemical shifts of the constituent monosaccharides of the oligosaccharide-alditols released from *Triturus alpestris* egg jelly coat mucins. Chemical shifts of methyl from acetyl groups which resonated between 23.5 to 23.8 ppm

Residue	Reporter group	**24**	**32**	**33**	**39**	**42**	**43**	**45**
GalNAc-ol I	C-1	~62.2	62.5	62.5	62.1	62.5	62.5	~62.2
	C-2	52.9	53.1	52.9	54.3	52.7	52.9	52.9
	C-3	**78.1**	**78.4**	**77.4**	**76.1**	**78.0**	**78.1**	**78.1**
	C-4	69.9	70.3	70.2	71.1	69.9	69.9	69.9
	C-5	69.3	68.5	68.8	69.3	69.2	68.8	68.3
	C-6	**66.7**	**71.5**	**72.3**	**71.8**	**72.0**	**70.9**	**70.9**
Gal(β1-3) II	C-1	104.7	105.6	105.5	101.3	104.7	105.5	104.7
	C-2	71.5	71.5	69.8	**74.3**	70.9	71.0	71.5
	C-3	**75.9**	**76.2**	**77.2**	**80.4**	**75.6**	**77.0**	**75.9**
	C-4	**78.6**	**79**	68.9	71.0	**78.0**	69.4	78.6
	C-5	75.4	76.1	76.8	74.9	75.3	75.9	75.4
	C-6	~67.2	~62	62.4	**70.3**	~62	~62	~62
GalNAc(β1-4) III	C-1	104.1	104.5	–	101.9*	104.0	–	104.1
	C-2	54.2	54.2	–	**77.7***	53.9	–	54.2
	C-3	~74	73.1	–	74.6*	73.9	–	~74
Or	C-4	69	69.2	–	70.6*	68.9	–	69.0
*Gal(β1-3) III	C-5	75.9	76.7	–	76.6*	75.8	–	75.9
	C-6	~62	~62	–	62.1*	~62	–	~62
GlcNAc(β1-6) II′	C-1	–	103.3	103.6	102.5	102.7	102.6	102.7
	C-2	–	56.5	55.9	56.7	56.2	~55.7	~55.7
	C-3	–	73.9	**82.6**	75.3	74	**82.4**	**82.4**
	C-4	–	**80.8**	70.3	71.1	**80.5**	69.9	69.9
	C-5	–	76.2	77.0	77.1	75.3	76.9	76.9
	C-6	–	61.9	62.4	62.1	~61	~61	~61
GlcNAc(β1-3) II″	C-1	–	–	102.7	–	–	102.9	102.7
	C-2	–	–	56.4	–	–	~56	~56
	C-3	–	–	72.6	–	–	~72.6	~72.6
	C-4	–	–	**81.1**	–	–	**80.5**	**80.5**
	C-5	–	–	76.1	–	–	75.9	75.9
	C-6	–	–	61.6	–	–	~61	~61
GalNAc(β1-4) III′	C-1	–	103.3	103.1	–	102.5	102.7	102.7
	C-2	–	54.2	54.0	–	54.0	52.4	52.4
	C-3	–	73.7	72.5	–	**80.3**	**80.6**	**80.6**
	C-4	–	69.2	68.9	–	68.9	~69.3	~69.3
	C-5	–	76.7	76.8	–	75.8	75.9	75.9
	C-6	–	62.4	62.4	–	~62	~62	~62
GlcNAc(β1-3) IV′	C-1	–	–	–	102.5*	102.7	102.9	102.9
	C-2	–	–	–	56.7*	56.2	~56	~56
	C-3	–	–	–	75.3*	73.0	~72.6	~72.6
Or	C-4	–	–	–	71.1*	**80.6**	**80.5**	**80.5**
*GlcNAc(β1-6) IV	C-5	–	–	–	77.1*	76.0	75.9	75.9
	C-6	–	–	–	62.1*	61.5	~61	~61
HSO_3(4)GalNAc(β1-4) V′	C-1	–	–	–	–	102.7	103.8	103.8
	C-2	–	–	–	–	54.2	54.2	54.2
	C-3	–	–	–	–	73.2	73.2	73.2
	C-4	–	–	–	–	**77.0**	**77.1**	**77.1**
	C-5	–	–	–	–	75.9	75.9	75.9
	C-6	–	–	–	–	~62	~62	~62

Countinued

Table 8 *Countinued*

Residue	Reporter group	24	32	33	39	42	43	45
Fuc(α1-2) **F^{2}II**/Fuc(α1-2) **F^{2}III**	C-1	–	–	–	98.8/100.7	–	–	–
	C-2	–	–	–	69.2/69.2	–	–	–
	C-3	–	–	–	70.6/70.6	–	–	–
	C-4	–	–	–	73.2/73.2	–	–	–
	C-5	–	–	–	67.6/68.3	–	–	–
	C-6	–	–	–	16.7/16.7	–	–	–
NeuAc **N^3**/NeuAc **N^6**	C-3	38.7/*41.4*	38.6	41.3	–	38.5	41.3	38.7
	C-4	69.9/69.9	70.2	70.2	–	69.7	69.9	69.9
	C-5	52.9/52.9	53.4	53.2	–	52.9	52.9	52.9
	C-6	74.3/73.5	74.8	74.8	–	74.4	74.8	74.3
	C-7	69.3/69.3	69.7	70.2	–	69.3	69.3	69.3
	C-8	73.7/73.2	72.3	73.3	–	73.2	73.3	73.7
	C-9	64.2/64.2	64.4	64.3	–	63.9	64.18	64.2

that the divergence between these two very closely related species, which hybridize readily in nature, occurred during geographic isolation estimed at 2–7 million years ago [25].

The relative expression level of Kdn-to-Neu5Acyl-containing glycoconjugates in trout testis may correlate with the relative expression of specific enzyme activities required for synthesis of ManNAc-6-P and Kdn-9-P synthase [26,27]. However, a comparison of mouse and human NeuAc-9-P synthase homologue of *E. coli* NeuAc synthase (*neu B* gene product) has shown the ability of the human enzyme to generate phosphorylated forms of NeuAc and Kdn, whereas the mouse enzyme had no activity to synthesize Kdn-9-P from Man-6-P and PEP [28,29]. So far, no experimental data permit to correlate the diversity of sialylation observed in amphibians with one of these biosynthetic pathways.

The remarkable species-specific structural diversity, which has been observed in carbohydrate chains of amphibian-egg jellies, was verified for *Triturus alpestris*. The examination of the Scheme 1 indicates the presence of a new family of glycan chains containing poly-GalNAc(β1-4)GlcNAc sequence (poly LacdiNAc or poly LDN). The LacdiNAc determinant is now considered as a common constituent of *N*- and *O*- linked glycans and largely widespread in invertebrates and particularly in many human pathogens [30]. It has been found in unsubstituted as well as terminally 4-*O*-sulphated, α1-3-fucosylated, or α2-6-sialylated forms at the distal end of conjugated glycans, but has not been reported previously at truly internal position of polylactosamine chains. Nevertheless, the conversion of GalNAc(β1-4)GlcNAc(β1-)R into Gal(β1-4)GlcNAc(β1-3)GalNAc(β1-4)GlcNAc(β1-)R was obtained using the serum-derived β1,3-*N*-acetylglucosaminidase along with bovine β4GalT [31], whereas authentic poly-LDN structures were recently produced *in vitro* by Chinese Hamster Ovary Lec8 cells transfected with the *Caenorhabditis elegans* β1,4-*N*-acetylgalactosaminyltransferase [32]. To our knowledge, the poly-LDN *O*-glycans isolated from *T. alpestris* egg jelly mucins represent the first reported occurrence of this type of polymer naturally synthesized by animal cells.

Acknowledgments This investigation was supported in part by the CNRS (Unité Mixte de Recherche CNRS/USTL 8576; Directeur Jean-Claude MICHALSKI), by the Ministère de la Recherche et de l'Enseignement Supérieur and by the Génopole of Lille [The mass spectrometry facility used in this study was funded by the European Community (FEDER), the Région Nord-Pas de Calais (France), the CNRS and the Université des Sciences et Technologies de Lille]. We gratefully thank B. Coddeville and Y. Leroy for their skillful technical assistance.

References

1. Inoue, S., Inoue, Y.: Fish glycoproteins, in: Glycoproteins, I.I., Montreuil, J., Vliegenthart, J.F.G., Schachter, H., (eds.) Elsevier Science 143–61 (1997)
2. Wieruszeski, J.M., Michalski, J.C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., van Halbeck, H., Mutsears, J.H.G.M., Vliegenthart, J.F.G.: Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the Chinese swiftlet (*Genus collocalia*). J. Biol. Chem. **262**, 6650–57 (1987)
3. Strecker, G.: Amphibian glycoproteins, In: Glycoproteins, I.I., Montreuil, J., Vliegenthart, J.F.G., Schachter, H., (eds.) Elsevier Science 163–72 (1997)
4. Laine, R.A.: A calculation of all possible oligosaccharide isomers both branched and linear yields $1.05 \times 10(12)$ structures for a reducing hexasaccharide: the isomer barrier to development of single-method saccharide sequencing or synthesis systems. Glycobiology **4**, 759–67 (1994)
5. Varki, A.: Biological roles of oligosaccharides: all the theories are correct. Glycobiology **3**, 97–130 (1993)
6. Zanetta, J.P., Timmerman, P., Leroy, Y.: Gas-liquid chromatography of the heptafluorobutyrate derivatives of the *O*-methyl glycosides on capillary columns: a method for the quantitative determination of the monosaccharide composition of glycoproteins and glycolipids. Glycobiology **9**, 255–66 (1999)
7. Zanetta, J.P., Pons, A., Iwersen, M., Mariller, C., Leroy, Y., Timmerman, P., Schauer, R.: Diversity of sialic acids revealed using gas chromatography/mass spectrometry of heptafluorobutyrate derivatives. Glycobiology **11**, 663–76 (2001)

8. Kamerling, J.P., Vliegenthart, J.F.G.: High-resolution ^{1}H-nuclear magnetic resonancespectroscopy of oligosaccharide-alditols released from mucin-type *O*-glycoproteins. Biological Magnetic Resonance **10**, 1–287 (1992)
9. Strecker, G., Wieruszeski, J.M., Michalski, J.C., Alonso, C., Boilly, B., Montreuil, J.: Characterization of Le(x), Le(y) and A Le(y) antigen determinants in KDN-containing *O*-linked glycan chains from *Pleurodeles waltlii* jelly coat eggs. FEBS Lett. **298**, 39–43 (1992)
10. Strecker, G., Wieruszeski, J.M., Michalski, J.C., Alonso, C., Leroy, Y., Boilly, B., Montreuil, J.: Primary structure of neutral and acidic oligosaccharide-alditols derived from the jelly coat of the *Mexican axolotl*. Occurrence of oligosaccharides with fucosyl(α1-3)fucosyl(α1-4)-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid and galactosyl(α1-4)[fucosyl(α1-2)]galactosyl(β1-4)-*N*-acetylglucosamine sequences. Eur. J. Biochem. **207**, 995–1002 (1992)
11. Maes, E., Wieruszeski, J. M., Plancke, Y., Strecker, G.: Structure of three Kdn-containing oligosaccharide-alditols released from oviducal secretions of *Ambystoma tigrinum*: characterization of the carbohydrate sequence Fuc (α1-5) [Fuc (α1-4)] Kdn (α2-3(6). FEBS Lett. **358**, 205–10 (1995)
12. Fontaine, M.D., Wieruszeski, J.M., Plancke, Y., Delplace, F., Strecker, G.: Structure of six 3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (Kdn)-containing oligosaccharide-alditols released from oviduct secretions of *Ambystoma maculatum*. Characterization of the sequence fucosyl(α1-2)[fucosyl(α1-3)]fucosyl(α1-4)-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid. Eur. J. Biochem. **231**, 424–33 (1995)
13. Maes, E., Florea, D., Delplace, F., Lemoine, J., Plancke, Y., Strecker, G.: Structural analysis of the oligosaccharide-alditols released by reductive β-elimination from oviducal mucins of *Rana temporaria*. Glycoconj. J. **14**, 127–46 (1997)
14. Lloyd, K.O., Kabat, E.A., Layug, E.J., Gruezo, F.: Immunochemical studies on blood groups. XXXIV. Structures of some oligosaccharides produced by alkaline degradation of blood group A, B, and H substances. Biochemistry **5**, 1489–1501 (1966)
15. Lloyd, K.O.: The chemistry and immunochemistry of blood group A, B, H, and Lewis antigens: past, present and future. Glycoconj. J. **17**, 531–41 (2000)
16. Florea, D., Maes, E., Haddad, M., Strecker, G.: Structural analysis of the oligosaccharidealditols released from the jelly coat of *Rana dalmatina* eggs by reductive β-elimination. Biochimie. **84**, 611–24 (2002)
17. Mourad, R., Morelle, W., Neveu, A., Strecker, G.: Diversity of *O*-linked glycosylation patterns between species. Characterization of 25 carbohydrate chains from oviducal mucins of *Rana ridibunda*. Eur. J. Biochem. **268**, 1990–2003 (2001)
18. Coppin, A., Maes, E., Flahaut, C., Coddeville, B., Strecker, G.: Acquisition of species-specific *O*-linked carbohydrate chains from oviducal mucins in *Rana arvalis* A case study. Eur. J. Biochem. **266**, 370–82 (1999)
19. Guerardel, Y., Kol, O., Maes, E., Lefebvre, T., Boilly, B., Davril, M., Strecker, G.: *O*-glycan variability of egg-jelly mucins from *Xenopus laevis*: characterization of four phenotypes that differ by the terminal glycosylation of their mucins. Biochem. J. **352**, 449–63 (2000)
20. Nadano, D., Iwasaki, M., Endo, S., Kitajima, K., Inoue, S., Inoue, Y.: A naturally occurring deaminated neuraminic acid, 3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (KDN). Its unique occurrence at the non reducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trait eggs. J. Biol. Chem. **261**, 11550–57 (1986)
21. Knirel, Y., Kocharova, N.A., Shashkov, A.S., Kochetkov, N.K., Mamonotova, V.A., Soloveva, T.F.: Structure of the capsular polysaccharide of *Klebsiella ozaenae* serotype K4 containing 3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid. Carbohydr. Res. **188**, 145–55 (1989)
22. Ziak, M., Qu, B., Zuo, X., Zuber, C., Kanamori, A., Kitajima, K., Inoue, S., Inoue, Y., Roth, J.: Occurrence of poly(alpha2,8-deaminoneuraminic acid) in mammalian tissues: widespread and developmentally regulated but highly selective expression on glycoproteins. Proc. Natl. Acad. Sci. USA **93**, 2759–63 (1996)
23. Inoue, S., Lin, S.L., Chang, T., Wu, S.H., Yao, C.W., Chu, T.Y., Troy, F.A. Inoue, Y.: Identification of free deaminated sialic acid (2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid) in human red blood cells and its elevated expression in fetal cord red blood cells and ovarian cancer cells. J. Biol. Chem. **273**, 27199–204 (1998)
24. Coppin, A., Florea, D., Maes, E., Cogalniceanu, D., Strecker, G.: Comparative study carbohydrate chains released -form the oviducal mucins of the two very closely related amphibian species *Bombina bombina* and *Bombina variegata*. Biochimie. **85**, 53–64 (2003)
25. Arntzen, J. W.: Some hypotheses on postglacial migrations of the fire-bellied toad, *Bombina bombina* (Linnaeus) and yellow-bellied toad, *Bombina variegata* (Linnaeus). J. Biogeog. **2**, 339–45 (1978)
26. Hinderlich, S., Stäsche, R., Zeitler, R., Reutter, W.: A bifuctional enzyme catalyzes the first two steps in N-acetylneuraminic acid biosynthesis of rat liver. Purification and characterization of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. J. Biol. Chem. **272**, 24313–318 (1997)
27. Angata, T., Nakata, D., Matsuda, T., Kitajima, K., Troy, F.A.: Biosynthesis of KDN (2-Keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid). Identification and characterization of a KDN-9-phosphate synthetase activity from trout testis. J. Biol. Chem. **274**, 22949–956 (1999)
28. Nakata, D., Close, B.E., Colley, K.J., Matsuda, T., Kitajima, K.: Molecular cloning and expression of the mouse N-acetylneuraminic acid 9-phosphate synthase which does not have deaminoneuraminic acid (Kdn) 9-phosphate synthase activity. Biochem. Biophys. Res. Commun. **273**, 642–48 (2000)
29. Lawrence, S.M., Huddleston, K.A., Pitts, L.R., Nguyen, N., Lee, Y.C., Vann, W.F., Coleman, T.A., Betenbaugh, M.J.: Cloning and expression of the human N-acetylneuraminic acid phosphate synthase gene with 2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid biosynthetic ability. J. Biol. Chem. **275**, 17869–877 (2000)
30. Van den Berg, T.K., Honing, H., Franke, N., van Remoortere, A., Schiphorst, W.E.C.M., Liu, F.-T., Deelder, A.M., Cummings, R.D., Hokke, C.H., van Die, I.: LacdiNAc-glycans constitute a parasite pattern for galectin-3-mediated immune recognition. J. Immunol. **173**, 1902–07 (2004)
31. Salo, H., Aitio, O., Ilves, K., Bencomo, E., Toivonen, S., Penttilä, L., Niemelä, R., Salminen, H., Grabenhorst, E., Renkonen, R., Renkonen, O.: Several polylactosamine-modifying glycosyltransferases also use internal GalNAcβ1-4GlcNAc units of synthetic saccharides as acceptors. Glycobiology **12**, 217–28 (2002)
32. Kawar, Z.S., Haslam, S.M., Morris, H.R., Dell, A., Cummings, R.D.: Novel poly-GalNAcβ1-4GlcNAc (LacdiNAc) and fucosylated poly-LacdiNAc *N*-glycans from mammalian cells expressing β1,4-*N*-acetylgalactosaminyltransferase and α1,3-fucosyltransferase. J. Biol. Chem. **280**, 12810–819 (2005)