Eur. J. Biochem. *137*, 319–323 (1983)

Structure determination of the carbohydrate chains of rat α-fetoprotein

Bernard BAYARD, Jean-Pierre KERCKAERT, Gérard STRECKER, Lambertus DORLAND, Herman VAN HALBEEK, and Johannes F. G. VLIEGENTHART

Institut de Recherches sur le Cancer de Lille, Unité 124 de l'Institut National de la Santé et de la Recherche Médicale, Lille; Laboratoire de Chimie Biologique, Université des Sciences et Techniques de Lille I, Villeneuve d'Ascq; and Department of Bio-Organic Chemistry, University of Utrecht

(Received May 19/August 29, 1983) – EJB 83 0511

By affinity chromatography on concanavalin A (ConA) linked to Sepharose, *S*-carboxymethylated rat α-fetoprotein could be separated into two molecular variants: a ConA-reactive and a ConA-nonreactive fraction. The carbohydrate chains were quantitatively released from the protein by hydrazinolysis. Based on methylation analysis and high-resolution ^{1}H-NMR spectroscopy of the re-*N*-acetylated hydrazinolysates, the carbohydrate structures of the two ConA-molecular variants of α-fetoprotein were established. The ConA-reactive species contains two *N*-glycosidic carbohydrate units per molecule, both having the following structure:

NeuAc(α2→6)Gal(β1→4)GlcNAc(β1→2)Man(α1→3)

Man(β1→4)GlcNAc(β1→4)GlcNAc.

NeuAc(α2→6)Gal(β1→4)GlcNAc(β1→2)Man(α1→6)

The ConA-nonreactive species possesses also two *N*-glycosidically linked oligosaccharide chains per molecule; each of these has the following structure:

NeuAc(α2→6)Gal(β1→4)GlcNAc(β1→2)Man(α1→3)

GlcNAc(β1→4)–Man(β1→4)GlcNAc(β1→4)GlcNAc.

NeuAc(α2→6)Gal(β1→4)GlcNAc(β1→2)Man(α1→6)

α-Fetoprotein is a transitory glycoprotein present in relatively high concentrations in fetal and neonatal sera and in amniotic fluid of many species. Distinct molecular variants of rat α-fetoprotein have been demonstrated using affinity chromatography on concanavalin A (ConA) [1] or crossed immunoaffino-electrophoresis [2]. Differences in the ConA-affinity pattern between α-fetoprotein from neonatal rat liver, from rat hepatoma and α-fetoprotein produced *in vitro* by fetal hepatocytes in culture [1, 3] prompted us to investigate the structure of the oligosaccharide chains. We have shown previously [4] that each α-fetoprotein variant, whether ConA-reactive or nonreactive, has two carbohydrate chains per molecule which are *N*-glycosidically linked to asparagine. Evidence was obtained that the ConA-reactive α-fetoprotein variant contains two identical ConA-reactive glycans. Also the ConA-nonreactive α-fetoprotein variant possesses two identical oligosaccharide units, which are obviously different from the ConA-reactive glycans.

Since the yield of carbohydrates released from the polypeptide backbone by enzymic digestion is low, we have employed the hydrazinolysis technique for the quantitative splitting of the carbohydrate chains from the glycoprotein [5]. Structure determination of the oligosaccharide chains was carried out on highly purified partial structures, by means of methylation analysis and high-resolution ^{1}H-NMR spectroscopy.

Abbreviations. Gal, D-galactose; GlcNAc, *N*-acetyl-D-glucosamine; Man, D-mannose; NeuAc, *N*-acetyl-D-neuraminic acid; Asn, L-asparagine; ConA, concanavalin A; NMR, nuclear magnetic resonance.

MATERIALS AND METHODS

Purification of α-fetoprotein

α-Fetoprotein was isolated from the amniotic fluid of pregnant (14–18 days of gestation) Wistar rats. The fluids, which contained 1.5–2.2 mg α-fetoprotein/ml, were clarified by centrifugation at 2000 × *g* for 30 min and stored at −20 °C. The glycoprotein was then purified by immuno-absorption on anti-(rat α-fetoprotein) IgG linked to Sepharose 4B [6].

Reduction and alkylation of α-fetoprotein

The glycoprotein was reduced and alkylated according to Crestfield et al. [7] with iodo-[1-^{14}C]acetic acid.

Concanavalin-A – Sepharose chromatography

ConA (Pharmacia Fine Chemicals) coupled with Sepharose 4B, previously activated according to the CNBr method of March et al. [8], was used. Labeled *S*-carboxymethylated rat

α-fetoprotein (30 mg) was passed through the ConA-Sepharose column (2 × 30 cm) equilibrated with 0.05 M Tris/HCl buffer (pH 7.6) containing 1 M NaCl, 1 mM $MnCl_2$, 1 mM $MgCl_2$ and 1 mM $CaCl_2$. The ConA-nonreactive material was removed by washing the column with the aforementioned buffer. Subsequently the ConA-reactive compound was eluted with 0.1 M methyl α-D-glucoside in the same buffer. Each α-fetoprotein fraction was detected by radioactivity, dialysed and lyophilised.

Isolation of oligosaccharide chains from ConA-nonreactive and ConA-reactive α-fetoprotein variants

Each ConA-affinity variant was subjected to hydrazinolysis, and then re-*N*-acetylated, as previously described [4, 5, 9], starting with 10 mg of glycoprotein.

Thin-layer chromatography

^{14}C-labeled oligosaccharide samples (6000 counts/min in 5 µl) were analyzed by thin-layer chromatography [4] and located by fluorography after impregnation with 5% 2.5-diphenyloxazol in ether.

Carbohydrate composition

The qualitative and quantitative carbohydrate composition of the oligosaccharides was estimated by gas-liquid chromatography according to Zanetta et al. [10].

Methylation analysis

Methylation analysis was performed according to Finne et al. [11]. Partially *O*-methylated methyl glycosides resulting from the methanolysis of permethylated glycans were *N* and *O*-acetylated with acetic anhydride/pyridine (1:1, v/v) and identified by gas-liquid chromatography in combination with mass spectrometry [12] on a Riber-Mag, model 10-10, mass spectrometer coupled with a capillary glass column (0.35 mm × 60 m) coated with OV-101 (temperature programme: 120–220 °C, 4 °C/min).

^{1}H-NMR spectroscopy

For ^{1}H-NMR spectroscopic analysis the α-fetoprotein carbohydrate preparations were repeatedly treated with ^2H_2O at room temperature. After each exchange treatment the materials were lyophilised.

360-MHz ^{1}H-NMR spectra were recorded on a Bruker HX-360 spectrometer and 500-MHz ^{1}H-NMR spectra on a Bruker WM-500 instrument (SON NMR facilities, Groningen and Nijmegen, The Netherlands). Both machines were run in the Fourier transform mode at a probe temperature of 300 K. For further details, see [13]. Chemical shifts are expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured by reference to internal acetone ($\delta = 2.225$ ppm).

RESULTS

Preparation and characterization of α-fetoprotein oligosaccharide units

As previously shown [2], ConA affinity chromatography of rat α-fetoprotein gives three fractions: a ConA-nonreactive, a ConA-weakly-reactive and a ConA-reactive species. The ConA-weakly-reactive species shows hydrophobic interaction with ConA. One can avoid this nonspecific carbohydrate interaction by *S*-carboxymethylation of the glycoprotein [4]. The ConA-nonreactive variant represents 60% and the ConA-reactive 40% of the alkylated α-fetoprotein. The carbohydrate chains from each *S*-carboxymethylated α-fetoprotein ConA variant were released by exhaustive hydrazinolysis followed by quantitative re-*N*-acetylation with [^{14}C]acetic anhydride. The glycans were detected by fluorography of the thin-layer chromatograms. Two kinds of oligosaccharide chain were characterized by thin-layer chromatography, a slow and a fast-moving carbohydrate, originating from the ConA-nonreactive and the ConA-reactive alkylated α-fetoprotein, respectively. These two carbohydrate chains possess the same neutral and acidic monosaccharide composition (2 NeuAc, 2 Gal and 3 Man residues each); they differ only in their *N*-acetylglucosamine content: 5 GlcNAc residues were found for the ConA-nonreactive α-fetoprotein species and 4 for the ConA-reactive α-fetoprotein species.

Table 1. *Molar ratios of partially O-methylated methylglycosides present after methylation analysis of the two ConA molecular variant oligosaccharides from α-fetoprotein*

3,4,6-Me_3-Man = 2-*O*-acetyl-3,4,6-tri-*O*-methyl-methylmannoside; other abbreviations similarly. The molar ratios were determined on the basis of two residues of 3,4,6-Me_3-Man

Partially *O*-methylated monosaccharide	Residues in	
	ConA-reactive oligosaccharide	ConA-nonreactive oligosaccharide
2,3,4,6-Me_4-Man	–	0.05 (0)
2,3,4,6-Me_4-Gal	–	0.1 (0)
3,4,6-Me_3-Man	2.00 (2)[a]	2.00 (2)[a]
2,3,4-Me_3-Gal	1.86 (2)	2.26 (2)
2,4-Me_2-Man	0.95 (1)	–
3,4,6-Me_3-GlcN(Me)Ac	–	1.15 (1)
3,6-Me_2-GlcN(Me)Ac	3.4 (4)[b]	2.60 (4)[b]
2-Me-Man	0.1 (0)	0.76 (1)
4,7,8,9-Me_4-Neu(Me)Ac	1.5 (2)	1.85 (2)

[a] The number given in brackets corresponds to the theoretical value.

[b] The low values for 3,6-Me_2-GlcN(Me)Ac (3.4 and 2.6 instead of 4) are due to the formation of GlcNAc-NAc hydrazone from GlcNAc-1 [15]

Methylation analysis

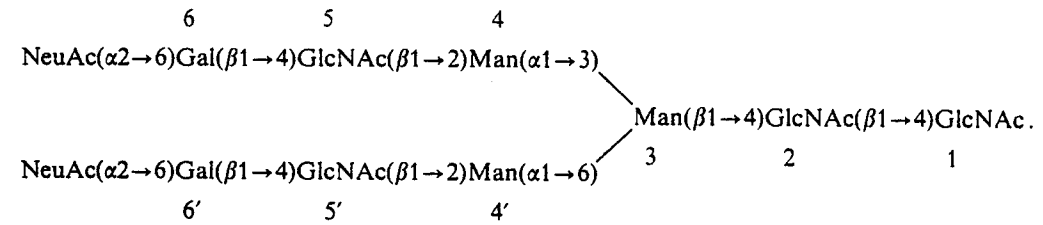
The results of the methylation analysis of the two ConA molecular variant oligosaccharides of α-fetoprotein are summarized in Table 1. The identity of the partially methylated monosaccharides was confirmed by mass spectrometry. The molar ratios of partially methylated monosaccharides obtained from the ConA-reactive α-fetoprotein variant (see Table 1) suggest a classical diantennary structure of the *N*-acetyllactosamine type, containing two α(2→6)-sialylated branches. The occurrence of traces of 2-*O*-methyl mannose (2-Me-Man) points to contamination of the ConA-reactive material with small amounts of the ConA-nonreactive oligosaccharide (see below).

As compared to the ConA-reactive material, the methylation of the ConA-nonreactive oligosaccharides revealed the additional presence of 3,4,6-Me_3-GlcN(Me)Ac, in conjunction with the presence of 2-Me-Man instead of 2,4-Me_2-Man. These results suggest that the ConA-nonreactive glycan is a sialo, diantennary structure of the *N*-acetyllactosamine type, which possesses an additional GlcNAc residue linked to C-4 of Man-3 (so called intersecting GlcNAc).

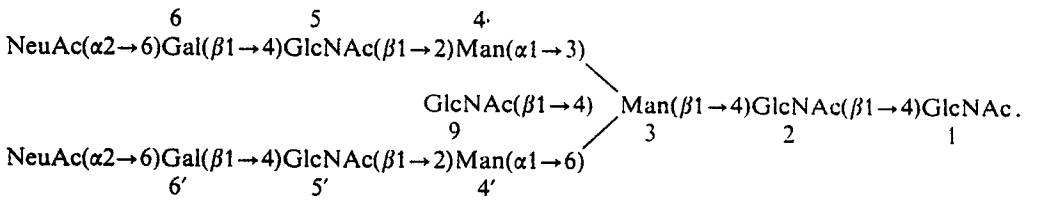
^{1}H-NMR spectroscopy

High-resolution ^{1}H-NMR spectra were recorded for the ConA-reactive and the ConA-nonreactive oligosaccharide preparations, obtained from rat α-fetoprotein by hydrazinolysis. The chemical shifts of pertinent structural reporter groups for these two glycans, together with those for some suitable reference compounds, are compiled in Table 2.

The NMR spectral parameters for the ConA-reactive preparation from α-fetoprotein show such a close resemblance with those acquired [13, 14] for the sialo, diantennary glycopeptide derived from human serotransferrin (see Table 2), that the peripheral part of the former glycan can be readily inferred to be identical to that of the reference compound. The structural reporter group signals of Man-3, -4 and -4′ demonstrate the presence of the mannotriose core unit, and its diantennary type of branching [13]. The chemical shifts of NeuAc H-3*ax* and H-3*eq* point to the presence of NeuAc in (α2→6)-linkage to Gal. The occurrence of NeuAc in both branches is revealed by the resonance positions of the signals for Gal-6 and -6′ H-1, GlcNAc-5 and -5′ *N*-acetyl and Man-4 and -4′ H-1 [13], as well as by the relative intensities of the NeuAc H-3*ax* and H-3*eq* signals, each accounting for two protons. The quality of the 360-MHz spectrum of the ConA-reactive oligosaccharide is not sufficient to derive detailed information about the structure of the GlcNAc-1 residue partially converted into GlcNAc hydrazone [15]. In conclusion, the structure of the ConA-reactive glycan from α-fetoprotein has been established to be the following:



Concerning the ConA-nonreactive molecular variant oligosaccharide of α-fetoprotein, the chemical shifts of the structural reporter groups of Man-3, -4 and -4′ indicate a diantennary type of branching of the mannotriose core, with the additional presence of GlcNAc-9 in (β1→4)-linkage to Man-3. This is corroborated by the presence of a GlcNAc-9 H-1 signal (at δ = 4.472 ppm) and an additional *N*-acetyl signal (at δ = 2.058 ppm). The NMR parameters of the branching unit of this α-fetoprotein glycan are similar to those for the β-anomer of the oligosaccharide '5', isolated from the urine of a patient with Morquio syndrome type B [16]. However, at the nonreducing terminus of both *N*-acetyllactosamine branches, NeuAc is present in (α2→6)-linkage to Gal (δH-3*ax* = 1.719 ppm; δH-3*eq* = 2.675 ppm). Comparison of the data for the α-fetoprotein preparation with those for the Morquio oligosaccharide (see Table 2) shows the well-known (compare with [13]) effects of extension of an *N*-acetyllactosamine branch with NeuAc in (α2→6)-linkage, on the H-1 signals of Man-4 ($\Delta\delta$ = 0.022 ppm), Man-4′ ($\Delta\delta$ = 0.012 ppm), GlcNAc-5 and -5′ ($\Delta\delta \approx$ 0.015 ppm), Gal-6 and -6′ ($\Delta\delta$ = −0.022 ppm), and on the *N*-acetyl signals of GlcNAc-5 and -5′ ($\Delta\delta \approx$ 0.015 ppm). Thus, the structure of the ConA-nonreactive α-fetoprotein oligosaccharide chain has been established to be:



The influences of the attachment of GlcNAc-9 on the chemical shifts of structural reporter groups of Man-3, -4 and -4′, as observed in the step from the ConA-reactive to the ConA-nonreactive α-fetoprotein oligosaccharide chain (see Table 2), are completely analogous to those described previously for similar elongations [16–19].

DISCUSSION

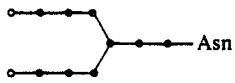
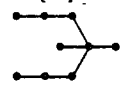
The structure of the diantennary oligosaccharide unit derived from the ConA-reactive variant of rat α-fetoprotein is a classical one which occurs frequently in glycoproteins [20]. The structure of the ConA-nonreactive glycan of rat α-fetoprotein which contains an additionnal GlcNAc residue linked (β1→4) to the β-Man is a less common carbohydrate structure; so far, it has been found only in a few glycan structures, such as IgA myeloma protein α_1 subclass [21], ovotransferrin [22] and glycophorin [23]. This intersecting GlcNAc residue is absent in human α-fetoprotein [24].

From the present work in conjunction with previous data [4] it can now be concluded that, in rat, (a) there are two molecular forms of α-fetoprotein, ConA-reactive and ConA-nonreactive; (b) each molecular form contains a pair of identical *N*-glycosidic glycans; (c) the ConA-reactive α-fetoprotein bears two ConA-reactive diantennary glycans and the ConA-nonreactive α-fetoprotein possesses two ConA-nonreactive glycans, which are diantennary with an intersecting GlcNAc.

Different structures and conclusions have been presented recently [25] for bovine α-fetoprotein glycans. This α-fetoprotein also contains two *N*-glycosidic glycans per molecule with two distinct, namely diantennary and a triantennary, structures. ConA-nonreactive bovine α-fetoprotein was shown

Table 2. *^1H chemical shifts of structural reporter groups of constituent monosaccharides for the two ConA molecular variant carbohydrate preparations obtained from rat α-fetoprotein and for some reference compounds*

Chemical shifts are given at 300 K, relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O (but were actually measured relative to internal acetone: δ = 2.225 ppm). Data for the α-fetoprotein ConA-reactive oligosaccharide were aquired at 360 MHz; those for the other three glycans at 500 MHz. For comparison with the α-fetoprotein ConA-nonreactive oligosaccharide possessing GlcNAc-2 in β-linkage to (the hydrazinolysis product of) GlcNAc-1, only the chemical shifts for the β-anomer of the Morquio disease oligosaccharide are given. ●—— = neutral or amino sugar residue; ○—— = NeuAc (α2→6). n.d. = not determined

Reporter group	Residue	Chemical shift in α-fetoprotein ConA-reactive preparation	serotransferrin glycopeptide [13, 14] 	α-fetoprotein ConA-nonreactive preparation	Morquio disease oligosaccharide '5' [16] 
		ppm			
H-1	GlcNAc-1	n.d	5.088	n.d.	–
	GlcNAc-2	4.613	4.616	4.614	4.721
	Man-3	4.77	4.773	n.d.	4.686
	Man-4	5.136	5.133	5.079	5.057
	Man-4′	4.943	4,949	5.032	5.020
	GlcNAc-5	4.60	4.603	4.599	4.583
	GlcNAc-5′	4.60	4.603	4.599	4.591
	Gal-6	4.440	4.442	4.446	4.468
	Gal-6′	4.440	4.447	4.446	4.477
	GlcNAc-9	–	–	4.472	4.468
H-2	Man-3	4.260	4.254	4.184	4.175
	Man-4	4.197	4.195	4.267	4.258
	Man-4′	4.121	4.116	n.d.	4.142
H-3*ax*	NeuAc	1.719[a]	1.716[a]	1.719[a]	–
H-3*eq*	NeuAc	2.663[a]	2.666 2.672	2.675[a]	–
Nac	GlcNAc-1	n.d.	2.002	n.d.	–
	GlcNAc-2	2.078	2.081	2.082	2.060
	GlcNAc-5	2.064	2.067	2.067	2.053
	GlcNAc-5′	2.064	2.063	2.065	2.039
	GlcNAc-9	–	–	2.058	2.064
	NeuAc	2.026[b]	2.029 2.028	2.031	–

[a] Signal of two protons
[b] Signal of two methyl groups

to contain two triantennary glycans whereas it was deduced that ConA-reactive α-fetoprotein consists of a mixture of two molecular sub-fractions: the first containing one triantennary and one diantennary glycan, the second bearing two diantennary glycans. The existence of such mixed glycan structures within rat α-fetoprotein molecular variants, i.e. one diantennary and one diantennary with an intersecting GlcNAc, has been excluded in our case [4]. Indeed, in certain conditions of ConA-Sepharose affinity chromatography [26] one can characterize the two types of glycan in the α-fetoprotein fraction which is bound on the column. However, we have shown that after reduction and *S*-carboxymethylation, this ConA-reactive α-fetoprotein fraction contains in fact about 15% ConA-nonreactive alkylated α-fetoprotein. Each of these alkylated ConA-reactive and ConA-nonreactive α-fetoprotein fractions were shown further to contain a pair of identical glycans [4].

The authors wish to thank Danielle Roux, Sabine Quief and Yves Leroy for their expert technical assistance. This investigation was supported by the *Institut National de la Santé et de la Recherche Médicale* (CRL 78.5.172.2), The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO), and by The Netherlands Foundation for Cancer Research (KWF, grant UUKC-OC 79-13).

REFERENCES

1. Smith, C. J. & Kelleher, P. C. (1980) *Biochim. Biophys. Acta, 605*, 1–32..
2. Kerckaert, J. P., Bayard, B. & Biserte, G. (1979) *Biochim. Biophys. Acta, 576*, 99–108.
3. Fournier, C., Kerckaert, J. P., Bayard, B., Collyn, M. & Biserte, G. (1979) *Protides Biol. Fluids Proc. 27th Colloq. Bruges, 27*, 623–626.
4. Bayard, B. & Kerckaert, J. P. (1981) *Eur. J. Biochem. 113*, 405–414.
5. Bayard, B. & Roux, D. (1975) *FEBS Lett. 55*, 206–211.
6. Kerckaert, J. P., Bayard, B., Debray, H., Sautière, P. & Biserte, G. (1977) *Biochim. Biophys. Acta, 493*, 293–303.
7. Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622–627.

8. March, S. C., Parikh, I. & Cuatrecasas, P. (1974) *Anal. Biochem. 60*, 149–152.
9. Bayard, B. & Fournet, B. (1976) *Carbohyd. Res. 46*, 75–86.
10. Zanetta, J. P., Breckenridge, W. C. & Vincendon, G. (1972) *J. Chromatogr. 69*, 291–304.
11. Finne, J., Krusius, T. & Rauvala, H. (1980) *Carbohydr. Res. 80*, 336–339.
12. Fournet, B., Strecker, G., Leroy, Y. & Montreuil, J. (1981) *Anal. Biochem. 116*, 489–502.
13. Vliegenthart, J. F. G., van Halbeek, H. & Dorland, L. (1981) *Pure Appl. Chem. 53*, 45–77.
14. Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J. C., Fournet, B., Spik, G. & Montreuil, J. (1978) *Eur. J. Biochem. 87*, 323–329.
15. Saeed, M. S. & Williams, J. M. (1980) *Carbohydr. Res. 84*, 83–94.
16. Michalski, J. C., Strecker, G., van Halbeek, H., Dorland, L. & Vliegenthart, J. F. G. (1982) *Carbohydr. Res. 100*, 351–363.
17. Strecker, G., Herlant-Peers, M. C., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G. & Farriaux, J. P. (1977) *Eur. J. Biochem. 81*, 165–171.
18. Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Spik, G., Fournet, B. & Montreuil, J. (1979) *Eur. J. Biochem. 100*, 569–574.
19. Herlant-Peers, M. C., Montreuil, J., Strecker, G., Dorland, L., van Halbeek, H., Veldink, G. A. & Vliegenthart, J. F. G. (1981) *Eur. J. Biochem. 117*, 291–300.
20. Montreuil, J. (1982) *Comprehensive Biochemistry*, vol. 19B/II, pp. 1–188, Elsevier, Amsterdam.
21. Baenziger, J. & Kornfeld, S. (1974) *J. Biol. Chem. 249*, 7260–7269.
22. Spik, G., Fournet, B. & Montreuil, J. (1979) *C. R. Acad. Sci. Paris 288D*, 967–970.
23. Yoshima, H., Furthmayr, H. & Kobata, A. (1980) *J. Biol. Chem. 255*, 9713–9718.
24. Yoshima, H., Mizuochi, T., Ishii, M. & Kobata, A. (1980) *Cancer Res. 40*, 4276–4281.
25. Krusius, T. & Ruoslahti, E. (1982) *J. Biol. Chem. 257*, 3452–3457.
26. Kerckaert, J. P. & Bayard, B. (1980) *Biochem. Biophys. Res. Commun. 92*, 95–101.

B. Bayard and J. P. Kerckaert, Groupe de Recherches Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses (Unité 124 de l'Institut National de la Santé et de la Recherche Médicale), Institut de Recherches sur le Cancer de Lille, Boîte postale 311, F-59020 Lille-Cedex, Nord, Fance

G. Strecker, Laboratoire de Chimie Biologique (et Laboratoire associé au Centre National de la Recherche Scientifique 217), Université des Sciences et Techniques de Lille I, Boîte postale 36, F-59650 Villeneuve-d'Ascq, Nord, France

L. Dorland, H. van Halbeek, and J. F. G. Vliegenthart, Bio-organisch-Chemisch Laboratorium, Rijksuniversiteit Utrecht, Croesestraat 79, NL-3522-AD Utrecht, The Netherlands