Volume 241, number 1,2, 246–250 FEB 06523 December 1988

Sulfated *N*-linked carbohydrate chains in porcine thyroglobulin

Johannis P. Kamerling, Irma Rijkse, Augustinus A.M. Maas, J. Albert van Kuik and Johannes F.G. Vliegenthart

Department of Bio-Organic Chemistry, Utrecht University, Transitorium III, PO Box 80.075, 3508 TB Utrecht, The Netherlands

Received 16 September 1988

N-linked carbohydrate chains of porcine thyroglobulin were released by the hydrazinolysis procedure. The resulting mixture of oligosaccharide-alditols was fractionated by high-voltage paper electrophoresis, the acidic fractions were further separated by high-performance liquid chromatography on Lichrosorb-NH_2, and analyzed by 500-MHz ^{1}H-NMR spectroscopy and, partially, by permethylation analysis. Of the acidic oligosaccharide-alditols, the following sulfated carbohydrate chains could be identified: NeuAcα2→6Galβ1→4GlcNAcβ1→2Manα1→3[(SO_3Na→3)Galβ1→4GlcNAcβ1→2-Manα1→6]Manβ1→4GlcNAcβ1→4[Fucα1→6]GlcNAc-ol and NeuAcα2→6Galβ1→4(SO_3Na→6)$_{0-1}$GlcNAcβ1→2-Manα1→3[NeuAcα2→6Galβ1→4(SO_3Na→6)$_{1.0}$GlcNAcβ1→2Manα1→6]Manβ1→4GlcNAcβ1→4[Fucα1→6]GlcNAc-ol. The sulfated structural elements for porcine thyroglobulin form novel details of *N*-linked carbohydrate chains. They contribute to the fine structure of these oligosaccharides and are another type of expression of microheterogeneity.

Thyroglobulin; Sulfated *N*-linked carbohydrate; (Porcine)

1. INTRODUCTION

Thyroglobulin, the major iodinated glycoprotein synthesized in the thyroid gland, is the polypeptide precursor of the thyroid hormones. Analysis of the carbohydrate chains of thyroglobulin (molecular mass ~660 kDa) from several species has demonstrated that mainly two types of chains occur, generally called unit-A type (oligomannose type) and unit-B type (*N*-acetyllactosamine type) [1–10]. For human thyroglobulin also unit-C type (mucin type) [4] and unit-D type (proteoglycan type) [11] chains have been indicated. In many cases α1⟶3Gal residues can form part of the unit-B type chains [12–14].

Concerning the *N*-linked oligosaccharide chains of porcine thyroglobulin [5,6], the literature data indicate the presence of $Man_{5-9}GlcNAc_2$ structures, whereby the trimming of Man residues seems to take place randomly [7]. For the *N*-acetyllactosamine type a series of partially sialylated di- (75%) and tri- (25%) antennary structures with an α1⟶6-linked Fuc residue at the Asn-bound GlcNAc unit have been reported [8]. In addition, a terminal Galα1⟶3Galβ1⟶4 element in the Manα1⟶6 branch of a monosialylated (α2⟶6-linked NeuAc; Manα1⟶3 branch) diantennary structure has been shown to occur [14].

The recent finding of phosphate [10,15–17] and, especially, sulfate [18] groups in thyroglobulins from different biological origins prompted us to report data on the structural identification of sulfated *N*-linked unit-B type carbohydrate chains in porcine thyroglobulin.

2. MATERIALS AND METHODS

2.1. *Porcine thyroglobulin*

Porcine thyroglobulin was obtained from Sigma. SDS-polyacrylamide gel electrophoresis gave the usual pattern of two main bands, together with three faint bands of lower molecular mass [19,20]. Its sugar composition (Fuc/Man/Gal/GlcNAc/

Correspondence address: J.P. Kamerling, Department of Bio-Organic Chemistry, Utrecht University, Transitorium III, PO Box 80.075, 3508 TB Utrecht, The Netherlands

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

NeuAc = 0.4:3.0:1.2:2.6:0.4) and carbohydrate content (7%) are in accordance with literature data [1,2,21].

2.2. *Preparation of oligosaccharide-alditols*

The hydrazinolysis procedure on porcine thyroglobulin (six 125-mg portions) followed by high-voltage paper electrophoresis was carried out essentially as described [22,23]. The oligosaccharide-alditols were recovered from the paper by elution with water, yielding one neutral and three acidic fractions. The acidic fractions were further subfractionated by HPLC on Lichrosorb-NH_2 using elution systems of acetonitrile/15–30 mM KH_2PO_4-K_2HPO_4, pH 5.2 [24], and monitored at 205 nm. Fractions were desalted on Bio-Gel P-2 (200–400 mesh) using water as eluent and refractive index detection.

2.3. *Monosaccharide analysis*

Carbohydrate samples were subjected to the methanolysis procedure and analyzed by GLC on CP Sil 5, as described [25].

2.4. *Methylation analysis*

Methylation analysis was carried out essentially as described [26]. Permethylated material was hydrolyzed with 4 M trifluoroacetic acid, the obtained mixtures of partially methylated monosaccharides were reduced with NaB^2H_4 in water, and the partially methylated alditol acetates analyzed by GLC-MS.

2.5. *500-MHz ^{1}H-NMR spectroscopy*

Oligosaccharide-alditols were repeatedly exchanged in ^2H_2O (99.96 atom% ^{2}H) with intermediate lyophilization. Resolution-enhanced 500 MHz ^{1}H-NMR spectra were recorded in ^2H_2O at 27°C on a Bruker WM-500 spectrometer (SON hf-NMR facility, Department of Biophysical Chemistry, University of Nijmegen, The Netherlands). Chemical shifts (δ) are given relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured indirectly to acetone (δ = 2.225 ppm) [27].

3. RESULTS

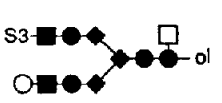
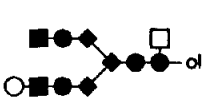
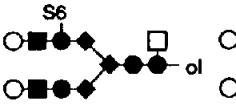
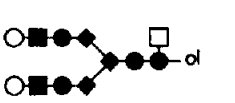
In the framework of our NMR studies on sulfated mono- [28] and oligo- [29,30] saccharides, porcine thyroglobulin was investigated for the presence of sulfated *N*-linked carbohydrate chains. The hydrazinolysis procedure in combination with high-voltage paper electrophoresis yielded one neutral (N) and three acidic (A1, A2, A3) fractions in the molar ratio of 27:35:28:10, respectively. As A2 and A3 turned out to contain the searched material, attention will only be paid to these fractions in this paper.

Subfractionation of A2, isolated from the 'double-negatively-charged' region of the paper electropherogram, on Lichrosorb-NH_2 yielded a series of peaks, of which A2b is of interest. The ^{1}H-NMR spectrum of A2b shows the presence of one major compound. The NMR features of this oligosaccharide-alditol, presented in table 1, resemble those of reference compound IgM, which is a classical diantennary structure terminated with β1⟶4-linked Gal in the Manα1⟶6 branch and α2⟶6-linked NeuAc in the Manα1⟶3 branch and having an α1⟶6-linked L-fucose residue at the Asn-bound *N*-acetylglucosamine [31]. However, of the structural-reporter-group signals the Gal-**6**′ H-1 doublet had shifted downfield from δ = 4.470 ppm to δ = 4.587 ppm. Furthermore, in contrast to the spectrum of IgM, in that of A2b two additional signals can be observed in the structural-reporter-group region, namely, a doublet of doublets at δ = 4.339 ppm (J values of 3.2 Hz and 10.0 Hz) and a doublet of doublets at δ = 4.294 ppm (J values of 3.2 Hz and ~1 Hz). The latter resonance has the typical shape of a Gal H-4 signal. By selective ^{1}H-decoupling of the signal at δ = 4.339 ppm, connected signals at δ = 4.294 ppm and δ = 3.684 ppm are found by difference spectroscopy. In the same way, selective ^{1}H-decoupling of the H-1 signal of Gal-**6**′ (δ = 4.587 ppm) and difference spectroscopy identified the resonance at δ = 3.684 ppm as Gal-**6**′ H-2. When a similar experiment is carried out for the H-1 signal of Gal-**6**, the corresponding H-2 resonance is found at δ = 3.534 ppm. In conclusion, the signals at δ = 4.339 ppm and δ = 4.294 ppm can be attributed to Gal-**6**′ H-3 and H-4, respectively.

Since the major compound in subfraction A2b is a monosialylated structure, an additional acidic substitution has to be present. In view of the literature data on acidic carbohydrate chains, it is obvious to suppose that either a phosphate or a sulfate group is involved [10,15–18]. The most obvious positions of attachment are Gal-**6**′ C-3 or C-4, as can be concluded from the downfield positions of Gal-**6**′ H-3 and H-4. Because of the absence of ^{31}P couplings on these signals, the occurrence of phosphate can be excluded. In the case of sulfate, the chemical shift values at δ = 4.339 ppm and δ = 4.294 ppm were compared to those of sulfated monosaccharide references. In table 2, the ^{1}H-NMR data of the methyl glycosides of α-D-galactopyranose, β-D-galactopyranose, 3-*O*-SO_3Na-α-D-galactopyranose and 4-*O*-SO_3Na-α-D-galactopyranose [28] are summarized. Going from methyl α-D-galactopyranoside to methyl

Table 1

^{1}H chemical shifts of structural-reporter-group protons of the constituent monosaccharides of fractions A2b and A3c obtained from porcine thyroglobulin, together with those of reference compounds IgM and PT

Reporter group	Residue[a]	Chemical shifts[b] in[c]			
		A2b	IgM	A3c[d]	PT
					
H1	GlcNAc-2	4.713	4.714	n.d.	4.715
	Man-3	4.778	4.760	n.d.	4.786
	Man-4	5.136	5.136	5.139	5.135
	Man-4′	4.924	4.924	4.942	4.942
	GlcNAc-5	4.607	4.605	4.602	4.605
	GlcNAc-5′	4.580[e]	4.581	4.643	4.605
	Gal-6	4.444	4.445	4.445	4.443
	Gal-6′	4.587[e]	4.470	4.486	4.443
	Fuc	4.898	4.896	4.896	4.896
H-2	GlcNAc-1-ol	4.213	4.220	4.212	4.219
	Man-3	4.255	4.258	4.258	4.260
	Man-4	4.198	4.197	4.205	4.198
	Man-4′	4.108	4.110	4.112	4.114
H-3a	NeuAc	1.718	1.720	1.719[f]	1.718[f]
H-3e	NeuAc	2.669	2.668	2.671[f]	2.669/2.673
H-3	Gal-6′	4.339	n.d.	n.d.	n.d.
H-4	Gal-6′	4.294	n.d.	n.d.	n.d.
H-5	Fuc	4.072	4.071	4.073	4.072
H-6	GlcNAc-5′	n.d.	n.d.	4.434	n.d.
H-6′	GlcNAc-5′	n.d.	n.d.	4.314	n.d.
CH_3	Fuc	1.223	1.224	1.225	1.224
NAc	GlcNAc-1-ol	2.056	2.057	2.056	2.056
	GlcNAc-2	2.089	2.088	2.089	2.090
	GlcNAc-5	2.070	2.071	2.072	2.071
	GlcNAc-5′	2.048	2.048	2.066	2.065
	NeuAc	2.030	2.031	2.030[f]	2.030[f]

[a] For numbering of the monosaccharide residues, see text

[b] Chemical shifts are given for neutral solutions at 27°C, in ppm downfield from internal 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O, acquired at 500 MHz (but were actually measured relative to internal acetone: δ = 2.225 ppm)

[c] Structures are represented by short-hand notation [27]: ●, GlcNAc; ◆, Man; ■, Gal; ○, NeuAcα2⟶6; □, Fuc; S, sulfate

[d] For reasons of simplicity, the δ values for GlcNAc-5/5′ H-1 and Gal-6/6′ H-1 have been grouped for a sulfate group at GlcNAc-5′

[e] Assignments may have to be interchanged

[f] Signal stems from two NeuAc residues

3-*O*-SO_3Na-α-D-galactopyranoside, similar downfield shifts for H-3 and H-4 ($\Delta\delta$ = +0.666 ppm and $\Delta\delta$ = +0.356 ppm, respectively) were observed, as going from methyl β-D-galactopyranoside to Gal-6′ in subfraction A2b (H-3, $\Delta\delta$ = +0.695 ppm; H-4, $\Delta\delta$ = +0.372 ppm). In the case of the 4-*O*-sulfated analog quite a different NMR peak pattern was observed. It has to be noted that similar positions for Gal H-3 and H-4 have recently been reported for a terminal 3-sulfated Gal-β1⟶4 residue in an *O*-linked chain [32]. Summarizing, the terminal Gal residue in the Manα1⟶6 antenna (Gal-6′) in A2b is sulfated at position C-3, leading to the following structure:

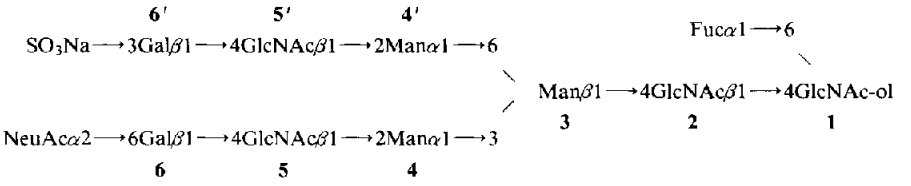


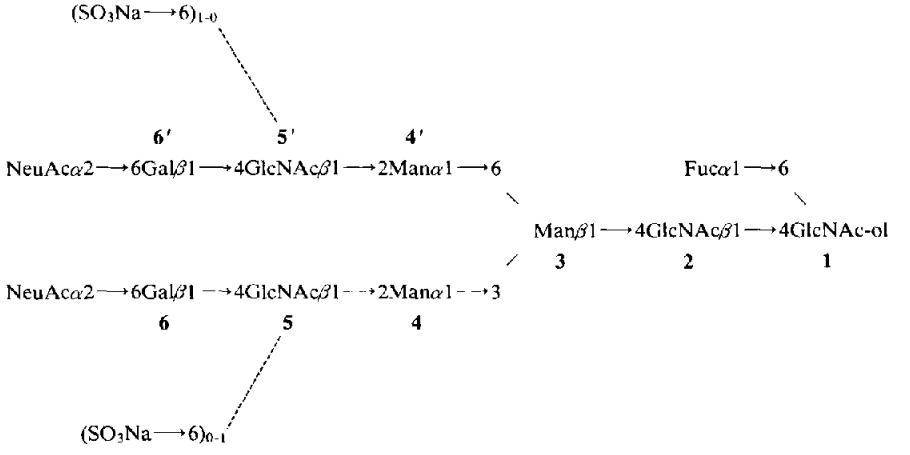
Table 2

^{1}H-NMR data for the methyl glycosides of α-D-galactopyranose, β-D-galactopyranose, 3-*O*-SO_3Na-α-D-galactopyranose and 4-*O*-SO_3Na-α-D-galactopyranose [28]

Protons	Chemical shifts[a] in			
	α-D-Gal	3-*O*-SO_3Na-α-D-Gal	4-*O*-SO_3Na-α-D-Gal	β-D-Gal
H-1	4.837	4.899	4.867	4.315
H-2	3.819	3.982	3.855	3.501
H-3	3.811	4.477	3.941	3.644
H-4	3.968	4.324	4.713	3.922
H-5	3.897	3.935	4.030	3.695
H-6	3.741	3.751	3.810	3.751
H-6′	3.752	3.759	3.755	3.796
OMe	3.414	3.430	3.425	3.573

[a] Chemical shifts are given at 27°C, in ppm downfield from internal 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O, acquired at 500 MHz (but were actually measured relative to internal acetone: δ = 2.225 ppm)

Subfractionation of fraction A3, isolated from the 'triple-negatively-charged' region of the paper electropherogram, on Lichrosorb-NH_2 resulted in several peaks, of which only fraction A3c is of interest. The ^{1}H-NMR spectrum of subfraction A3c shows a nearly pure compound. Comparison of the various resonances in the anomeric region of the spectrum as well as in the region δ < 2.8 ppm, with the spectral features of a classical α2→6 disialylated diantennary carbohydrate chain with an α1→6-linked Fuc residue at the Asn-bound GlcNAc unit, denoted PT, shows that the basic structure of A3c is similar to PT. The intensities of the GlcNAc-**5**/**5′** and Gal-**6**/**6′** H-1 doublets in A3c have only half the intensities of those of the corresponding signals in PT. But additionally, for GlcNAc as well as for Gal, a more downfield doublet is observed, namely, at δ = 4.643 ppm and at δ = 4.486 ppm, respectively. This means that in one of the antennae a structural element has changed, as compared to PT. Furthermore, two additional non-anomeric downfield signals, resonating at δ = 4.314 ppm (*J* values of 4.8 Hz and 11.0 Hz) and at δ = 4.434 ppm (*J* values of 1.2 Hz and 11.0 Hz) are observed. By selective irradiation it was found that both resonance patterns are coupled with each other. In view of earlier reports [29,33], such a typical coupled pattern belongs to H-6 and H-6′ of one of the monosaccharide residues, bearing a sulfate group at C-6. Methylation analysis of fraction A3c yielded a 4,6-disubstituted GlcNAc residue, indicating that GlcNAc-**5** or GlcNAc-**5′** bears a sulfate group at C-6. Combining the various data, the following two structures for A3c are possible:



Also the occurrence of a mixture of both structures cannot be excluded so far. The recently reported δ values of GlcNAc H-6 and H-6′ in the Galβ1 ⟶ 4(SO_3Na ⟶ 6)GlcNAc element, present in an *O*-linked chain, are in the same range as the values included here [34]. Recent studies on carbohydrate chains liberated from various ^{35}S-labeled mammalian cell lines suggested the occurrence of a NeuAcα2⟶6/3Galβ1⟶4(SO_3Na⟶6)GlcNAc element in *N*-linked oligosaccharides [35].

Acknowledgements: This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO), and by the Netherlands Foundation for Cancer Research (KWF, grant UUKC 83-13).

REFERENCES

[1] McQuillan, M.T. and Trikojus, V.M. (1972) in: The Glycoproteins, vol.B (Gottschalk, A. ed.) pp.926–963, Elsevier, Amsterdam.
[2] Spiro, R.G. and Spiro, M.J. (1965) J. Biol. Chem. 240, 997–1001.
[3] Spiro, R.G. (1965) J. Biol. Chem. 240, 1603–1610.
[4] Arima, T., Spiro, M.J. and Spiro, R.G. (1972) J. Biol. Chem. 247, 1825–1835.
[5] Fukuda, M. and Egami, F. (1971) Biochem. J. 123, 407–414.
[6] Fukuda, M. and Egami, F. (1971) Biochem. J. 123, 415–420.
[7] Tsuji, T., Yamamoto, K., Irimura, T. and Osawa, T. (1981) Biochem. J. 195, 691–699.
[8] Yamamoto, K., Tsuji, T., Irimura, T. and Osawa, T. (1981) Biochem. J. 195, 701–713.
[9] Ito, S., Yamashita, K., Spiro, R.G. and Kobata, A. (1977) J. Biochem. 81, 1621–1631.
[10] Yamamoto, K., Tsuji, T., Tarutani, O. and Osawa, T. (1984) Eur. J. Biochem. 143, 133–144.
[11] Spiro, M.J. (1977) J. Biol. Chem. 252, 5424–5430.
[12] Spiro, R.G. and Bhoyroo, V.D. (1984) J. Biol. Chem. 259, 9858–9866.
[13] Dorland, L., Van Halbeek, H. and Vliegenthart, J.F.G. (1984) Biochem. Biophys. Res. Commun. 122, 859–866.
[14] Maas, A.A.M., Rijkse, I., Van Halbeek, H., Kamerling, J.P. and Vliegenthart, J.F.G. (1985) Abstr. Third Eur. Symp. Carbohydr. (Defaye, J. ed.) p.20, Central University Library, Grenoble.
[15] Yamamoto, K., Tsuji, T., Tarutani, O. and Osawa, T. (1985) Biochim. Biophys. Acta 838, 84–91.
[16] Consiglio, E., Acquaviva, A.M., Formisano, S., Liguoro, D., Gallo, A., Vittorio, T., Santisteban, P., De Luca, M., Shifrin, S., Yeh, H.J.C. and Kohn, L.D. (1987) J. Biol. Chem. 262, 10304–10314.
[17] Herzog, V., Neumuller, W. and Holzmann, B. (1987) EMBO J. 6, 555–560.
[18] Spiro, M.J. and Spiro, R.G. (1988) Endocrinology 123, 56–65.
[19] Rolland, M. and Lissitzky, S. (1972) Biochim. Biophys. Acta 278, 316–336.
[20] Spiro, M.J. (1973) J. Biol. Chem. 248, 4446–4460.
[21] Ronin, C., Fenouillet, E., Hovsepian, S., Fayet, G. and Fournet, B. (1986) J. Biol. Chem. 261, 7287–7293.
[22] Takasaki, S., Mizuochi, T. and Kobata, A. (1982) Methods Enzymol. 83, 263–268.
[23] Van Kuik, J.A., Van Halbeek, H., Kamerling, J.P. and Vliegenthart, J.F.G. (1985) J. Biol. Chem. 260, 13984–13988.
[24] Joziasse, D.H., Blanken, W.M., Koppen, P.L. and Van den Eijnden, D.H. (1983) Carbohydr. Res. 119, 303–309.
[25] Kamerling, J.P. and Vliegenthart, J.F.G. (1982) Cell Biol. Monogr. 10, 95–125.
[26] Waeghe, T.J., Darvill, A.G., McNeill, M. and Albersheim, P. (1983) Carbohydr. Res. 123, 281–304.
[27] Vliegenthart, J.F.G., Dorland, L. and Van Halbeek, H. (1983) Adv. Carbohydr. Chem. Biochem. 41, 209–374.
[28] Ruiz Contreras, R., Kamerling, J.P., Breg, J. and Vliegenthart, J.F.G. (1988) Carbohydr. Res. 179, 411–418.
[29] Van Kuik, J.A., Breg, J., Kolsteeg, C.E.M., Kamerling, J.P. and Vliegenthart, J.F.G. (1987) FEBS Lett. 221, 150–154.
[30] Sugahara, K., Yamashina, I., De Waard, P., Van Halbeek, H. and Vliegenthart, J.F.G. (1988) J. Biol. Chem. 263, 10168–10174.
[31] Cahour, A., Debeire, P., Hartmann, L., Montreuil, J., Van Halbeek, H. and Vliegenthart, J.F.G. (1984) FEBS Lett. 170, 343–349.
[32] Capon, C., Cache, P., Leroy, Y., Wieruszeski, J.-M., Strecker, G., Montreuil, J. and Fournet, B. (1987) Proc. IXth Int. Symp. Glycoconjugates (Montreuil, J. et al. eds) p.A97, Lerouge Tourcoing.
[33] Hounsell, E.F., Feeney, J., Scudder, P., Tang, P.W. and Feizi, T. (1986) Eur. J. Biochem. 157, 375–384.
[34] Strecker, G., Wieruszeski, J.-M., Martel, C. and Montreuil, J. (1988) Glycoconj. J. 4, 329–337.
[35] Roux, L., Holojda, S., Sundblad, G., Freeze, H.H. and Varki, A. (1988) J. Biol. Chem. 263, 8879–8889.