

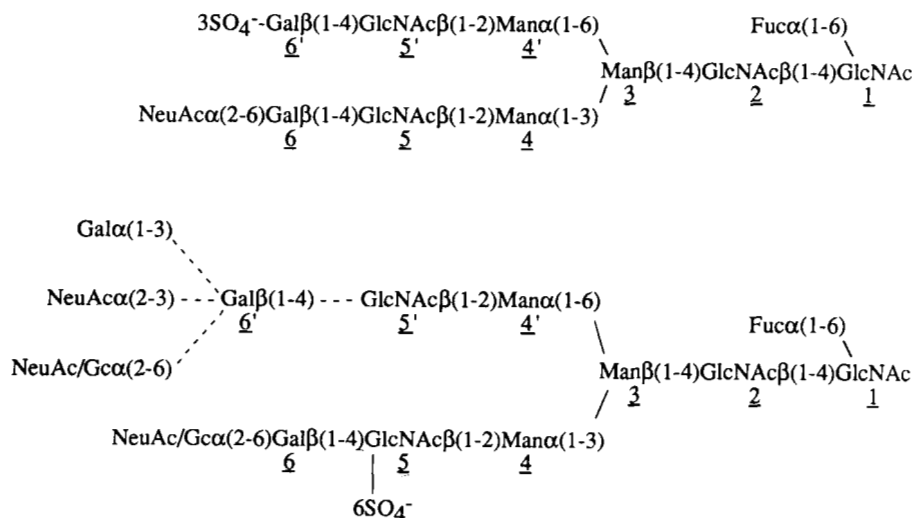
Structure Determination by ^1H NMR Spectroscopy of (Sulfated) Sialylated *N*-Linked Carbohydrate Chains Released from Porcine Thyroglobulin by Peptide- N^4 -(*N*-acetyl- β -glucosaminyl)asparagine Amidase-F*

(Received for publication, April 26, 1990)

Pieter de Waard, Anita Koorevaar, Johannes P. Kamerling, and Johannes F. G. Vliegthart

From the Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P. O. Box 80.075, NL-3508 TB Utrecht, The Netherlands

The *N*-linked carbohydrate chains of porcine thyroglobulin were released by peptide- N^4 -(*N*-acetyl- β -glucosaminyl)asparagine amidase-F (PNGase-F). The resulting oligosaccharides were fractionated by a combination of fast protein liquid chromatography and high performance liquid chromatography and analyzed by 500-MHz ^1H NMR spectroscopy. The major acidic compounds are mono- and disialylated, fucosylated diantennary compounds terminated with $\alpha(2\text{--}6)$ -linked sialic acid on the $\text{Man}\alpha(1\text{--}3)$ branch. The $\text{Man}\alpha(1\text{--}6)$ branch shows a large heterogeneity. It can be terminated with Man-4' , GlcNAc-5' , or Gal-6' , whereas the Gal-6' residue may be extended with $\text{Gal}\alpha(1\text{--}3)$, $\text{NeuAc}\alpha(2\text{--}3)$, or $\text{Sia}\alpha(2\text{--}6)$. In the major structures 8% of $\alpha(2\text{--}6)$ -linked sialic acid was found as NeuGc instead of NeuAc. The main compounds have sulfated homologues bearing a sulfate group (6–20%) at C-3 of Gal-6' or at C-6 of GlcNAc-5 as follows.



Thyroglobulin is the major iodinated glycoprotein in the thyroid gland. In porcine thyroglobulin only *N*-linked oligosaccharides are found, and two types of chains can be distinguished, generally denoted unit A (oligomannose type) and unit B (*N*-acetylglucosamine type) (1–3). For the oligomannose type a series of $\text{Man}_n\text{GlcNAc}_2$ structures have been identified (4), and for the *N*-acetylglucosamine type a series of $\alpha(1\text{--}6)$ -fucosylated di- and triantennary structures (5). The latter type of structures is partially sialylated, and the pres-

ence of NeuGc, besides NeuAc, has been suggested (1). Furthermore, a terminal $\text{Gal}\alpha(1\text{--}3)\text{Gal}\beta(1\text{--}4)$ element occurs in the $\text{Man}\alpha(1\text{--}6)$ branch (6, 7). In human and calf thyroglobulin 3-*O*-sulfated Gal and 6-*O*-sulfated GlcNAc residues have been demonstrated (8, 9).

The presence of non-carbohydrate substituents can give rise to additional microheterogeneity of carbohydrate structures. ^1H NMR spectroscopy is suited to determine the presence, type and location of substituents, like acetyl (10, 11), methyl (12, 13), phosphate (14, 15), and sulfate groups (16, 17). Sulfate groups give rise to a downfield shift of approximately 0.5 ppm for the proton attached to the substituted carbon atom (18). One- and two-dimensional HOHAHA¹

* This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and the Netherlands Foundation for Cancer Research (KWF, Grant UUKC 88-14). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ The abbreviations used are: HOHAHA, homonuclear Hartmann Hahn; PNGase-F, peptide- N^4 -(*N*-acetyl- β -glucosaminyl)asparagine amidase-F; MLEV, Malcom Levitt; ROESY, rotating frame Overhauser enhancement spectroscopy.

HPLC of FPLC fraction **S₃** gives rise to three fractions, denoted **S₃-1** to **S₃-3** (Fig. 2E). The ¹H NMR data are compiled in Table II, the structures are given in Scheme 4, and the ¹H NMR spectrum of fraction **S₃-2** is shown in Fig. 3D. The spectra of fractions **S₃-1**, **S₃-2**, and **S₃-3** are similar to those of fractions **2-1**, **2-2**, and **2-3**, respectively, except for additional signals, that have shifted out of the bulk-region, and downfield shifts for one of the Gal and one of the outer GlcNAc H-1 signals. Two-dimensional HOHAHA spectroscopy of **S₃-2** proves the interconnection of the additional signals at δ = 4.43 ppm and 4.31 ppm with the shifted GlcNAc H-1 signal (Fig. 4B), and they could be assigned as GlcNAc H-6 and H-6', respectively. The chemical shift values for this GlcNAc residue are in accordance with those of 6-*O*-sulfated GlcNAc residues reported in studies on keratan sulfate (36) and *O*-linked chains (37, 38). In the corresponding oligosaccharide-alditol of **S₃-2** it was not possible to assign the branch bearing the 6-*O*-sulfate group (19). For the assignment of the branch location in **S₃-2** a one-dimensional rotating frame Overhauser enhancement spectroscopy experiment was performed, with a selective pulse on the Man-4 H-1 signal. The spinlock pulse will allow magnetization transfer through space to neighboring protons. In this way it is demonstrated that the anomeric proton of the sulfated GlcNAc residue is close in space to Man-4 H-1 (data not shown). Therefore, the disialylated sulfated compounds of fraction **S₃** bear a 6-*O*-sulfated GlcNAc-5 residue on the Manα(1-3) branch, similar

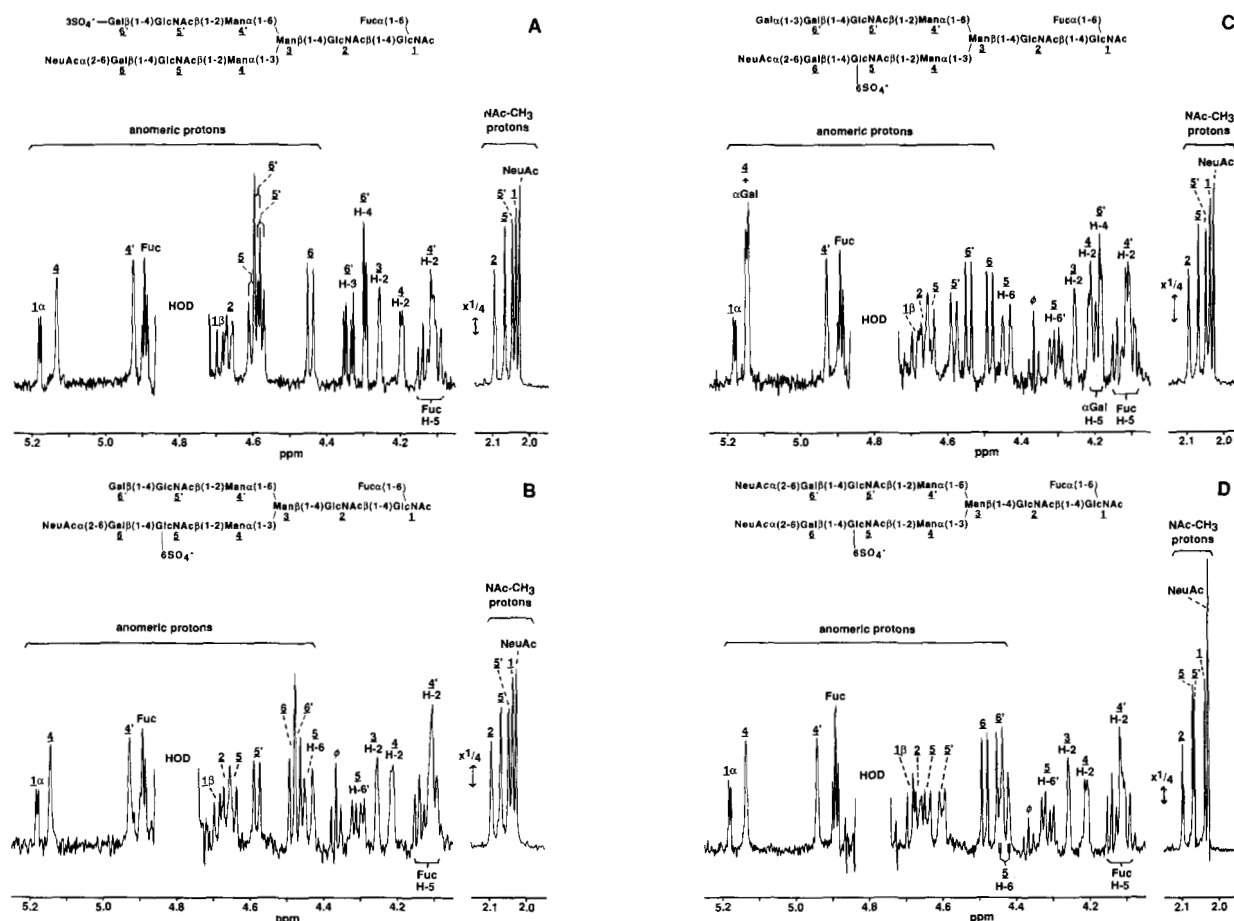


FIG. 3. Structural-reporter-group regions of the resolution-enhanced 500-MHz ^1H NMR spectra of sulfated oligosaccharides obtained from porcine thyroglobulin. The spectra were measured at 24 $^\circ\text{C}$. The sialic acid H-3 regions are not shown. A, fraction S_1 -1; B, fraction S_2 -2; C, fraction S_2 -3; D, fraction S_3 -2.

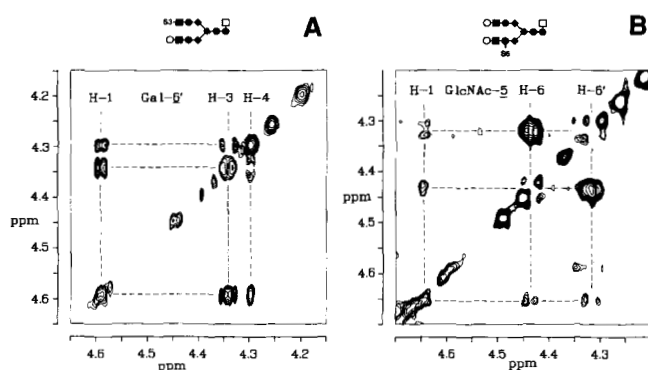


FIG. 4. Two-dimensional HOHAHA spectra of the major sulfated oligosaccharides derived from porcine thyroglobulin. The mixing time is 120 ms. Lines are drawn to show the interconnection between protons of the sulfated residues. A, fraction S_1 -1; B, fraction S_3 -2.

to the monosialylated sulfated compounds of fraction S_2 . These data indicate that the compounds in fractions S_3 -1, S_3 -2, and S_3 -3 are 6-*O*-sulfated analogues of those in fractions 2-1, 2-2, and 2-3, respectively.

DISCUSSION

In recent years it has been demonstrated that sulfate can occur as anionic substituent of Man, Gal, GlcNAc, or GalNAc residues in glycoprotein glycans (9, 16, 19, 34, 37, 38, 40-43).

fraction number	molar percentage of acidic material	structure
S_3 -1	0.5%	$\begin{array}{c} \text{NeuAc}\alpha(2-3)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-6) \\ \text{NeuAc}\alpha(2-6)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-3) \\ \text{Man}\beta(1-4)\text{GlcNAc}\beta(1-4)\text{GlcNAc} \\ \text{Fuc}\alpha(1-6) \\ \text{6SO}_4^- \end{array}$
S_3 -2	2.5%	$\begin{array}{c} \text{NeuAc}\alpha(2-6)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-6) \\ \text{NeuAc}\alpha(2-6)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-3) \\ \text{Man}\beta(1-4)\text{GlcNAc}\beta(1-4)\text{GlcNAc} \\ \text{Fuc}\alpha(1-6) \\ \text{6SO}_4^- \end{array}$
S_3 -3	0.4%	$\begin{array}{c} \text{NeuAc/Glc}\alpha(2-6)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-6) \\ \text{NeuGc/Ac}\alpha(2-6)\text{Gal}\beta(1-4)\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-3) \\ \text{Man}\beta(1-4)\text{GlcNAc}\beta(1-4)\text{GlcNAc} \\ \text{Fuc}\alpha(1-6) \\ \text{6SO}_4^- \end{array}$

SCHEME 4. Disialylated monosulfated structures obtained from porcine thyroglobulin.

For *N*-linked carbohydrate chains primary structures have been proposed in case of *Panulirus interruptus* hemocyanin (6-*O*-sulfated Man-4') (16), hen egg albumin (4-*O*-sulfated Man-*B* and/or Man-*A*) (42), and lutropin (4-*O*-sulfated GalNAc in GalNAc β (1-4)GlcNAc) (43). In the course of our

structural studies on porcine thyroglobulin, data became available, demonstrating that the *N*-linked oligosaccharides of calf and human thyroglobulin contain sulfated *N*-acetylglucosamine elements with either a sulfate group at C-3 of Gal or at C-6 of GlcNAc (9). However, no conclusive information was presented about the branch location of these disaccharide elements. Based on concanavalin A binding studies of human and calf thyroglobulin-derived glycopeptides, it was proposed that 3-*O*-sulfated *N*-acetylglucosamine occurs in diantennary as well as in higher branched glycan chains, whereas 6-*O*-sulfated *N*-acetylglucosamine is located in glycan chains with three or more branches (9). Human thyroglobulin contains a much higher percentage of 3-*O*-sulfated Gal than calf thyroglobulin, which is explained by the complete absence of the terminal Gal α 1-3Gal β 1-sequence in the human glycoprotein (9). Using ^1H NMR spectroscopy it has now been possible to identify the exact location of the sulfated *N*-acetylglucosamine structural elements in *N*-linked oligosaccharide chains of porcine thyroglobulin (Ref. 19 and this study).

The major acidic *N*-linked carbohydrate chains in porcine thyroglobulin are monosialylated or disialylated (Schemes 1 and 3). The sulfation of the monosialylated compounds can be summarized as follows. Comparison of the molar percentages of fractions 1-2 and 1-5 (Scheme 1) with those of fractions S₂-1 and S₂-3 (Scheme 2), respectively, reveals that both parent compounds are sulfated at C-6 of GlcNAc-5 for about 10%. The compound in fraction 1-3 is found as 6-*O*-sulfated (GlcNAc-5) homologue for 6% (fraction S₂-2) and as 3-*O*-sulfated (Gal-6') homologue for 12% (fraction S₁-1). For the possible presence of 3-*O*- or 6-*O*-sulfated homologues of the minor components, some evidence could be obtained from the NMR spectra of several minor fractions. For the disialylated compounds, comparison of the molar percentages of fractions 2-1, 2-2, and 2-3 (Scheme 3) with those of fractions S₃-1, S₃-2, S₃-3 (Scheme 4), respectively, shows that the parent compounds contain up to 20% sulfate at C-6 of GlcNAc-5. Terminal α Gal residues have been detected for 24% of the acidic carbohydrate chains, which is in accordance with earlier reports (6, 44). Furthermore, sialic acid occurs for 8% in the *N*-glycolyl form, which is in agreement with the earlier reported value of 10% (45), but contrasts the high value of 40% (1). For the major NeuAc-containing carbohydrate structures NeuGc-containing analogues have been identified. It is noteworthy that α 2-3-linked NeuGc has not been found in the various carbohydrate chains.

From a biosynthetic point of view the characterized series of carbohydrate chains are highly interesting. The identified oligosaccharides fit the known branch specificities of α 2-6-sialyltransferase (32) and α 1-3-galactosyltransferase (46). When both α 2-6- and α 2-3-linked sialic acid are present, the latter type of linkage is generally found in the Man α 1-6 branch of diantennary compounds. The existence of 6-*O*-sulfated GlcNAc in the Man α 1-3 branch of the diantennary structures only, suggests also a branch specificity for the yet not established *N*-acetylglucosamine-6-*O*-sulfotransferase. Although sulfate addition seems to be a late event in the biosynthesis of the *N*-linked carbohydrate chains, it is still too early to speculate about possible biosynthetic sequences of attachment. Because of the recent finding of a nonsialylated 6-*O*-sulfated *N*-acetylglucosamine element in *O*-linked carbohydrate chains (37, 38), it is tempting to suppose that 6-*O*-sulfation of GlcNAc-5 can take place before sialic acid is attached at C-6 of Gal-6. Calf thyroid microsomes have been shown to contain a 3-*O*-sulfotransferase which catalyzes the transfer of sulfate from 3'-phosphoadenosine 5'-phospho[^{35}S] sulfate to C-3 of Gal in the *N*-acetylglucosamine element (47).

Because in this study 3-*O*-sulfated Gal has only been found in the Man α 1-6 branch of diantennary structures, also for the galactose-3-*O*-sulfotransferase a branch specificity is indicated. Therefore, with respect to the capping of Gal-6' in porcine thyroglobulin several biosynthetic routes are followed, in which α 2-6-sialyltransferase, α 2-3-sialyltransferase, α 1-3-galactosyltransferase, and galactose-3-*O*-sulfotransferase compete.

The biological role of the sulfated variants of conventional *N*-linked carbohydrate chains remains to be clarified. With reference to the present knowledge about the function of Man 6-phosphate (48), it is possible that these variants have a function as recognition signals. With the detailed knowledge of these minor structures in mind, further studies with respect to biosynthetic pathways and biological significance are feasible.

REFERENCES

1. Fukuda, M., and Egami, F. (1971) *Biochem. J.* **123**, 407-414
2. Fukuda, M., and Egami, F. (1971) *Biochem. J.* **123**, 415-420
3. Malthiery, Y., Marriq, C., Bergé-Lefranc, J.-L., Franc, J.-L., Henry, M., Lejeune, P.-J., Ruf, J., and Lissitzky, S. (1989) *Biochimie* **71**, 195-210
4. Tsuji, T., Yamamoto, K., Irimura, T., and Osawa, T. (1981) *Biochem. J.* **195**, 691-699
5. Yamamoto, K., Tsuji, T., Irimura, T., and Osawa, T. (1981) *Biochem. J.* **195**, 701-713
6. Spiro, R. G., and Bhoyroo, V. D. (1984) *J. Biol. Chem.* **259**, 9858-9866
7. Maas, A. A. M., Rijkse, I., Van Halbeek, H., Kamerling, J. P., and Vliegthart, J. F. G. (1985) in *Abstracts of the Third European Symposium on Carbohydrates* (Defaye, J., ed) p. 20, Central University Library, Grenoble, France
8. Spiro, M. J., and Spiro, R. G. (1988) *Endocrinology* **123**, 56-65
9. Spiro, R. G., and Bhoyroo, V. D. (1988) *J. Biol. Chem.* **263**, 14351-14358
10. Damm, J. B. L., Voshol, H., Hård, K., Kamerling, J. P., and Vliegthart, J. F. G. (1989) *Eur. J. Biochem.* **180**, 101-110
11. Damm, J. B. L., Hård, K., Kamerling, J. P., Van Dedem, G. W. K., and Vliegthart, J. F. G. (1990) *Eur. J. Biochem.* **189**, 175-183
12. Van Kuik, J. A., Sijbesma, R. P., Kamerling, J. P., Vliegthart, J. F. G., and Wood, E. J. (1987) *Eur. J. Biochem.* **169**, 399-411
13. Gerwig, G. J., de Waard, P., Kamerling, J. P., Vliegthart, J. F. G., Morgenstern, E., Lamed, R., and Bayer, E. A. (1989) *J. Biol. Chem.* **264**, 1027-1035
14. de Waard, P., Vliegthart, J. F. G., Kozutsumi, Y., Kawasaki, T., and Yamashina, I. (1989) *J. Biol. Chem.* **264**, 12141-12144
15. Hernandez, L. M., Ballou, L., Alvarado, E., Tsai, P.-K., and Ballou, C. E. (1989) *J. Biol. Chem.* **264**, 13648-13659
16. Van Kuik, J. A., Breg, J., Kolsteeg, C. E. M., Kamerling, J. P., and Vliegthart, J. F. G. (1987) *FEBS Lett.* **221**, 150-154
17. Sugahara, K., Yamashina, I., de Waard, P., Van Halbeek, H., and Vliegthart, J. F. G. (1988) *J. Biol. Chem.* **263**, 10168-10174
18. Harris, M. J., and Turvey, J. R. (1970) *Carbohydr. Res.* **15**, 57-63
19. Kamerling, J. P., Rijkse, I., Maas, A. A. M., Van Kuik, J. A., and Vliegthart, J. F. G. (1988) *FEBS Lett.* **241**, 246-250
20. Freeze, H. H., and Varki, A. (1986) *Biochem. Biophys. Res. Commun.* **140**, 967-973
21. Bräulke, T., Hille, A., Hüttner, W. B., Hasilik, A., and Von Figura, K. (1987) *Biochem. Biophys. Res. Commun.* **143**, 178-185
22. Maley, F., Trimble, R. B., Tarentino, A. L., and Plummer Jr, T. H. (1989) *Anal. Biochem.* **180**, 195-204
23. Damm, J. B. L., Kamerling, J. P., Van Dedem, G. W. K., and Vliegthart, J. F. G. (1987) *Glycoconjugate J.* **4**, 129-144
24. Ernst, R. R. (1966) *Adv. Magn. Reson.* **2**, 1-135
25. Vliegthart, J. F. G., Dorland, L., and Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41**, 209-374
26. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65**, 355-360
27. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **63**, 207-213

28. Kessler, H., Oschkinat, H., and Griesinger, C. (1986) *J. Magn. Reson.* **70**, 106–133
29. Damm, J. B. L., Voshol, H., Hård, K., Kamerling, J. P., Van Dedem, G. W. K., and Vliegthart, J. F. G. (1988) *Glycoconjugate J.* **5**, 221–233
30. Damm, J. B. L., Bergwerff, A. A., Hård, K., Kamerling, J. P., and Vliegthart, J. F. G. (1989) *Recl. Trav. Chim. Pays-Bas.* **108**, 351–359
31. Dorland, L., Van Halbeek, H., and Vliegthart, J. F. G. (1984) *Biochem. Biophys. Res. Commun.* **122**, 859–866
32. Joziassse, D. H., Schiphorst, W. E. C. M., Van den Eijnden, D. H., Van Kuik, J. A., Van Halbeek, H., and Vliegthart, J. F. G. (1987) *J. Biol. Chem.* **262**, 2025–2033
33. Ruiz Contreras, R., Kamerling, J. P., Breg, J., and Vliegthart, J. F. G. (1988) *Carbohydr. Res.* **179**, 411–418
34. Capon, C., Leroy, Y., Wieruszski, J.-M., Ricart, G., Strecker, G., Montreuil, J., and Fournet, B. (1989) *Eur. J. Biochem.* **182**, 139–152
35. Nagai, K., Roberts, D. D., Toida, T., Matsumoto, H., Kushi, Y., Handa, S., and Ishizuka, I. (1989) *J. Biol. Chem.* **264**, 16229–16237
36. Hounsell, E. F., Feeney, J., Scudder, P., Tang, P. W., and Feizi, T. (1986) *Eur. J. Biochem.* **157**, 375–384
37. Strecker, G., Wieruszski, J.-M., Martel, C., and Montreuil, J. (1987) *Glycoconjugate J.* **4**, 329–337
38. Strecker, G., Wieruszski, J.-M., Martel, C., and Montreuil, J. (1989) *Carbohydr. Res.* **185**, 1–13
39. Monsigny, M., Petit, C., and Roche, A.-C. (1988) *Anal. Biochem.* **175**, 525–530
40. Roux, L., Holojda, S., Sundblad, G., Freeze, H. H., and Varki, A. (1988) *J. Biol. Chem.* **263**, 8879–8889
41. Freeze, H. H., and Wolgast, D. (1986) *J. Biol. Chem.* **261**, 127–134
42. Yamashita, K., Ueda, I., and Kobata, A. (1983) *J. Biol. Chem.* **258**, 14144–14147
43. Baenziger, J. U., and Green, E. D. (1988) *Biochim. Biophys. Acta* **947**, 287–306
44. Ronin, C., Fenouillet, E., Hovsepian, S., Fayet, G., and Fournet, B. (1986) *J. Biol. Chem.* **261**, 7287–7293
45. Spiro, R. G., and Spiro, M. J. (1965) *J. Biol. Chem.* **240**, 997–1001
46. Joziassse, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. J., and Shaper, N. L. (1989) *J. Biol. Chem.* **264**, 14290–14297
47. Kato, Y., and Spiro, R. G. (1989) *J. Biol. Chem.* **264**, 3364–3371
48. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631–664

Supplemental Material to:

Structure Determination by ^1H NMR Spectroscopy of (Sulfated) Sialylated N-Linked Carbohydrate Chains Released from Porcine Thyroglobulin by PNGase-F*

by: Pieter de Waard, Anita Koorevaar, Johannes P. Kamerling, and Johannes F. G. Vliegthart

EXPERIMENTAL PROCEDURES

Materials. Porcine thyroglobulin was obtained from Sigma, and checked for purity and sugar composition as described earlier (19). PNGase-F from *Flavobacterium meningosepticum* was obtained from Boehringer Mannheim.

Liberation of the Carbohydrate Chains. The N-linked carbohydrate chains were enzymatically released from the glycoprotein essentially according to (23). Briefly, 250 mg thyroglobulin was dissolved in 12.5 ml 50 mM Tris/HCl (pH 7.2), containing 50 mM EDTA, 2.5 % (m/v) SDS and 1 % (v/v) 2-mercaptoethanol, and incubated for 1 h at 37°C. After cooling to room temperature (Noidet P-40) was added to a final concentration of 4 % (m/v) and the sample was incubated with PNGase-F for 24 h at room temperature in an end-over-end mixer. Aliquots of 25, 12.5, and 12.5 μl of the enzyme were added at 0, 4, and 10 h, respectively. After lyophilization, the residue was dissolved in 50 mM NH_4HCO_3 , adjusted to pH 7.2 with HCl, centrifuged, and the supernatant fractionated on a Bio-Gel P-100 column (2.0 x 46 cm, 200–400 mesh, Bio-Rad) using the same buffer as eluent at a flow rate of 16 ml/h. Collected 2.6-ml fractions were stained for carbohydrate with orcinol/ H_2SO_4 , and carbohydrate-positive fractions were pooled and lyophilized. The residue was dissolved in water, and residual SDS and NP-40 were removed on an Extraction gel D column (0.5 x 1.5 cm, Pierce). Finally, the sample was desalted on a Bio-Gel P-2 column (1.4 x 28 cm, 200–400 mesh, Bio-Rad) using water as eluent.

Con A Binding Assay. The degree of N-deglycosylation was determined with a Con A binding assay (J.P.M. Langeveld, unpublished results). Glycoprotein samples, before and after digestion with PNGase-F, were equally diluted with 0.05 M NaHCO_3 / Na_2CO_3 buffer, pH 9.6, containing 0.02% (m/v) NaN_3 . Aliquots of 200 μl of each dilution, ranging from 0.05 to 2 μg of sample per ml, were pipetted into the wells of flat bottom, polystyrene microtiter plates with high adsorptive (coating) capacity (Nunc, Denmark) and incubated overnight. After washing, the wells were incubated with 200 μl Con A (40 $\mu\text{g}/\text{ml}$; Sigma, type IV) in 20 mM Tris/HCl, pH 7.5, containing 0.05 M NaCl, 1 mM CaCl_2 , 0.05% (m/v) Tween 20, 0.2% (m/v) bovine serum albumin, and 0.02% (m/v) NaN_3 for 4 h. The plates were washed, incubated with 200 μl horse radish peroxidase (10 $\mu\text{g}/\text{ml}$; Sigma, type VI) in the same buffer for 1 h, and washed again. All washings were performed with 0.15 M NaCl, containing 0.05% (m/v) Tween 20, 1 mM CaCl_2 , and 1 mM MnCl_2 . Then, the plates were developed with 200 μl substrate solution containing o-phenylenediamine dihydrochloride (0.4 mg/ml) and 0.012% (m/v) H_2O_2 in 24.3 mM citric acid / 51.4 mM Na_2HPO_4 . The reaction was stopped after 30 min by adding 50 μl 2 M H_2SO_4 , and the absorbance was measured at 492 nm, reflecting residual N-glycosylation. All steps were carried out at ambient temperature. When bovine serum albumin was used for coating as a control, essentially no binding of the lectin was observed, indicating that nonspecific binding to protein does not occur; PNGase-F does not react either.

FPLC-fractionation. Fractionation according to charge of the enzymatically released carbohydrate chains was carried out on a Mono Q HR 5/5 anion-exchange column (Pharmacia FPLC system), at a flow rate of 120 ml/h. A linear concentration gradient from 0–50 mM NaCl in 8 ml water (HPLC-quality) was used, followed by a steeper gradient from 50–500 mM NaCl in 8 ml water (23). The fractionation was monitored at 214 nm, and carbohydrate-containing fractions (orcinol/ H_2SO_4) were lyophilized, desalted on Bio-Gel P-2, and lyophilized again.

HPLC-fractionation. Subfractionation of the carbohydrate-containing Mono Q fractions was carried out on a Kratos SF 400 HPLC system (ABI Analytical, Kratos Division) equipped with a 10- μm Lichrosorb-NH $_2$ column (0.46 x 25 cm, Chrompack). Elutions were performed with 30 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, pH 7.0 / acetonitrile (35:65, v/v) at a flow rate of 1.5 ml/min and room temperature. The injection volume was 50 μl . The HPLC fractions were desalted on Bio-Gel P-2.

500-MHz ^1H NMR spectroscopy. Prior to ^1H NMR spectroscopic analysis oligosaccharide samples were repeatedly exchanged in H_2O (99.96% ^1H , Aldrich) with intermediate lyophilization. ^1H NMR spectra were recorded on a Bruker AM-500 spectrometer (Department of Chemistry, Utrecht University) operating at 500 MHz at a probe temperature of 24°C. Resolution enhancement of 1D spectra was achieved by Lorentzian-to-Gaussian transformation (24). Chemical shifts are given relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured indirectly to acetone in H_2O ($\delta = 2.225$ ppm) (25). For the 2D HOHAHA spectra (26), a MLEV-17 mixing sequence of 120 ms was used; the 90° ^1H pulse width was 27 μs . The HOD signal was suppressed by presaturation during 1 s. 512 Spectra of 2048 data points were recorded, with 64 scans per t_1 value. The total measuring time for each spectrum was 15 h. The time domain data were multiplied with a phase shifted sine-bell. After Fourier transformation the resulting dataset of 1024 x 1024 points was baseline corrected in both frequency domains by a third order polynomial fit. The 2D ROESY pulse sequence (27) was modified into a 1D ROESY sequence, wherein the first 90° pulse together with the increment t_1 were replaced by a selective 270° sine pulse (53 ms) (28). The ROESY spinlock pulse had a duration of 200 ms.

RESULTS

Porcine thyroglobulin was deglycosylated with PNGase-F. Since SDS-PAGE could not be used to follow the N-deglycosylation, due to precipitation of deglycosylated thyroglobulin, a Con A binding assay was chosen to follow the degree of N-deglycosylation during the incubation. After 7 h the digestion was almost complete. The addition of extra PNGase-F did not result in a further release of carbohydrate chains. After filtration on Bio-Gel P-100 the desalted carbohydrate-containing fraction was fractionated by FPLC (Fig. 1). Fractions having the same retention time as neutral or sialic-acid-containing reference compounds (29) are denoted 0, 1, 2, and 3, respectively. On basis of their elution positions three peaks presumably contain sulfated compounds, and they are denoted S $_1$, S $_2$, and S $_3$. The remaining peaks denoted with 6 contain trace amounts of carbohydrates, being too low for structural identification by ^1H NMR spectroscopy. The FPLC fractions were subfractionated by HPLC.

HPLC of the neutral FPLC fraction 0, comprising 33% of the total amount of carbohydrates, gives rise to 5 fractions, denoted 0-1 to 0-5. ^1H NMR spectroscopy of these fractions reveals a series of oligomannose chains (data not shown), namely, Man $_2$ GlcNAc $_2$ up to and including Man $_5$ GlcNAc $_2$. The fractions 0-3 and 0-4 are heterogeneous with respect to the terminal Man $_2$ (1-2) residues. These results are in accordance with (4).

The relevant ^1H NMR data of the acidic fractions are compiled in Tables I and II. All acidic compounds have the fucosylated N-acetylchitobiose unit in common, as is evident from the anomeric signals at $\delta = 5.18$ ppm (GlcNAc-1), 4.69 ppm (GlcNAc-1), 4.67 ppm (GlcNAc-2), and 4.89 ppm (Fuc), together with the N-acetyl signals at $\delta = 2.038$ –2.039 ppm (GlcNAc-1), 2.094–2.099 ppm (GlcNAc-2), the Fuc CH $_3$ signals at $\delta = 1.21$ ppm, and the Fuc H-5 signals at $\delta = 4.10$ –4.13 ppm (23). The anomization effect results in doubling of the GlcNAc-2 H-1 and of the Fuc H-1, H-5 and CH $_3$ structural-reporter-group signals.

HPLC of FPLC fraction 1 yields 7 fractions, denoted 1-1 to 1-7 (Fig. 2A). On guidance of the ^1H NMR spectra it was concluded that the fractions contain only monosialylated compounds (Scheme 1, Table I). In the spectrum of fraction 1-1 the typical pattern of the Man H-1 and H-2 signals indicates the presence of a Man $_2$ (1-3) [Man $_2$ (1-6)]Man $_2$ (1-4) unit (25). The NeuAc $_2$ (2-6)Gal $_2$ (1-4)GlcNAc $_2$ (1-2)Man $_2$ (1-3) branch is characterized by the Man-4 H-1 signal at $\delta = 5.135$ ppm together with the anomeric signals of GlcNAc-2 and Gal-4, the H-3e and H-3a signals of NeuAc, and the N-acetyl signals of GlcNAc-2 and NeuAc. The Man $_2$ (1-6) branch is terminated with Man-4, since the Man-4' H-2 signal is not shifted out of the bulk region, due to the absence of GlcNAc-5 (compare compound 23 in ref. 25). In the spectrum of fraction 1-2 the H-2 signal of Man-4' is now resonating at $\delta = 4.109$ ppm, indicating the presence of a GlcNAc-5' residue. This residue is in terminal position, as follows from its structural-reporter-group signals (10, 25). The compound in fraction 1-3 is an extension of that in 1-2 with Gal-6' in terminal position (compare compound 27 in ref. 25). Fraction 1-4 is a mixture of three components, and the structural reporter groups of the compounds identified in 1-3 and 1-5 are present. The presence of a N-glycolyl signal at $\delta = 4.118$ ppm indicates the occurrence of NeuGc in the major component. This is also reflected in the slightly altered chemical shifts of the H-3e and H-3a signals, and of the H-1 signal of Gal-6' (30). The retention time of 1-4 compared to that of 1-3 is in accordance with the replacement of NeuAc by NeuGc (30). The compound in fraction 1-5 is extended at Gal-6' with an α Gal residue as compared to that in fraction 1-3. This is evident from the α Gal H-1 signal at $\delta = 5.146$ ppm and the influence on Gal-6' H-4, which has shifted out of the bulk region to $\delta = 4.185$ ppm (31). In fraction 1-6 a triantennary structure, bearing NeuAc $_2$ (2-6) at Gal-6', is identified as the major component. The assignment is based on the structural-reporter groups of GlcNAc-2 and Gal-6' and the Man H-1 and H-2 pattern (32). Compared to the compound in fraction 1-5, the major compound in fraction 1-7 contains NeuGc in stead of NeuAc. This results in a novel structure with a NeuGc residue in the Man $_2$ (1-3) branch and a terminal α Gal residue in the Man $_2$ (1-6) branch.

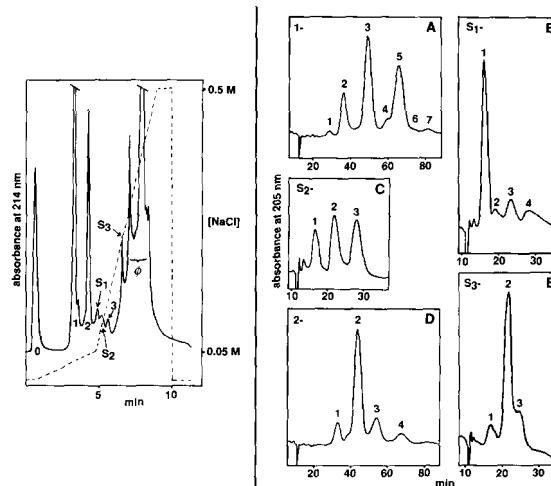


Fig. 1.

Fig. 1. Fractionation pattern of the PNGase-F digestion products derived from porcine thyroglobulin on a FPLC HR 5/5 Mono Q column. The column was eluted with a linear concentration gradient (---) from 0–50 mM NaCl in 8 ml H_2O , followed by a steeper gradient of 50–500 mM NaCl in 8 ml H_2O at a flow rate of 2 ml/min. The injection volume was 0.5 ml. Fractions were collected as indicated.

Fig. 2. Fractionation patterns of porcine thyroglobulin FPLC fractions on HPLC. The 10- μm Lichrosorb-NH $_2$ column (0.46 x 25 cm) was eluted isocratically with 30 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, pH 7.0 / acetonitrile (35:65, v/v) at a flow rate of 1.5 ml/min at room temperature. The injection volume was 50 μl . Fractions were collected as indicated. (A) FPLC fraction 1; (B) fraction S $_1$; (C) fraction S $_2$; (D) fraction 2; (E) fraction S $_3$.

SCHEME 1. Monosialylated nonsulfated structures obtained from porcine thyroglobulin.

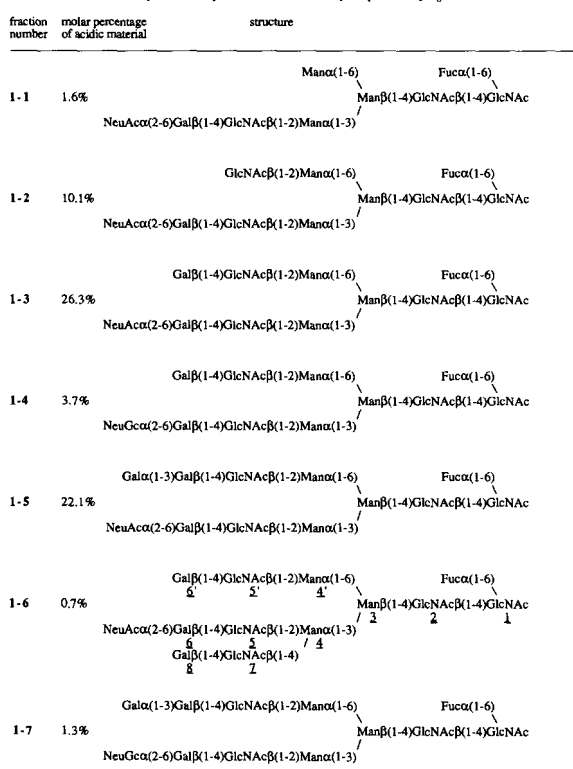


TABLE 1. ^1H Chemical shifts of structural-reporter-group protons of the constituent monosaccharides of the monosialylated nonsulfated and monosialylated sulfated compounds derived from porcine thyroglobulin. Chemical shifts are given at 24°C, in ppm downfield from internal 4,4-dimethyl-4-silapentane-1-sulfonate in $^2\text{H}_2\text{O}$, acquired at 500 MHz (but were actually measured relative to internal acetone: $\delta = 2.225$ ppm). For the complete structures of the compounds, see Schemes 1 and 2. In the table-heading, the structures are represented by short-hand symbolic notation (25): ●, GlcNAc; ◆, Man; ■, β Gal; □, α Gal; ○, NeuAc α (2-6); ◊, NeuGc α (2-6); □, Fuc; S3, 3-O-sulfate; S6, 6-O-sulfate.

FPLC peak:		1 2 3 4 5 6 7 S1 1 S2 3											
HPLC peak:		1 2 3 4 5 6 7 S1 1 S2 3											
Reporter group													
Residue													
H-1	1 α	5.181	5.181	5.181	5.181	5.180	5.181	5.180	5.180	5.181	5.181	5.181	
	1 β	4.692	4.692	4.693	4.692	4.692	4.692	4.692	4.692	4.691	4.691	4.691	
	2 α 1	4.664	4.665	4.665	4.665	4.664	4.665	4.664	4.664	4.665	4.664	4.664	
	2 β 1	4.669	4.669	4.669	4.669	4.669	4.669	4.670	4.668	4.670	4.669	4.668	
	3	n.d.	n.d.	4.773	n.d.	4.772	n.d.	n.d.	n.d.	4.774	4.775	4.775	
	4	5.135	5.135	5.135	5.134	5.135	5.132	5.135	5.134	5.145	5.145	5.143	
	4'	4.919	4.919	4.928	4.928	4.929	4.925	4.927	4.925	4.921	4.928	4.930	
	5	4.603	4.606	4.606	4.607	4.605	4.592	4.607	4.605	4.645	4.645	4.646	
	6	---	---	4.554	4.582	4.582	4.583	4.581	4.582	4.580	4.554	4.582	
	6'	4.445	4.446	4.446	4.450	4.445	4.444	4.450	4.444	4.486	4.486	4.486	
H-2	7	---	---	---	4.472	4.472	4.544	4.472	4.543	4.589	---	4.471	
	7'	---	---	---	---	---	---	---	---	---	---	---	
	8	---	---	---	---	---	5.146	---	5.146	---	---	5.146	
	8'	---	---	---	---	---	---	---	---	---	---	---	
	α Gal	---	---	---	---	---	---	---	---	---	---	---	
	Fuca1	4.886	4.887	4.890	4.889	4.890	4.890	4.889	4.892	4.886	4.889	4.890	
	Fuc1	4.892	4.894	4.897	4.897	4.897	4.898	4.898	4.899	4.894	4.897	4.897	
	10	---	---	---	---	---	---	---	---	---	---	---	
	11	---	---	---	---	---	---	---	---	---	---	---	
	12	---	---	---	---	---	---	---	---	---	---	---	
H-2	3	4.258	4.257	4.257	4.256	4.257	4.20	4.256	4.257	4.257	4.255	4.256	
	4	4.194	4.197	4.198	4.196	4.195	4.222	4.195	4.198	4.215	4.215	4.213	
	5	<4.0	4.109	4.109	4.112	4.111	4.112	4.111	4.112	4.108	4.109	4.110	
	6	---	---	---	---	---	---	---	---	---	---	---	
H-3a NeuAc	1	1.721	1.722	1.722	---	1.722	1.721	---	1.722	1.723	1.724	1.724	
	2	---	---	---	1.739	---	---	1.738	---	---	---	---	
	3	2.667	2.667	2.667	---	2.666	2.667	---	2.667	2.667	2.666	2.667	
	4	---	---	---	2.685	---	---	2.683	---	---	---	---	
H-3a NeuGc	1	---	---	---	---	---	---	---	---	---	---	---	
	2	---	---	---	---	---	---	---	---	---	---	---	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	---	---	---	---	---	---	---	---	---	---	---	
H-3c NeuAc	1	---	---	---	---	---	---	---	---	---	---	---	
	2	---	---	---	---	---	---	---	---	---	---	---	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	---	---	---	---	---	---	---	---	---	---	---	
H-4	1	---	---	---	n.d.	n.d.	n.d.	n.d.	n.d.	4.341	---	n.d.	
	2	---	---	---	n.d.	n.d.	4.185	n.d.	4.186	4.296	---	n.d.	
	3	---	---	---	---	---	---	---	---	---	---	4.183	
	4	---	---	---	---	---	---	---	---	---	---	---	
H-5	1	4.098	4.097	4.098	4.098	4.098	4.099	4.099	4.098	4.099	4.098	4.100	
	2	4.134	4.134	4.135	4.135	4.134	4.133	4.133	4.133	4.133	4.134	4.132	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	---	---	---	---	---	---	---	---	---	---	---	
H-6	5	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	4.442	4.440	4.440	
	6	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	4.308	4.307	4.306	
	7	---	---	---	---	---	---	---	---	---	---	---	
	8	---	---	---	---	---	---	---	---	---	---	---	
CH3	1	1.209	1.209	1.209	1.209	1.209	1.210	1.209	1.210	1.209	1.208	1.209	
	2	1.220	1.220	1.221	1.221	1.220	1.222	1.220	1.221	1.221	1.220	1.221	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	---	---	---	---	---	---	---	---	---	---	---	
NAc	1	2.038	2.038	2.039	2.039	2.038	2.038	2.038	2.039	2.038	2.038	2.038	
	2	2.094	2.095	2.097	2.096	2.097	2.096	2.096	2.097	2.095	2.096	2.097	
	3	2.068	2.069	2.069	2.070	2.069	2.069	2.072	2.069	2.070	2.070	2.070	
	4	---	---	2.052	2.048	2.048	2.048	2.047	2.048	2.048	2.053	2.049	
NeuAc	1	---	---	---	---	---	---	---	---	---	---	---	
	2	---	---	---	---	---	---	---	---	---	---	---	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	2.029	2.029	2.030	---	2.029	2.030	---	2.029	2.029	2.028	2.029	
Gc	1	---	---	---	4.118	---	---	---	4.117	---	---	---	
	2	---	---	---	---	---	---	---	---	---	---	---	
	3	---	---	---	---	---	---	---	---	---	---	---	
	4	---	---	---	---	---	---	---	---	---	---	---	

¹ α and β refer to the anomeric configuration of GlcNAc-1.

SCHEME 3. Di- and trisialylated nonsulfated structures obtained from porcine thyroglobulin.

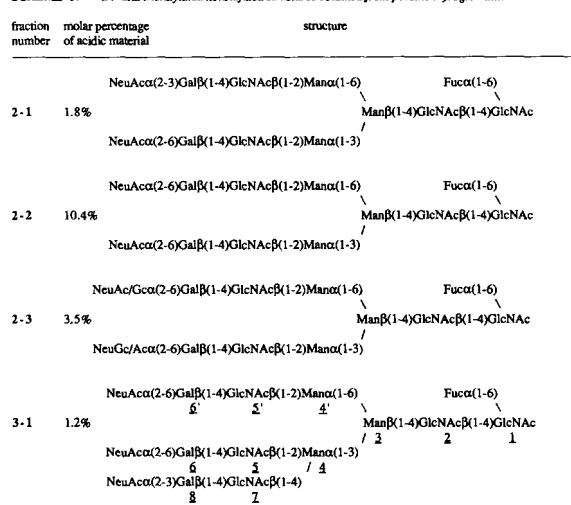


TABLE II. ¹H Chemical shifts of structural-reporter-group protons of the constituent monosaccharides of the di- and trisialylated nonsulfated and disialylated sulfated compounds derived from porcine thyroglobulin. Chemical shifts are given at 24°C, in ppm downfield from internal 4,4-dimethyl-4-silapentane-1-sulfonate in D₂O, acquired at 500 MHz (but were actually measured relative to internal acetone: δ = 2.225 ppm). For the complete structures of the compounds, see Schemes 3 and 4. In the table-heading, the structures are represented by shorthand symbolic notation (25): ●, GlcNAc; ◆, Man; ■, Gal; ○, NeuAcα(2-6); △, NeuAcα(2-3); ⊙, NeuGcα(2-6); □, Fuc; S3, 3-O-sulfate; S6, 6-O-sulfate.

FPLC peak: HPLC peak:		1	2	3	1	S3	2	3	3	1
Reporter group		Residue								
H-1	1α	5.181	5.181	5.181	5.181	5.181	5.181	5.181	5.181	5.181
	1β	4.692	4.691	4.690	4.691	4.690	4.689	4.689	4.689	4.690
	2α ¹	4.667	4.667	4.667	4.667	4.667	4.667	4.667	4.667	4.667
	2β ¹	4.670	4.670	4.670	4.670	4.670	4.670	4.670	4.671	4.671
	3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	4	5.133	5.132	5.133	5.138	5.138	5.136	5.136	5.130	5.130
	4'	4.923	4.942	4.941	4.920	4.943	4.941	4.941	4.933	4.933
	5	4.604	4.604	4.605 ²	4.643	4.643	4.643	4.643	4.592	4.592
	5'	4.573	4.604	4.605 ²	4.572	4.603	4.603	4.603	4.603	4.603
	6	4.444	4.445	4.446 ²	4.486	4.486	4.486	4.486	4.445	4.445
	6'	4.550	4.445	4.446 ²	4.550	4.445	4.449	4.449	4.445	4.445
	7	---	---	---	---	---	---	---	4.547	4.547
	8	---	---	---	---	---	---	---	4.547	4.547
	Fucα ¹	4.893	4.889	4.889	4.894	4.889	4.889	4.889	4.890	4.890
	Fucβ ¹	4.901	4.896	4.898	4.901	4.897	4.900	4.898	4.898	4.898
H-2	3	4.256	4.260	4.259	4.257	4.258	4.256	4.22	4.22	4.22
	4	4.198	4.198	4.198	4.210	4.210	4.209	4.22	4.22	4.22
	4'	4.111	4.112	4.112	4.110	4.112	4.111	4.111	4.111	4.111
H-3a	NeuAc	1.721	1.721	---	1.722	1.722	---	1.721	1.721	1.721
	NeuAc'	1.804	1.721	---	1.805	1.722	---	1.721	1.721	1.721
	NeuAc ³	---	---	1.721	---	---	1.723	---	---	---
	NeuGc ³	---	---	1.737	---	---	1.737	---	---	---
H-3e	NeuAc	---	---	---	---	---	---	1.803	1.803	1.803
	NeuAc'	2.666	2.667	---	2.668	2.667	---	2.667	2.667	2.667
	NeuAc ³	2.758	2.672	---	2.758	2.673	---	2.672	2.672	2.672
	NeuGc ³	---	---	2.67	---	---	2.67	---	---	---
H-5	Fucα ¹	---	---	2.69	---	---	2.69	---	---	---
	Fucβ ¹	---	---	---	---	---	---	2.759	2.759	2.759
	NeuAc	---	---	---	---	---	---	---	---	---
H-6	Fucα ¹	4.099	4.099	4.098	4.100	4.099	4.099	4.100	4.100	4.100
	Fucβ ¹	4.134	4.134	4.134	4.132	4.134	4.134	4.132	4.132	4.132
H-6'	5	n.d.	n.d.	n.d.	4.434	4.433	4.432	n.d.	n.d.	n.d.
	5'	n.d.	n.d.	n.d.	4.313	4.315	4.314	n.d.	n.d.	n.d.
CH ₃	Fucα ¹	1.211	1.210	1.210	1.211	1.210	1.210	1.211	1.211	1.211
	Fucβ ¹	1.223	1.221	1.222	1.223	1.221	1.221	1.223	1.223	1.223
NAc	1	2.039	2.038	2.038	2.039	2.038	2.038	2.038	2.038	2.038
	2	2.097	2.099	2.099	2.097	α ¹ 2.096 β ¹ 2.099	2.098	2.098	2.098	2.098
	3	2.069	2.069	2.069	2.071	2.071	2.070	2.067	2.067	2.067
	3'	2.044	2.069	2.069	2.045	2.067	2.070	2.075	2.075	2.075
	4	---	---	---	---	---	---	---	---	---
	NeuAc	2.030	2.029	2.030	2.030	2.029	2.029	2.030	2.030	2.030
	NeuGc	---	---	4.118	---	---	4.118	---	---	---

¹ α and β refer to the anomeric configuration of GlcNAc-1.

² The resolution of the spectrum did not allow to interpret the differences due to the type of sialic acid.

³ These residues can occur in both branches.

⁴ NeuAc linked to the Gal-β residue.

HPLC of FPLC fraction 2 yields 4 fractions, denoted 2-1 to 2-4 (Fig. 2D). On guidance of the ¹H NMR spectra it was concluded that the fractions contain only disialylated compounds (Scheme 3, Table II). The structural-reporter groups in the spectrum of 2-2 are indicative of a diantennary structure containing two α(2-6)-linked NeuAc residues (compare compounds 29 and 33 in ref. 25). As compared to the compound in 2-2, the α(2-6)-linked NeuAc residue at Gal-β is replaced by α(2-3)-linked NeuAc in the compound of 2-1, which is reflected by the structural-reporter groups of the Gal-β, GlcNAc-2', and Man-4' residues (see compound 39 in ref. 25). As compared to the compound in 2-2, one of the NeuAc residues in that of 2-3 is replaced by a NeuGc residue (compare fraction 1-5 and 1-7). The NeuAc and NeuGc residues afford different chemical shifts for the H-3e atoms when attached to different branches, resulting in 4 possible multiplets for the H-3e atoms (30). Since the H-3e signals are coinciding in the ¹H NMR spectrum of 2-3, both possible structures with a NeuGc residue in one of the branches can occur in fraction 2-3. The retention time of fraction 2-3 compared to that of 2-2 is in accordance with the replacement of one of the NeuAc residues by NeuGc (30). Fraction 2-4 contains too low an amount of material for unambiguous assignment of the ¹H NMR spectrum. However, structural elements of disialylated di- and triantennary structures, α(2-6)-linked NeuAc/Gc, and terminal αGal could be recognized.

HPLC of FPLC-fraction 3 resulted in one major fraction denoted 3-1 (data not shown). The ¹H NMR data of this fraction (Table II) are in full agreement with literature data (compound 41 in ref. 25) of a trisialylated triantennary compound with α(2-6)-linked NeuAc residues at Gal-β and Gal-β' and α(2-3)-linked NeuAc at Gal-β (Scheme 3).

The molar percentage of each fraction compared to the total amount of acidic material, based on the peak areas of the FPLC and HPLC elution patterns, are included in Schemes 1-4. The minor fractions, for which no detailed structure could be elucidated, make up to 100%. A resorcinol sulfuric acid test (39) on a small part of each HPLC fraction revealed similar percentages.