Vol. 99, No. 3, 1981 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 15, 1981 Pages 886-892

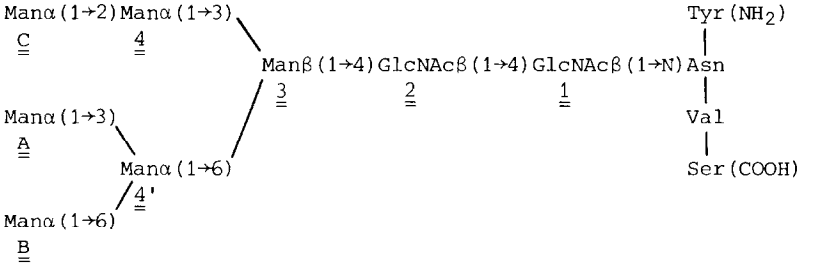
PRIMARY STRUCTURE OF THE ASPARAGINE-563-LINKED CARBOHYDRATE CHAIN OF AN IMMUNOGLOBULIN M FROM A PATIENT WITH WALDENSTROM'S MACROGLOBULINEMIA

A REINVESTIGATION BY 500-MHz ^{1}H-NMR SPECTROSCOPY

Herman VAN HALBEEK, Lambertus DORLAND and
Johannes F. G. VLIEGENTHART
Department of Bio-Organic Chemistry, University of Utrecht,
Croesestraat 79, NL-3522 AD Utrecht, The Netherlands
Jacqueline JOUANNEAU* and Roland BOURILLON
Laboratoire de Biochimie (U. 180 INSERM), Faculté de Médecine,
45 rue des Saints-Pères, 75006 Paris, France

Received January 27,1981

Summary. Earlier studies on the oligomannoside-type carbohydrate chain linked to Asn-563 of a human IgM protein from blood plasma of a patient (Du) with Waldenström's macroglobulinemia (Jouanneau, J. and Bourillon, R. (1979) *Biochem. Biophys. Res. Commun. 91*, 1057-1061) showed the occurrence of an unusual core structure. It was suggested that only one *N*-acetylglucosamine residue was present in stead of two. However, reinvestigation of this glycopeptide sample by 500-MHz ^{1}H-NMR spectroscopy lent no support for an unusual core structure. On the contrary, the NMR spectrum shows all spectral features characteristic for an *N,N'*-diacetylchitobiose unit. Moreover, it revealed a heterogeneity in both the carbohydrate and the peptide moiety. The structure of the major component is the following¶:



Besides, larger glycopeptides are present having one or two additional mannose residues α(1→2)-linked to terminal residues B and/or C in the above structure. It should be noted that for this sample, containing only 25 nmoles of glycopeptide, the microheterogeneity of the carbohydrate chain could excellently be revealed by 500-MHz ^{1}H-NMR spectroscopy.

---

* Present address: Institut Pasteur, Département de Biologie Moléculaire, 25 rue du Docteur-Roux, 75724 Paris Cedex 15, France.

¶ All sugars mentioned possess the D-configuration; all amino acids mentioned possess the L-configuration.

0006-291X/81/070886-07$01.00/0

Introduction. It is generally assumed that the carbohydrate moieties of glycoproteins account for several biological functions displayed by these compounds, *e.g.*, for recognition phenomena, for immunological events and for determining the life-span of cells and glycoproteins. The structure determination of the oligosaccharide chains of both soluble and membrane-bound glycoproteins is a prerequisite to gain insight into these biological events at the molecular level. Despite the functional diversity of glycoproteins, their carbohydrate components have many common structural features [1-3]. In particular, the oligosaccharide chains *N*-glycosidically linked to asparagine in a polypeptide backbone, appear to possess a virtually invariant core structure of the pentasaccharide, $Man_3GlcNAc_2$.

One of the exceptions hitherto reported [4], is the carbohydrate structure linked to Asn-563 of the heavy chain of a human IgM from blood plasma of a patient (Du) with Waldenström's macroglobulinemia. After isolation and purification [5], the glycopeptide was characterized by chemical and enzymic methods. On the basis of the carbohydrate compositions of the total glycopeptide and of its enzymic digestion products, it was proposed that an oligomannoside-type structure was concerned containing 6 mannoses but, surprisingly, only one *N*-acetylglucosamine residue in the core region. The question arose whether or not this deviating structure is correlated with the patient's disease, implicating another way of biosynthesis and/or processing of the carbohydrate chain than generally accepted to occur in man [6].

It seemed to be worthwhile to reinvestigate the reported core structure utilizing an independent method suitable for structural analysis on microscale. It was decided to apply high-resolution ^{1}H-NMR spectroscopy which was proved to be a powerful method for the structure determination of the carbohydrate chains of glycoproteins [7,8]. Recently, considerable progress has been made in the elucidation of oligomannoside-type structures by ^{1}H-NMR spectroscopy [8-13], making it reasonable to expect this method to give also structural details on the peripheral part of this IgM oligosaccharide chain.

Materials and methods. The isolation, purification and partial characterization of the IgM Asn-563 glycopeptide from blood plasma of a patient (Du) with Waldenström's macroglobulinemia has been described [4,5].

For NMR spectroscopy about 100 µg of the glycopeptide ([^{14}C]labelled by acetylation of the *N*-terminal amino acid with [^{14}C]acetic anhydride [4]) was repeatedly exchanged in D_2O (99.96 atom% D, Aldrich, Milwaukee, U.S.A.), with intermediate lyophilization. ^{1}H-NMR spectroscopy was performed at 500 MHz on a Bruker WM-500 spectrometer operating in the Fourier transform mode at a probe temperature of 300 K. Resolution enhancement of the spectrum was achieved by Lorentzian to Gaussian transformation from quadrature phase detection. Chemical shifts are given relative to sodium-2,2-dimethyl-2-silapentane-5-sulphonate (DSS) (indirectly to acetone in D_2O : δ = 2.225 ppm).

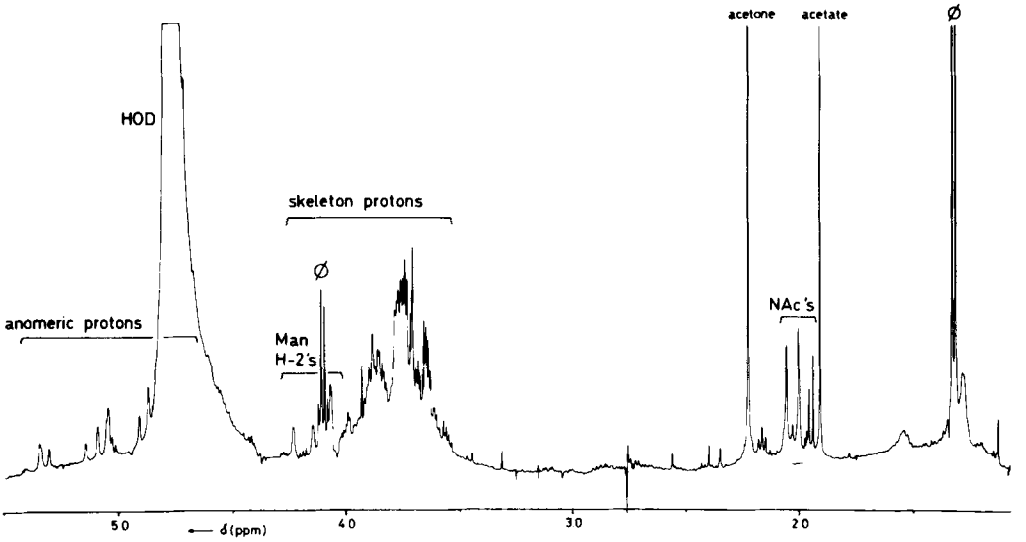


Fig. 1. Overall 500-MHz ^{1}H-NMR spectrum of the human IgM Asn-563 glycopeptide mixture from blood plasma of a patient (Du) with Waldenström's macroglobulinemia, in D_2O at 300 K. The signals marked by Φ stem from a non-protein non-carbohydrate contaminant.

Results and discussion. The overall 500-MHz ^{1}H-NMR spectrum of the human IgM Asn-563 glycopeptide, stemming from blood plasma of a patient (Du) with Waldenström's macroglobulinemia, is depicted in Fig. 1. From the acetyl-region of the spectrum ($1.9 < \delta < 2.1$ ppm) it is evident that four *N*-acetyl methyl singlets are present at $\delta \simeq 1.94$, 1.96, 2.00 and 2.06 ppm with intensity ratio 2:1:3:3, respectively. (Besides, a singlet of contaminating acetate is observed at δ = 1.910 ppm.) The signals at δ = 1.939 and 1.957 ppm are derived from the [^{14}C]-*N*-acetyl groups introduced by labelling of the *N*-terminal amino acids with [^{14}C]acetic anhydride. The remaining two signals belong to the GlcNAc residues $\underline{\underline{1}}$ and $\underline{\underline{2}}$ of the pentasaccharide core of *N*-glycosidically linked oligosaccharides. For GlcNAc-$\underline{\underline{1}}$ (δ = 2.001 ppm) this assignment is based on the spectral data of a wide variety of glycopeptides of both the *N*-acetyllactosamine and the oligomannoside type [12,14,15]. The chemical shift of the fourth *N*-acetyl signal (δ = 2.055 ppm) is typical for GlcNAc-$\underline{\underline{2}}$ in the core pentasaccharide extended with mannose residues only (*cf.* [12,16,17]). This assignment is corroborated by the chemical shift of H-1 of this GlcNAc residue (δ = 4.621 ppm) (*cf.* [12,14-17]). Therefore, the carbohydrate moiety of this IgM Asn-563 glycopeptide belongs to the oligomannoside-type glycan chains having the normal *N*,*N*'-diacetylchitobiose unit in the core region. It should be noted that recently for the carbohydrate chain located at Asn-402 of the heavy chain of the IgM of this patient (Du)

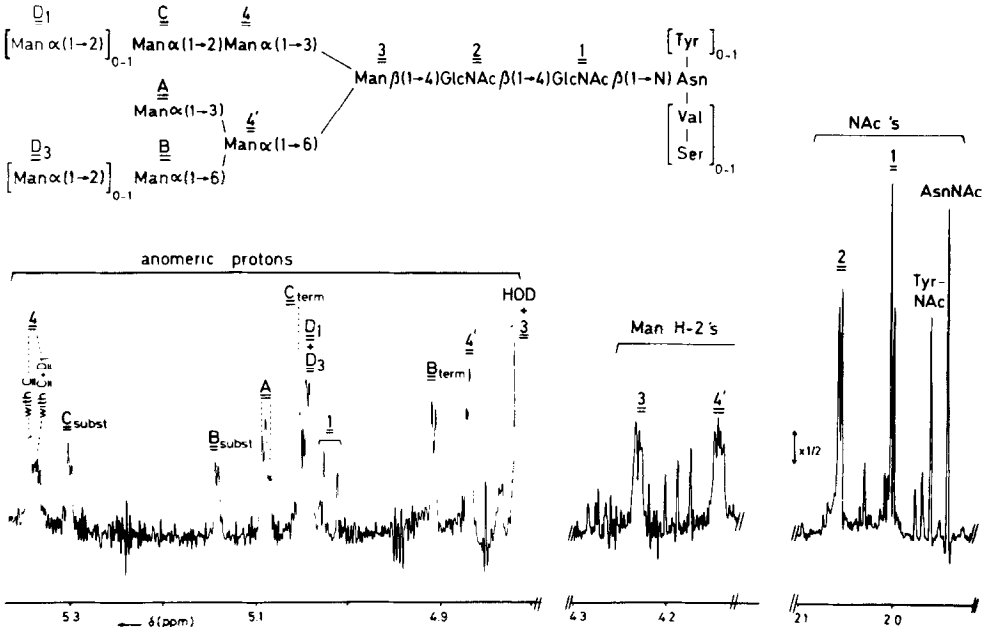


Fig. 2. Structural reporter group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectrum of the human IgM Asn-563 glycopeptide mixture in D_2O at 300 K, together with its comprehensive structure. The numbers and letters in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the *N*-acetyl proton region differs from that of the other parts of the spectrum as indicated. The *N*-acetyl signals of GlcNAc-1 and -2 both are split into two singlets, reflecting the earlier [4] described heterogeneity of the peptide moiety.

also the normal core pentasaccharide was found [18], utilizing essentially the same methods as in [4]. It is not yet understood why these afforded the incorrect molar ratio Asn : GlcNAc = 1:1 for the Asn-563-linked chain [4].

From the resolution-enhanced 500-MHz ^{1}H-NMR spectrum of the IgM glycopeptide the structure of the oligomannan part of the chain can be inferred on the basis of the resonances of the structural reporter groups, in particular the mannose H-1 signals, which are given in Fig. 2. Relevant NMR parameters for the glycopeptide sample are listed in table 1.

Besides the β-linked Man-3, clearly characterized by its H-2 signal at δ = 4.232 ppm, only α-linked mannose residues occur in the peripheral part of the glycan chain, on the basis of the chemical shifts of their H-1's in combination with their $J_{1,2}$'s [15] . The relatively high-field resonance position of H-1 of Man-4' (δ = 4.869 ppm) in conjunction with the relatively downfield position of its H-2 (δ = 4.145 ppm) is indicative of a disubstitution of this residue at

Table 1. ^{1}H chemical shifts of structural reporter groups of constituent monosaccharides of the human IgM Asn-563 glycopeptides obtained from blood plasma of a patient (Du) with Waldenström's disease

reporter group	residue[a]	compound and schematic structure[b] Man$_6$GP	Man$_7$GP	Man$_7$GP*	Man$_8$GP
H-1 of	1	≈5.02 [c]	≈5.02 [c]	≈5.02 [c]	≈5.02 [c]
	2	4.621	4.621	4.621	4.621
NAc of	1	2.001[d]	2.001[d]	2.001[d]	2.001[d]
		1.998	1.998	1.998	1.998
	2	2.054[d]	2.054[d]	2.054[d]	2.054[d]
		2.057	2.057	2.057	2.057
H-1 of	3	≈4.78 [e]	≈4.78 [e]	≈4.78 [e]	≈4.78 [e]
	4	5.345	5.336	5.345	5.336
	4'	4.869	4.869	4.869	4.869
	A	5.093[d]	5.093[d]	5.093[d]	5.093[d]
		5.088	5.088	5.088	5.088
	B	4.908	4.908	5.145	5.145
	C	5.052	5.304	5.052	5.304
	D_1	-	5.044	-	5.044
	D_3	-	-	5.044	5.044
H-2 of	3	4.232	4.232	4.232	4.232
	4'	4.145	4.145	4.145	4.145

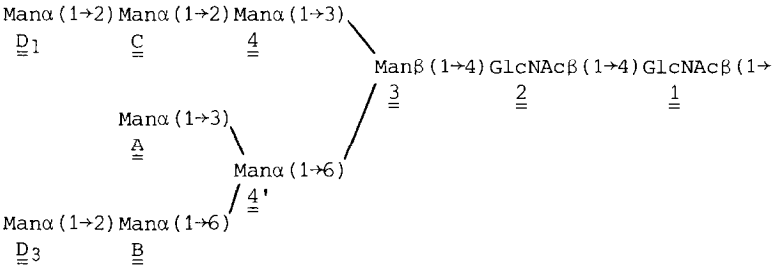
a for complete structures and coding of monosaccharide residues, see Fig. 2.
b ●— = mannose; O— = *N*-acetylglucosamine.
c value can not be determined more accurately (±0.01 ppm) due to heterogeneity of the peptide moiety (see legend Fig. 2).
d see text and legend Fig. 2.
e value can not be determined more accurately (±0.01 ppm) due to interference of the HOD-line.

C-3 and C-6 by Man-A and -B, respectively [8,11-13,16,17]. The H-1 of Man-A gives rise to two doublets at δ = 5.093 and 5.088 ppm, in a ratio 2:1, resonating in an area which is characteristic for a terminal non-reducing position of Man-A [11,12]. The doubling of the signal can be explained in terms of heterogeneity of the sample. Also H-1 of Man-B gives rise to two doublets, at δ = 4.908 (terminal B) and 5.145 ppm (C-2 substituted B) [11-13], in the same ratio 2:1. From these values it can be concluded that 33% of the Man-B residues in

the glycopeptide sample bear an α(1→2)-linked Man-$\underline{\underline{D}}_3$ (δ H-1 = 5.044 ppm), whereas in the remaining part Man-$\underline{\underline{B}}$ occupies a terminal position in the chain.

For H-1 of Man-$\underline{\underline{4}}$ in the upper branch of the chain two doublets are observed at δ = 5.345 and 5.336 ppm, in a ratio 1:1. The approximate chemical shift value (5.34 ppm) indicates that in all components of the glycopeptide mixture Man-$\underline{\underline{4}}$ bears the α(1→2)-linked Man-$\underline{\underline{C}}$ [13,16]. Also H-1 of Man-$\underline{\underline{C}}$ gives rise to two doublets, at δ = 5.304 (C-2 substituted $\underline{\underline{C}}$) and 5.052 ppm (terminal $\underline{\underline{C}}$) [11, 13,16]. These features can be explained by the presence of Man-$\underline{\underline{D}}_1$, α(1→2)-linked to Man-$\underline{\underline{C}}$, in 50% of the structures in the mixture. The presence of Man-$\underline{\underline{D}}_1$ (δ H-1 = 5.044 ppm) gives rise to a downfield shift of H-1 of Man-$\underline{\underline{C}}$ (Δδ = 0.252 ppm), and to a slight upfield shift of H-1 of Man-$\underline{\underline{4}}$ (Δδ ≃ - 0.01 ppm), compared to the $\underline{\underline{4}}$-$\underline{\underline{C}}$ branch without $\underline{\underline{D}}_1$ (*cf.* [13]).

Based on the above findings it can be concluded that the largest possible structure of the carbohydrate moiety of this IgM (Du) glycopeptide preparation is:



In the mixture of IgM glycopeptides also at least one, maybe two isomeric $Man_7GlcNAc_2$-oligosaccharides occur, possessing either Man-$\underline{\underline{D}}_1$ or Man-$\underline{\underline{D}}_3$. In addition, a $Man_6GlcNAc_2$-structure is present, without Man-$\underline{\underline{D}}_1$ and -$\underline{\underline{D}}_3$, which is the most abundant.

The same oligomannoside moieties as described above were reported for oligosaccharides derived from glycosylation site Asn-563 of an IgM from another patient (Ca) with Waldenström's macroglobulinemia, after laborious separation of the components of the prepared oligosaccharide mixture [11]. Furthermore, the structure of the predominant glycopeptide in the IgM sample from patient Du is identical with that reported for the Asn-563-linked $Man_6GlcNAc_2$-oligosaccharide, found in a human IgM myeloma protein [19] and in mouse IgM secreted by a plasmocytoma [20]. The location of additional mannose residues, *e.g.*, in $Man_8GlcNAc_2$-structures also occurring at the same glycosylation site in the latter two IgM's, was not specified. Moreover, these oligomannoside-type structures occur in a wide variety of glycoproteins such as calf thyroglobulin, ovalbumin, taka-amylase, bovine lactotransferrin, soybean agglutinin, ribonucleases and the glycoprotein from chinese hamster ovary cell membranes [1-3,11,16,17].

Acknowledgements. Thanks are due to Dr W.E. Hull (Bruker Analytische Messtechnik, GmbH, Rheinstetten, F.R.G.) for recording the 500-MHz ^{1}H-NMR spectrum.

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), and by the Netherlands Foundation for Cancer Research (KWF, grant UUKC-OC 79-13).

References.

[1] Montreuil, J. (1980) *Adv. Carbohydr. Chem. Biochem. 37*, 157-223.
[2] Kornfeld, R., and Kornfeld, S. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W.J. ed.) pp. 1-34, Plenum Press, New York.
[3] Sharon, N., and Lis, H. (1981) in *The Proteins*, Third Edition, Vol. V (Neurath, H., and Hill, R.L. eds.) in press, Academic Press, New York.
[4] Jouanneau, J., and Bourillon, R. (1979) *Biochem. Biophys. Res. Commun. 91*, 1057-1061.
[5] Jouanneau, J., Razafimahaleo, E., and Bourillon, R. (1970) *Eur. J. Biochem. 17*, 72-77.
[6] Schachter, H., and Roseman, S. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W.J. ed.) pp. 85-160, Plenum Press, New York.
[7] Vliegenthart, J.F.G., Van Halbeek, H., and Dorland, L. (1980) in *IUPAC 27th Int. Congr. Pure Appl. Chem.*, Helsinki 1979 (Varmavuori, A. ed.) pp. 253-262, Pergamon Press, Oxford.
[8] Vliegenthart, J.F.G., Van Halbeek, H., and Dorland, L. (1981) *Pure Appl. Chem. 53*, 45-77.
[9] Van Halbeek, H., Dorland, L., Vliegenthart, J.F.G., Strecker, G., and Montreuil, J. (1980) *Abstr. 13th FEBS Meeting*, Jerusalem 1980, 182.
[10] Narasimhan, S., Harpaz, N., Longmore, G., Carver, J.P., Grey, A.A., and Schachter, H. (1980) *J. Biol. Chem. 255*, 4876-4884.
[11] Cohen, R.E., and Ballou, C.E. (1980) *Biochemistry 19*, 4345-4358.
[12] Van Halbeek, H., Dorland, L., Veldink, G.A., Vliegenthart, J.F.G., Michalski, J.-C., Montreuil, J., Strecker, G., and Hull, W.E. (1980) *FEBS Lett. 121*, 65-70.
[13] Van Halbeek, H., Dorland, L., Veldink, G.A., Vliegenthart, J.F.G., Strecker, G., Michalski, J.-C., Montreuil, J., and Hull, W.E. (1980) *FEBS Lett. 121*, 71-77.
[14] 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.
[15] 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.
[16] Van Halbeek, H., Dorland, L., Vliegenthart, J.F.G., Spik, G., Chéron, A., and Montreuil, J. (1981) *Biochim. Biophys. Acta*, submitted for publication.
[17] Dorland, L., Van Halbeek, H., Vliegenthart, J.F.G., Lis, H., and Sharon, N. (1981) *J. Biol. Chem.* in press.
[18] Jouanneau, J., Fournet, B., and Bourillon, R. (1981) *Biochim. Biophys. Acta 667*, in press.
[19] Chapman, A., and Kornfeld, R. (1979) *J. Biol. Chem. 254*, 816-823.
[20] Brenckle, R., and Kornfeld, R. (1980) *Arch. Biochem. Biophys. 201*, 160-173.