STRUCTURE ELUCIDATION OF OLIGOMANNOSIDE-TYPE ASPARAGINE-BOUND CARBOHYDRATE CHAINS OF GLYCOPROTEINS BY 500 MHz ^{1}H NMR SPECTROSCOPY

H. VAN HALBEEK, L. DORLAND, G. A. VELDINK, J. F. G. VLIEGENTHART, J.-C. MICHALSKI+, J. MONTREUIL+, G. STRECKER+ and W. E. HULL*

*Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, 3522 AD Utrecht, The Netherlands, +Laboratoire de Chimie Biologique, Université des Sciences et Techniques de Lille I, Villeneuve d'Ascq, France and *Bruker Analytische Messtechnik, GmbH, D-7512 Rheinstetten-Fo., FRG*

Received 15 September 1980

1. Introduction

High-resolution ^{1}H NMR spectroscopy is an efficient and powerful technique for the structure determination of a wide variety of carbohydrate chains derived from *N*- and *O*-type glycoproteins [1–6]. However, for oligomannoside-type [7] glycopeptides and oligosaccharides from *N*-type glycoproteins the fully discriminative deduction of the primary structure from the ^{1}H NMR spectrum remained difficult due to the similarity of the constituting units.

The availability of a 500 MHz ^{1}H NMR spectrometer in conjunction with a more sophisticated computer resolution enhancement routine afforded a significant refinement of the spectral data of *N*-acetyllactosamine-type carbohydrate chains of glycoproteins, and the linewidth of the signals in the spectrum was made a useful parameter for structural assignments [5].

This new development opened the possibility to derive the structures of oligomannoside-type carbohydrate chains from their ^{1}H NMR spectra. This will be illustrated for 3 glycoasparagines isolated from the urine of a patient with Gaucher's disease, a glucocerebrosidase deficiency [8].

2. Materials and methods

The glycopeptides dealt with here were isolated from the urine of a patient suffering from Gaucher's disease, and fractionated in the usual way [7]. Details will be described elsewhere.

The molar ratios of neutral monosaccharides and hexosamines were determined after methanolysis [9]. Molar ratios of glucosamine and aspartic acid were determined on a Beckman amino acid analyzer, after hydrolysis in 4 N CF_3COOH for 4 h at 100°C.

For NMR spectroscopy the glycopeptides were repeatedly exchanged in D_2O (99.96 atom% D, Aldrich) with intermediate lyophilization.

The 500 MHz ^{1}H NMR spectra of neutral solutions of the compounds were recorded on a Bruker WM-500 spectrometer, operating in the Fourier transform mode at a probe temperature of 300 K. Chemical shifts are given relative to sodium-2,2-dimethyl-2-silapentane-5-sulphonate (indirectly to acetone in D_2O: δ = 2.225 ppm). Resolution enhancement of the 500 MHz spectra was achieved by Lorentzian to Gaussian transformation according to [10].

3. Results and discussion

From the urine of a patient with Gaucher's disease 3 major glycopeptides were isolated. Sugar and amino acid analyses revealed that only mannose, *N*-acetylglucosamine and asparagine were present in these compounds.

To elucidate the primary structures of the glycoasparagines, 500 MHz ^{1}H NMR spectra of the compounds in D_2O were recorded. The structural reporter group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectra of the 3 glycopeptides, viz. the signals of the anomeric protons, the mannose H-2 resonances and the *N*-acetyl proton singlets are given in fig. 1–3. In view of their molar carbohydrate com-

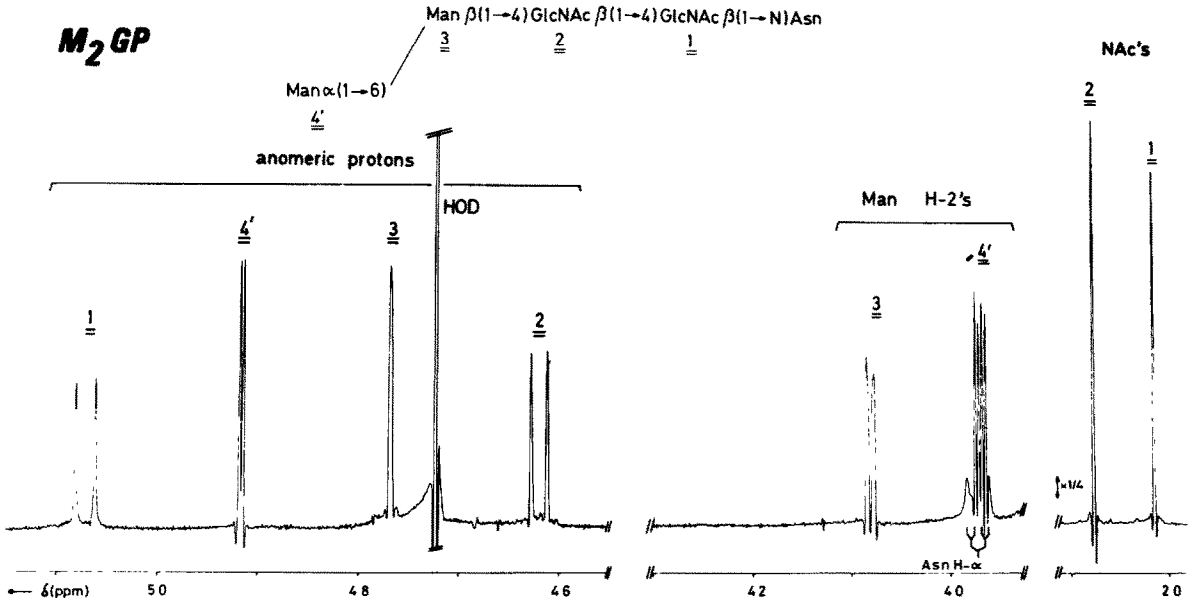


Fig.1. Structural reporter group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectrum of M_2GP, a glycoasparagine containing 2 mannoses, in D_2O at 300 K. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated.

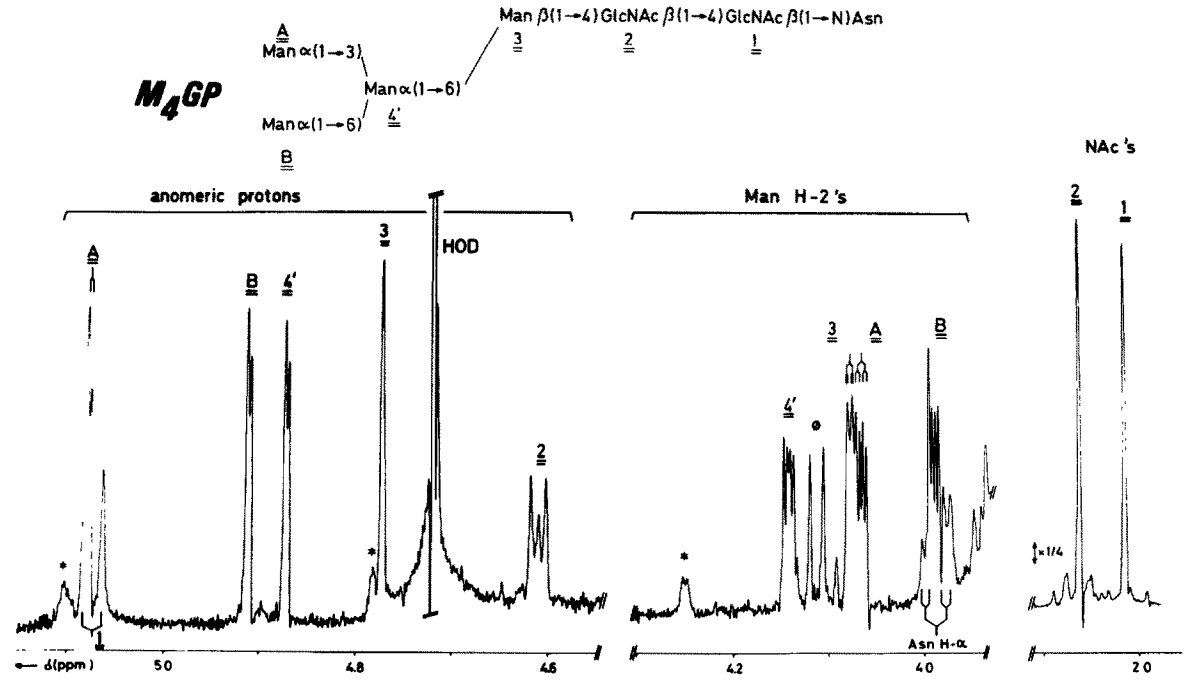


Fig.2. Structural reporter group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectrum of M_4GP, a glycoasparagine of the oligomannoside type containing 4 mannoses, in D_2O at 300 K. The numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The spectrum shows signals of low intensity stemming from contaminants with related carbohydrate structures; those marked by asterisks belong to M_5GP. The quartet at $\delta \simeq 4.11$ ppm, indicated by ϕ, stems from a non-carbohydrate non-protein contaminant (see also fig.4).

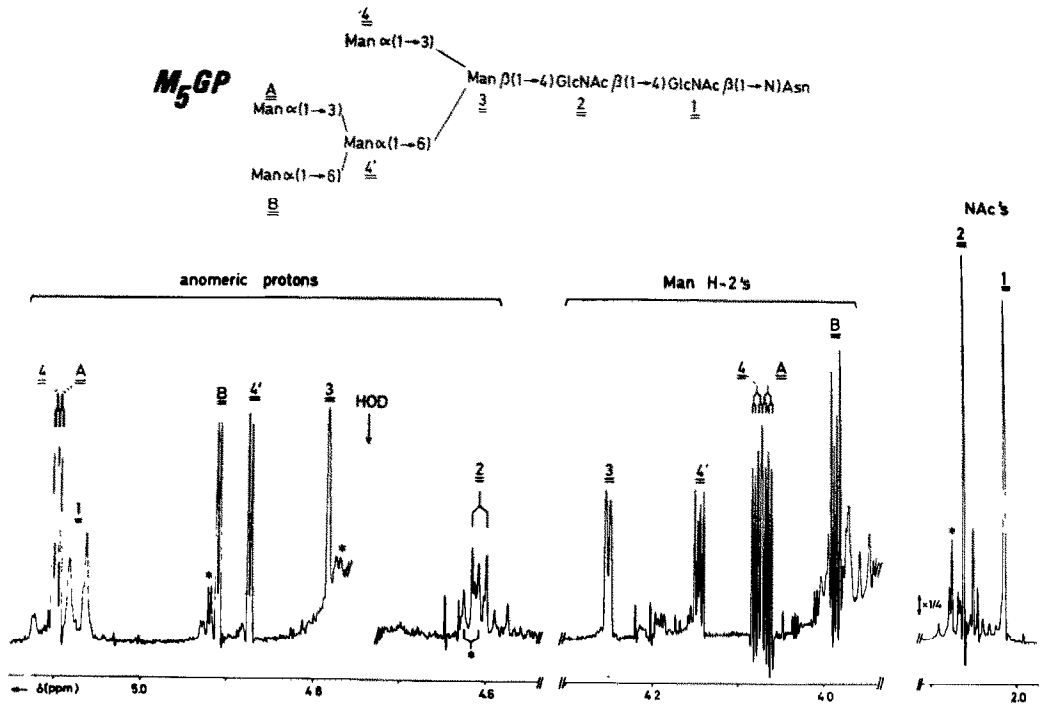


Fig.3. Structural reporter group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectrum of M_5GP, a glycoasparagine of the oligomannoside type containing 5 mannoses, in D_2O at 300 K. The numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The HOD-resonance is left out from the spectrum; its position is indicated by an arrow. Compound M_2GP is present in small amount in the sample (see signals marked by asterisks), which contains a few other low amount carbohydrate contaminants.

positions, estimated by integration of the anomeric proton regions of their ^{1}H NMR spectra (see table 1), the glycopeptides are designated with subscripts after their mannose content M_2GP, M_4GP and M_5GP. As an example for the 3 compounds, the overall spectrum of M_4GP is depicted in fig.4. Relevant NMR parameters for the 3 glycopeptides are listed in table 2.

Comparison of the 500 MHz ^{1}H NMR spectra of M_2GP, M_4GP and M_5GP reveals that the resonance positions as well as the patterns of the signals belonging to corresponding anomeric protons (at $\delta \simeq 5.07$ ppm; $J_{1,2} = 9.8$ Hz and at $\delta \simeq 4.61$ ppm; $J_{1,2} = 8.2$ Hz) and to corresponding *N*-acetyl methyl proton (at $\delta \simeq 2.01$ and $\simeq 2.06$ ppm) of the GlcNAc residues 1 and 2, respectively, are identical. The same holds for the asparagine proton signals (δH-$\alpha \simeq 3.98$ ppm; δH-$\beta \simeq 2.86$ ppm and δH-$\beta' \simeq 2.93$ ppm). These similarities in the spectra indicate that the glycopeptides have in common the (●→4)GlcNAcβ(1→4)-GlcNAcβ(1→N)Asn moiety, as usual for *N*-glycosidic carbohydrate structures [5].

The two additional mannose residues of M_2GP can be characterized as follows. The anomeric proton resonating at $\delta = 4.767$ ppm belongs to the β-linked

Table 1
Carbohydrate and amino acid compositions of three oligomannoside-type glycopeptides isolated from Gaucher's disease urine[a]

Constituent	Glycopeptides		
	M_2GP	M_4GP	M_5GP
Mannose	2	4	5
N-Acetylglucosamine	2	2	2
Asparagine	1	1	1

[a] The molar compositions were estimated by means of NMR spectral integration, on the basis of one residue of asparagine

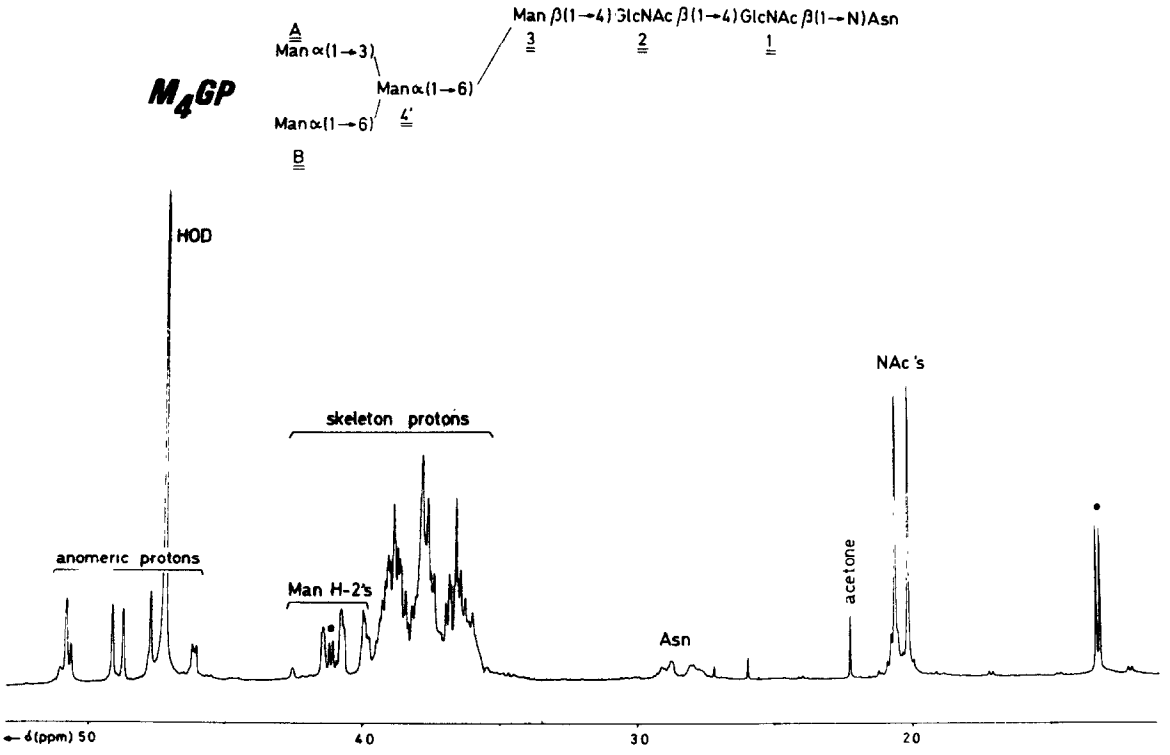


Fig.4. The overall 500 MHz ^{1}H NMR spectrum of M_4GP, a glycoasparagine of the oligomannoside type containing 4 mannoses, in D_2O at 300 K. The quartet at $\delta \simeq 4.11$ ppm, together with the doublet at $\delta \simeq 1.32$ ppm, indicated by ϕ, stem from a non-carbohydrate non-protein contaminant.

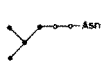
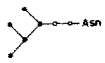
Man 3 (cf. [11]). Its coupling constant $J_{1,2}$ (≃0.6 Hz) is indicative of the β-type of linkage between Man 3 and GlcNAc 2. The relatively small value of $J_{1,2}$ gives rise to an apparent broad lined singlet for H-1, which can easily be distinguished from the well-resolved doublet of H-1 of an α-linked mannose residue [5].

The chemical shift of H-2 of Man 3 for compound M_2GP (δ = 4.080 ppm) reflects a mono-α(1→6)-substitution of Man 3 by another mannose, which is usually numbered Man 4'. Type and configuration of the Man→Man linkage in M_2GP are proved to be α(1→6) by the chemical shift of the H-1 of Man 4' (δ = 4.915 ppm) [11,12] and its $J_{1,2}$ (1.8 Hz) [5]. In accordance with the terminal position of Man 4' the chemical shift of its H-2 is δ = 3.968 ppm. It should be noted that in more complex structures possessing this element this H-2 signal is buried in the bulk of non-anomeric sugar skeleton protons, as in [11,12]. The signals of the (structural reporter group) protons of Man 4' are marked by relatively narrow linewidths, corresponding with the expected mobility of the residue. The structure of M_2GP as given in fig.1 is the afuco-analogue of a glycoasparagine isolated from the urine of a patient with fucosidosis [11].

In the spectrum of M_4GP (fig.2) the signals at δ = 4.770 ppm (H-1) and at δ = 4.076 ppm (H-2) point again to a mono-α(1→6)-substitution of Man 3. Therefore M_4GP contains Man 4'. The well-resolved doublet at δ = 4.909 ppm ($J_{1,2}$ = 1.8 Hz) and the narrow doublet of doublets at δ = 3.988 ppm closely resemble the Man 4' structural reporter group signals in the spectrum of M_2GP, also with respect to their linewidths. Therefore, a terminal α(1→6)-linked mannose residue is present in M_4GP, designated Man B.

The H-1 doublet at δ = 5.076 ppm ($J_{1,2}$ = 1.8 Hz), in conjunction with the H-2 resonance at δ = 4.064 ppm, are characteristic for a terminal mannose, α(1→3)-linked to another mannose residue, as can be derived from earlier observations, for example, Man 4 in the monosialo-oligosaccharide NeuAcα(2→6)-Galβ(1→4)GlcNAcβ(1→2)Manα(1→6)[*Manα(1→3)*]-

Table 2
^{1}H chemical shifts of structural reporter groups of constituent monosaccharides for three oligomannoside-type glycopeptides isolated from Gaucher's disease urine

Reporter group	Residue[a]	Compound and schematic structure[b]		
		M_2GP	M_4GP	M_5GP
				
H-1 of	1	5.071	5.069	5.071
	2	4.618	4.608	4.606
NAc of	1	2.014	2.013	2.012
	2	2.076	2.061	2.060
H-1 of	3	4.767	4.770	4.781
	4	–	–	5.099
	4′	4.915	4.870	4.872
	A	–	5.076	5.093
	B	–	4.909	4.908
H-2 of	3	4.080	4.076	4.251
	4	–	–	4.077
	4′	3.968	4.140	4.144
	A	–	4.064	4.066
	B	–	3.988	3.985

[a] For coding of monosaccharide residues and complete structures see fig.1–3
[b] (●—) Mannose; (○—) *N*-acetylglucosamine

Manβ(1→4)GlcNAc, isolated from human meconium [12] and Man 4 in the trisaccharide *Man*α(1→3)-Manβ(1→4)GlcNAc [13]. The α(1→3)-linked mannose residue in M_4GP cannot be attached to Man 3, since the latter is mono-α(1→6)-substituted, nor to Man B which is a terminal residue. The fourth mannose in M_4GP, designated Man A, must therefore be attached to C-3 of Man 4′.

The H-1 signal of Man 4′ is found at δ = 4.870 ppm ($J_{1,2}$ = 1.8 Hz). The presence of an α(1→3)-linked mannose residue hardly influences the resonance position of the H-1 of the residue to which it is attached (cf. [5] vs. [11]). The observed shift decrement for H-1 of Man 4′, as compared to M_2GP, Δδ = −0.045 ppm, must therefore be ascribed to the substitution of this residue at C-6. A similar shift decrement is observed for H-1 of Man 4′ in the step from tri- to tetra-antennary glycopeptide of the *N*-acetyllactosamine type [5], involving extension with a Galβ(1→4)GlcNAc moiety, β(1→6)-linked to Man 4′. In contrast, the shift increment for H-2 of Man 4′ as compared to M_2GP, can be attributed to the attachment of Man A. A similar shift increment is observed for H-2 of Man 3 extending M_2GP to the meconium oligosaccharide described above [12]. An α(1→6) substitution of one mannose by another hardly affects the former's H-2 chemical shift [2]. The above findings lead to the structure for M_4GP shown in fig.2.

The 500 MHz ^{1}H NMR spectrum of M_5GP (fig.3) differs in the region of the anomeric proton signals from that of M_4GP (fig.2) only by the occurrence of an additional doublet at δ = 5.099 ppm ($J_{1,2}$ = 1.9 Hz). This suggests that M_5GP is an extension of M_4GP with another terminal α(1→3)-linked mannose residue. The residue to which this fifth mannose is attached, can be inferred from the chemical shifts of the mannose H-2s. The characteristically-shaped H-2 signal of the β-Man 3 is shifted from δ = 4.076 ppm to δ = 4.251 ppm going from the spectrum of M_4GP to M_5GP, indicating a further substitution of Man 3 at C-3. The other H-2 signals remain at essentially the same positions and an additional H-2 resonance is found at δ = 4.077 ppm. By consequence, the fifth mannose residue in M_5GP is identified as Man 4, α(1→3)-linked to Man 3. The chemical shifts of H-1 and H-2 of this Man 4 are essentially the same as those described for the terminal Man 4 in the trisaccharide Manα(1→3)Manβ(1→4)GlcNAc [13]. The set of chemical shifts of H-1 and H-2 of Man 3 in M_5GP reflect the completeness of the trimannosyl (3,4,4′)-*N*,*N*′-di-acetylchitobiose-Asn core (cf. [5]). The observed change of the chemical shift of H-1 of Man A, going from M_4GP to M_5GP, is remarkable; it may reflect a spatial effect. The δ-values of the anomeric protons and the H-2s of the 2 terminal α(1→3)-linked mannose residues are very similar (cf. [14]). The most likely assignment has been given in table 2.

The structure of M_5GP proposed on the basis of its NMR data, is given in fig.3.

4. Concluding remarks

The resolution-enhanced 500 MHz ^{1}H NMR spectra of 3 glycoasparagines isolated from the urine of a patient with Gaucher's disease, could be interpreted in terms of complete primary structures of these

compounds. In fact, the proposed structures are mainly based upon interpretation of the sets of chemical shift values of H-1s and H-2s of constituting mannose residues. Such data are in fact sensitive to the type and configuration of the glycosidic linkage of the mannose residue and to the position of mannose in the chain.

This study provided us with the NMR characteristics of the second branching point (A–4′–B), also occurring in more complex oligomannoside-type structures; it thereby disclosed this type of carbohydrate chains for high-resolution ^{1}H NMR structural analysis.

For biosynthetic studies the non-destructive identification of partial structures, as shown in [15], is quite helpful. Now this seems also to be possible for oligomannoside-type carbohydrate structures.

Acknowledgements

This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), the Netherlands Foundation for Cancer Research (KWF, grant UUKC-OC 79-13), the Centre National de la Recherche Scientifique (Laboratoire Associé no. 217: Biologie physicochimique et moléculaire des glucides libres et conjugués) and the Institut National de la Santé et de la Recherche Médicale (contract no. 78.1.0523). J.-C. M. is Attaché de Recherche de l'INSERM; G. S. is Maître de Recherche de CNRS.

References

[1] Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Binette, J. P. and Schmid, K. (1978) Biochemistry 17, 5206–5214.
[2] Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J.-C., Fournet, B., Spik, G. and Montreuil, J. (1978) Eur. J. Biochem. 87, 323–329.
[3] Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Fiat, A.-M. and Jollès, P. (1980) Biochim. Biophys. Acta 623, 295–300.
[4] Vliegenthart, J. F. G., Van Halbeek, H. and Dorland, L. (1980) in: IUPAC 27th Int. Congr. Pure and Applied Chemistry, Helsinki 1979 (Varmavuori, A. ed) pp. 253–262, Pergamon, Oxford.
[5] Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Schmid, K., Montreuil, J., Fournet, B. and Hull, W. E. (1980) FEBS Lett. 114, 11–16.
[6] Van Halbeek, H., Dorland, L., Haverkamp, J., Veldink, G. A., Vliegenthart, J. F. G., Fournet, B., Ricart, G., Montreuil, J. and Aminoff, D. (1980) submitted.
[7] Strecker, G. and Montreuil, J. (1979) Biochimie 61, 1199–1246.
[8] Brunngraber, E. G. (1979) in: Glycoproteins and glycolipids in disease processes (Walborg, E. F. jr, ed) p. 143, Am. Chem. Soc., Washington.
[9] Zanetta, J. P., Breckenridge, W. C. and Vincendon, G. (1972) J. Chromatog. 69, 291–301.
[10] Ernst, R. R. (1966) in: Advances in Magnetic Resonance (Waugh, J. S. ed) vol. 2, pp. 1–135, Academic Press, London, New York.
[11] Strecker, G., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G. and Dubesset, D. (1978) Biochimie 60, 725–734.
[12] Herlant-Peers, M.-C., Montreuil, J., Strecker, G., Dorland, L., Van Halbeek, H., Veldink, G. A. and Vliegenthart, J. F. G. (1980) Eur. J. Biochem. in press.
[13] Dorland, L., Haverkamp, J., Schut, B. L., Vliegenthart, J. F. G., Spik, G., Strecker, G., Fournet, B. and Montreuil, J. (1977) FEBS Lett. 77, 15–20.
[14] Narasimhan, S., Harpaz, N., Longmore, G., Carver, J. P., Grey, A. A. and Schachter, H. (1980) J. Biol. Chem. 255, 4876–4884.
[15] Van den Eijnden, D. H., Joziasse, D. H., Dorland, L., Van Halbeek, H., Vliegenthart, J. F. G. and Schmid, K. (1980) Biochem. Biophys. Res. Commun. 92, 839–845.