*Carbohydrate Research*, 221 (1991) 169–177
Elsevier Science Publishers B.V., Amsterdam



# Characterisation of the N-linked oligosaccharides of the light chain of human glycoprotein IIb by f.a.b.-m.s.

Andrew J. Reason, Anne Dell, Howard R. Morris,
*Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ (Great Britain)*

Mark E. Rogers,
*M- Scan Inc, 137 Brandywine Parkway, West Chester, PA 19380 (U.S.A.)*

Juan J. Calvete, and José González-Rodríguez
*Instituto de Química Física, C.S.I.C., Serrano 119, Madrid 28006 (Spain)*

(Received April 10th, 1991; accepted for publication June 12th, 1991)

ABSTRACT

The glycosylation of the light chain (GPIIbL) of glycoprotein IIb, one of the glycoproteins constituting the receptor for fibrinogen, fibronectin, and the von Willebrand factor on platelet cell surfaces, was investigated using fast-atom-bombardment mass spectrometry (f.a.b.-m.s.). Complex-type N-glycans were observed, attached to Asn-60. The most abundant oligosaccharide is a disialylated biantennary structure substituted with fucose on the chitobiose core. Mono-sialylated biantennary, and di- and tri-sialylated triantennary structures were found as minor constituents of the N-glycan population. The amino acid sequence of GPIIbL was fully mapped by f.a.b.-m.s., thereby providing the first direct evidence for the absence of *O*-glycosylation.

## INTRODUCTION

Glycoproteins IIb (GPIIb) and IIIa (GPIIIa), the two major bitopic glycoproteins at the surface of blood platelets, form the $\alpha$ and $\beta$ subunits, respectively, of the $Ca^{2+}$-dependent non-covalent heterodimer GPIIb/IIIa, also known as integrin $\alpha_{IIb}\beta_3$, which serves as the receptor for fibrinogen and other adhesive proteins upon platelet activation, and plays a primary role in platelet adhesion and spreading on subendothelium, in platelet–platelet aggregation, and in clot retraction[1]. The GPIIb or $\alpha_{IIb}$ subunit (136 kDa) contains 16% of sugars and is made up of two chains joined by a single disulphide bond[2-4]. The heavy chain GPIIbH (114 kDa) contains >16% by weight of sugar, including *N*-acetylgalactosamine. It is fully extracellular and has four potential N-linked glycosylation sites as predicted from the cDNA-derived amino acid sequence[5]. The light chain GPIIbL (22–23 kDa) shows electrophoretic size-heterogeneity and contains 10% of carbohydrate with no *N*-acetylgalactosamine[3]. It has a single transmembrane segment, a short cytoplasmic tail, and a single extracellular *N*-glycosylation point at Asn-60, as predicted from the cDNA-derived amino acid sequence[5]. Biosynthetic studies in human megakaryocytes[6] and HEL cells[7] confirmed the observation that GPIIb is synthesised as a single chain precursor[8] which is cleaved proteolytically into the heavy and light chains in the golgi apparatus. Post-ribosomal

0008-6215/91/$ 03.50 

modifications of GPIIb have been investigated using protein chemical and mass-spectrometric techniques, and the previously observed electrophoretic size-heterogeneity of GPIIbL has been rationalised[9] as arising from dual proteolytic processing of GPIIb, giving rise to $GPIIbL_1$ and $GPIIbL_2$ (Fig. 1).

The glycosylation of GPIIb has been investigated using biosynthetic inhibitors[10], endoglycosidases[11], and sugar analyses[12]. These experiments have indicated that all putative glycosylation sites are occupied by complex-type oligosaccharides, but detailed structural information has not yet been reported. We are currently using mass-spectrometric strategies for investigating post-ribosomal modifications of GPIIbL and H, and we now report on our f.a.b.-m.s. studies of GPIIbL. The protein sequence deduced from cDNA studies, together with the previously reported N-terminal heterogeneity, were confirmed by f.a.b. mapping, and the absence of *O*-glycosylation was established. The major glycan attached to Asn-60 was shown to be a disialylated biantennary structure. Monosialylated biantennary and triantennary components are present in minor proportions.

## EXPERIMENTAL

*Materials.* — PNGase F was a gift from Dr. G. W. Hart (Johns Hopkins Medical School, Baltimore, U.S.A.). Sep-Pak cartridges were purchased from Waters Ltd. The GPIIbL glycoprotein was prepared from human platelet plasma membranes, then reduced and carboxymethylated as described[3].

*Chymotryptic digest.* — Reduced carboxymethylated GPIIbL (560 μg) was incubated with chymotrypsin (1:50 enzyme:substrate) in 50 mM ammonium hydrogencarbonate (125 μL, pH 8.4) for 5 h at 37°. The reaction was stopped by freeze-drying and the products were purified by h.p.l.c.

*Tryptic digest.* — The same conditions as the chymotryptic digest were used except that the reaction time was 1.5 h.

*PNGase F digest*[12]. — The tryptic digest (560 μg) was dissolved in 50mM ammonium hydrogencarbonate (100 μL, pH 8.3) and digested with 2.5 μL of PNGase F solution at 37° for 16 h. The reaction was terminated by freeze-drying. The products were dissolved in aqueous 5% acetic acid and loaded onto a C18 Sep-Pak cartridge primed with methanol, 1-propanol, and aqueous 5% acetic acid. The cartridge was eluted with aqueous 5% acetic acid (the void) followed by 10, 20, 30, and 40% 1-propanol in aqueous 5% acetic acid. Oligosaccharides were eluted in the void volume and the tryptic peptides in 10–30% 1-propanol.

*H.p.l.c. of the chymotryptic digest.* — The reverse-phase h.p.l.c. involved a Kontron system (450 data system, 420 pump, 430 detector, and M800 mixer) equipped with an Ultrasphere ODS column (4.6 × 250 mm) eluted at 1 mL/min. Solvent *A* was aqueous 0.1% trifluoroacetic acid, and solvent *B* was 0.1% of trifluoroacetic acid in aqueous 90% acetonitrile. The column was equilibrated in solvent *A*, then eluted with a 3-step gradient: 10→20% of solvent *B* in solvent *A* in 10 min, 20→50% in the next 30 min, and 50→100% in the next 20 min. Fractions were collected every 1 min, then

```
-12                -1  1                  10                  20
(  Q I F L P E P E Q P S R) L Q D P V L V S C D S A P C T V V Q C D

                   30                  40                  50
     L Q E M A R G Q R A M V T V L A F L W L P S L Y Q R P L D Q

                   60                  70                  80
     F V L Q S H A W F N V S S L P Y A V P P L S L P R G E A Q V

                   90                 100                 110
     W T Q L L R A L E E R A I P I W W V L V G V L G G L L L L T

                  120                 130          137
     I L V L A M W K V G F F K R N R P P L E E D D E E G E
```

Fig. 1. The cDNA-derived amino acid sequence of GPIIbL$_1$ (residues −12 to 137) and GPIIbL$_2$ (residues 1–137). The N-terminus of GPIIbL$_1$ is cyclised to pyroglutamic acid in the mature protein. The 12 amino acid N-terminal extension (bracketed) present in GPIIbL$_1$ is annotated with negative numbers in order to preserve the original numbering system for the light chain, in which Asn-60 is the site of *N*-glycosylation. N-linked glycosylation site (Asn-60).

concentrated, and each residue was dissolved in 10 $\mu$L of aqueous 5% acetic acid for f.a.b.-m.s.

*Methylation.* — The void fraction from the Sep-Pak elution was transferred to a screw-capped glass tube and methylated by the addition of a NaOH–$Me_2SO$ slurry (1 mL) made by grinding 8 pellets of NaOH in a dried mortar with dry distilled $Me_2SO$ (1 mL). This was followed by the addition of methyl iodide (1 mL), and the mixture was shaken at room temperature for 10 min before quenching slowly with water (~2 mL). The permethylated oligosaccharides were extracted into chloroform (1 mL), the extract was washed with water (5 × 1 mL), and the solvent was evaporated in a stream of nitrogen.

*F.a.b.-m.s.* — A VG Analytical High-Field ZAB-HF mass spectrometer was used for peptide mapping and for the permethylated oligosaccharides, and a High-Field ZAB-2SE instrument was used for the glycopeptides and the high-mass regions of permethylated oligosaccharides. The ZAB-HF isntrument was fitted with an M-Scan f.a.b. gun which was operated at 10 kV. The ZAB-2SE instrument was fitted with a cesium ion gun operated at 30 kV. The matrix was glycerol–thioglycerol (1:1). Samples were dissolved in aqueous 5% acetic acid (peptides and glycopeptides) or methanol (permethylated oligosaccharides) prior to loading into the matrix. Spectra acquired on the ZAB-2SE instrument were computer-processed and the assignments correspond to average chemical masses, whilst spectra acquired on the ZAB-HF instrument were recorded on oscillographic chart paper and the nominal masses were assigned by counting.

## RESULTS

*F.a.b.-m.s. analyses of chymotryptic and tryptic peptides.* — Reduced carboxymethylated GPIIbL was subjected to chymotryptic or tryptic digestion and the digests were purified by reverse-phase h.p.l.c. (chymotryptic digest) or on a Sep-Pak cartridge (tryptic digest). Each h.p.l.c. or Sep-Pak fraction was screened by f.a.b.-m.s. and the

TABLE I

Quasimolecular ions[a] $(M + H]^+$ observed in the f.a.b. spectra of chymotryptic and tryptic digests of reduced carboxymethylated GPIIbL

| *Chymotryptic digest* | | *Tryptic digest* | |
|---|---|---|---|
| m/z | *Assignment* | m/z | *Assignment* |
| 474 | 82–85 | 597 | 119–123 |
| 684 | 1–6 | 617 | 87–91 |
| 767 | 116–121 | 851 | 30–37 |
| 778 | 39–44 | 949 | 67–75 |
| 789 | 43–48 | 1150 | 30–39 |
| 826 | 98–106 | 1300 | 76–86 |
| 891 | 38–44 | 1423 | (−12)–(−1) |
| 903 | 45–51 | 1528 | 125–137 |
| 987 | 52–59 | 2348 | 27–46 |
| 1142 | 72–81 | 2819 | 1–26 |
| 1173 | 107–117 | | |
| 1276 | 111–121 | | |
| 1536 | (−12)–1 | | |
| 1619 | 67–81 | | |
| 1812 | 123–137 | | |
| 1994 | 82–97 | | |
| 3131 | 60–66 (glycopeptide) | | |

[a] These were mapped onto the cDNA-deduced protein sequence and the residue numbers are given for each peptide corresponding to the quasimolecular ions.

pseudomolecular ions observed were mapped onto the predicted protein sequence (Fig. 1), using a computer program developed in this laboratory[13]. The masses of the pseudomolecular ions and their assignments are given in Table I. The chymotryptic glycopeptide was observed in fraction 11, and the high-mass region of the f.a.b. spectrum of this fraction is reproduced in Fig. 2. The major cluster of pseudomolecular ions at $m/z$ 3131 $[M + H]^+$, 3153 $[M + Na]^+$, and 3175 $[M - H + 2Na]^+$ corresponds to a glycan of composition $Fuc_1NeuAc_2Hex_5HexNAc_4$ attached to the $Asn_{60}-Tyr_{66}$ peptide. The cluster centred at $m/z$ 2862 is separated from the major molecular ions by a mass interval corresponding to a sialic acid residue.

*F.a.b.-m.s. analyses of permethylated oligosaccharides.* — The N-linked oligosaccharides were liberated from the tryptic digest by digestion with PNGase F and separated from the tryptic peptides by Sep-Pak purification. The oligosaccharides were permethylated and subjected to f.a.b.-m.s. Spectra below $m/z$ 3000 were recorded with the ZAB-HF mass spectrometer (Fig. 3a), and higher mass molecular ions were detected using the ZAB-2SE instrument (Fig. 3b). It should be noted that the mass increment in the region common to Figs. 3a and 3b is due to the former being a nominal (H = 1, C = 12, N = 14, O = 16) mass and the latter being a computer-assigned accurate chemical mass. Each number is equally valid, and assignment depends upon inputting the nominal or accurate mass value of the saccharide, as appropriate. In Fig. 3b, the $[M + H]^+$ quasimolecular ions at $m/z$ 2946 (major), 3395, and 3757 correspond to composi-

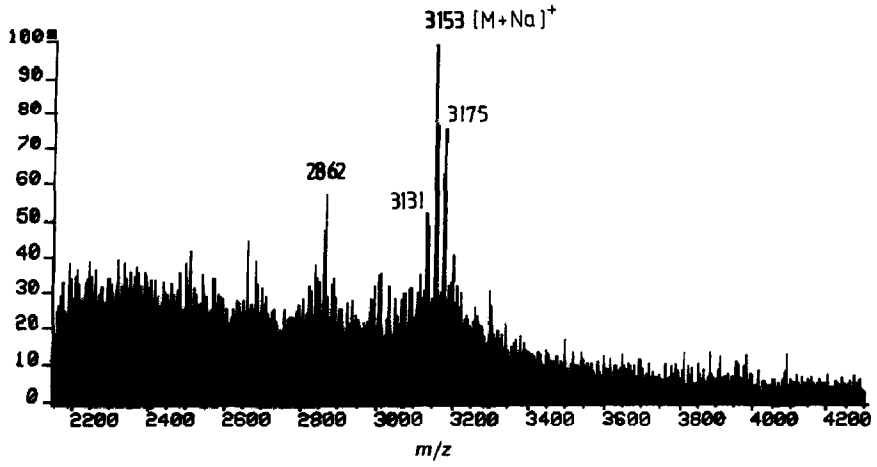
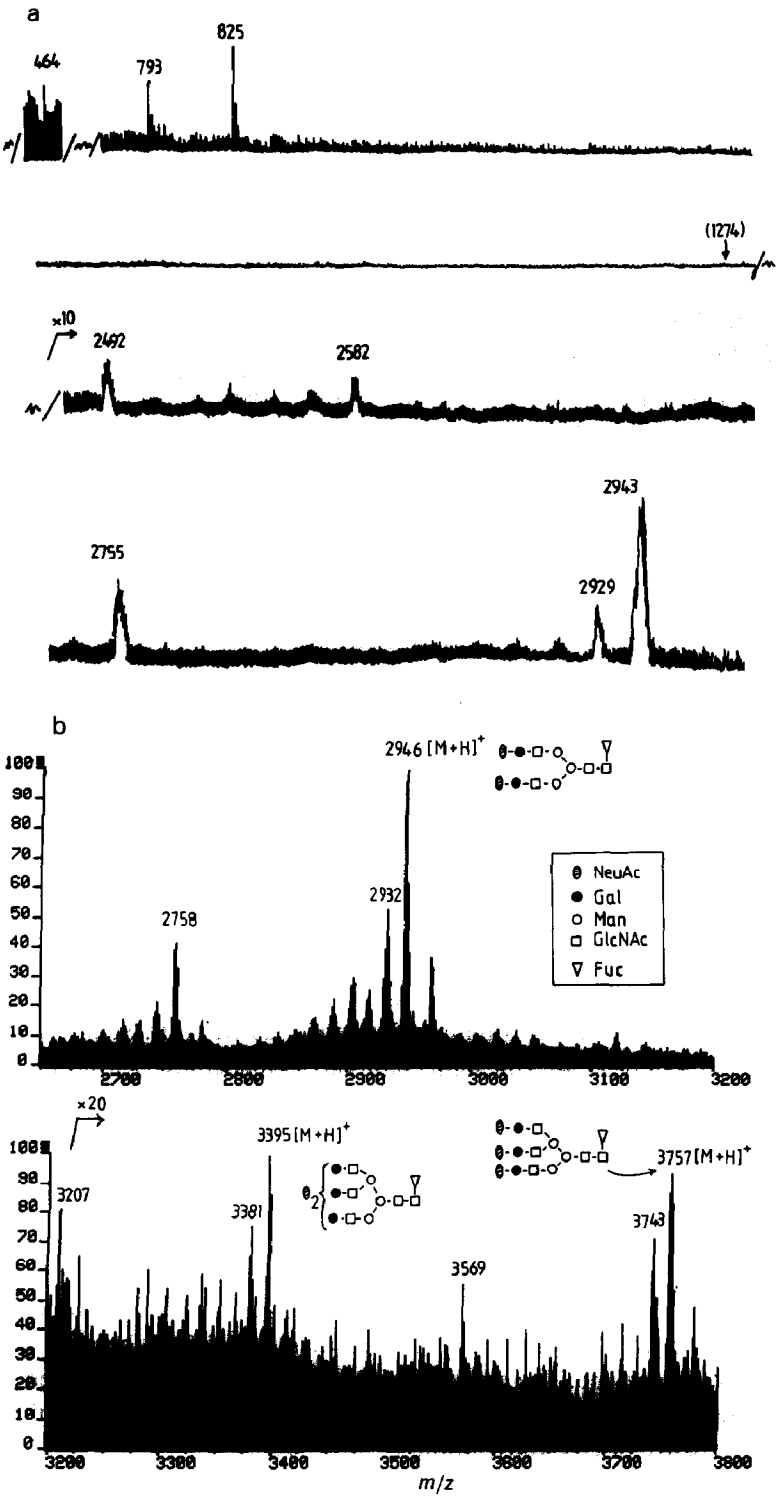



Fig. 2. Molecular ion region of the positive f.a.b.-mass spectrum of the chymotryptic glycopeptide from GPIIbL. The data were acquired on the ZAB 2SE instrument, and the spectrum is computer-calibrated.

tions $Fuc_1NeuAc_2Hex_5HexNAc_4$, $Fuc_1NeuAc_2Hex_6HexNAc_5$, and $Fuc_1NeuAc_3Hex_6HexNAc_5$, respectively. Each ion is accompanied by a signal 188 m.u. to lower mass (*m/z* 2758, 3207, and 3569). These are fragment ions derived by loss of fucose *via* a *β*-cleavage reaction[14]. Fragment ions formed by A-type cleavages[14] at amino sugar residues are present in Fig. 3a at *m/z* 464 (very minor; $HexHexNAc^+$), 825 ($NeuAcHexHexNAc^+$), and 2492 ($NeuAc_2Hex_5HexNAc_3^+$). The minor signal at *m/z* 2852 in Fig. 3a is most probably the quasimolecular ion for the monosialylated counterpart of *m/z* 2946 in Fig. 3b.

## DISCUSSION

The light chain of GPIIb is heterogeneous at the N-terminus due to dual proteolysis of the precursor (Fig. 1). The smaller constituent ($GPIIbL_2$), which comprises the last 137 residues of the GPIIb precursor, was originally identified[7] as GPIIbL. Subsequent work showed that approximately 50% of the light chains are extended by 12 residues at the N-terminus and are blocked by cyclisation of the N-terminal glutamine[9]. In order to preserve the original numbering system for GPIIbL, in which the glycosylated residue occurs at position 60, we have given negative annotations to the N-terminal extension (Fig. 1). The sequence of GPIIbL, including the blocked N-terminal extension, was fully mapped by our f.a.b.-m.s. studies of the tryptic and chymotryptic digests (Table I). The existence of $GPIIbL_2$ in our preparation of GPIIbL was confirmed by the observation of chymotryptic peptides commencing at position 1. These peptides cannot be produced by chymotryptic digestion of the N-terminally extended form of GPIIbL because the R–L bond is resistant to this enzyme. The data in Table I validate the cDNA sequence assignment and the reported N-terminal heterogeneity of GPIIbL, and provide the first direct evidence for the absence of *O*-glycosylation.


a
825
464
793
(1274)
×10
2492
2582
2943
2755
2929
b
2946 [M+H]+
2758
2932
NeuAc
Gal
Man
GlcNAc
Fuc
×20
3395 [M+H]+
3757 [M+H]+
3207
3381
3743
3569
m/z

The N-glycan structures present in GPIIbL were probed by f.a.b.-m.s. of the underivatised chymotryptic glycopeptide. Cleavage occurred on the N-terminal side of the glycosylation site to give the $Asn_{60}$–$Tyr_{66}$ glycopeptide, which afforded abundant quasimolecular ions (Fig. 2) whose compositions are consistent with a disialylated biantennary oligosaccharide substituted with one fucosyl residue. Additional signals in the molecular ion region occur at the calculated mass for its monosialylated counterpart. These ions could be formed by loss of sialic acid during the f.a.b. experiment or they could be pseudomolecular ions of mono-sialylated components present in the sample. In order to resolve this ambiguity, f.a.b.-m.s. analyses were carried out on the N-glycans after their release from the protein and subsequent permethylation. Molecular ions and fragment ions formed by $\beta$-cleavages are readily differentiated in spectra of permethylated samples because the latter carry free hydroxyl groups at the sites of $\beta$-cleavage. Preparation of the permethylated derivative also improved the sensitivity, and this allowed the detection of minor components that were not observed in the spectra of the glycopeptide. The characteristic A-type fragmentation that occurs at each amino sugar linkage in permethylated samples is an additional advantage of this derivative because of the sequence information that can be deduced from the compositions of the fragment ions.

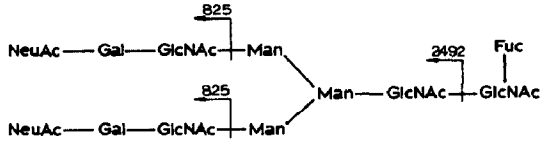
The N-glycans were liberated from GPIIbL by use of peptide *N*-glycosidase F (PNGase F), which hydrolyses the GlcNAc–Asn bond of N-glycans[12]. In order to ensure that the glycosylation site was accessible to the enzyme, the glycoprotein was first proteolytically digested. Detergents and such agents as guanidinium hydrochloride are frequently used to denature proteins prior to PNGase F digestion, but these are not compatible with high-sensitivity f.a.b.-m.s.. We chose to employ trypsin, rather than chymotrypsin, prior to PNGase digestion, because our f.a.b.-m.s. experiments had revealed that the glycopeptide produced in the latter digest has the glycosylation site at the N-terminus. Other workers have shown that glycans attached to N-terminal Asn are resistant to PNGase F digestion[12].

The data afforded by the permethylated derivative (Fig. 3) confirmed the presence of a disialylated structure as the major N-glycan in GPIIbL. This gives the quasimolecular ion signals at *m/z* 2943 and 2946 in Figs. 3a and 3b, respectively. The apparent shift in mass in Fig. 3b is a consequence of computer, rather than manual, mass assignment (see Experimental). The low-mass end of the spectrum in Fig. 3a (below *m/z* 1300) contains the A-tpe fragment ions that define the non-reducing moieties present in the complex oligosaccharides. One prominent A-type ion is present, namely $NeuAcHexHexNAc^+$ at *m/z* 825. The very minor signal at *m/z* 464 corresponds

Fig. 3. (*a*) Positive f.a.b.-mass spectrum of permethylated oligosaccharides from GPIIbL acquired using the ZAB HF instrument. The spectrum was counted manually, giving nominal mass assignments. Signals are assigned in the text, with the exception of *m/z* 793 which corresponds to loss of methanol from *m/z* 825. (*b*) High-mass region of the f.a.b.-mass spectrum of the sample giving the data in (*a*). The spectrum was acquired using the ZAB 2SE instrument, and the chemical masses are computer-assigned. Signals are assigned in the text, except for undermethylated components which afford the signals 14 a.m.u. below each quasimolecular ion.

to $HexHexNAc^+$. The low abundance of the latter ion is consistent with complete sialylation of the major component. The absence of signals at *m/z* 638 ($FucHexHexNAc^+$) and 999 ($NeuAcFucHexHexNAc^+$) is notable, indicating that the fucosyl residue contributing to the molecular ion composition is unlikely to be located in the antennae and is most probably on the chitobiose core. This view is corroborated by the chitobiose cleavage ion at *m/z* 2492 (Fig. 3a), which is assigned the composition $NeuAc_2Hex_5HexNAc_3^+$.

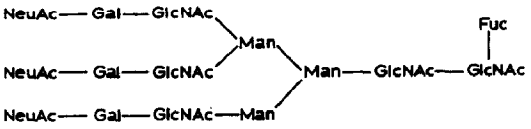
The major ions observed in the f.a.b. spectra of the permethylated derivative are fully consistent with the disialylated biantennary structure (**1**), which fragments as shown at each GlcNAc residue.




1

The monosialylated counterpart of **1** is also present, but it is only a minor component as judged by the abundance of both the A-type fragment ion at *m/z* 464 and the quasimolecular ion at *m/z* 2582. It is therefore likely that the relatively abundant signal corresponding to the monosialylated glycopeptide observed in Fig. 2 is largely derived from loss of sialic acid during the f.a.b. experiment. This outcome is not surprising, because previous studies have shown that glycosidic linkages involving sialic acids are readily cleaved in underivatised glycopeptides[15].

The composition assigned to the quasimolecular ion at *m/z* 3757 (Fig. 3b) is consistent with the trisialylated triantennary structure **2**. Its disialylated counterpart gives the quasimolecular ion at *m/z* 3395. An alternative explanation for the ion at *m/z* 3395, namely a biantennary structure containing NeuAc–Gal–GlcNAc–Gal–GlcNAc in one antenna, is considered to be less likely because of the absence of an A-type ion at *m/z* 1274 for $NeuAcHex_2HexNAc_2^+$ (Fig. 3a). The triantennary components are relatively minor (note the 20-fold increase in gain above *m/z* 3000 in Fig. 3b) and did not give convincing data on the ZAB-HF instrument because of the drop in sensitivity consequent upon the lowered accelerating voltage required to focus ions above *m/z* 3300 on this instrument.




2

In conclusion, our f.a.b. studies show that Asn-60 of GPIIbL is substituted with a family of complex-type structures, one of which, the disialylated biantennary structure **1**, is very much more abundant than its triantennary relative **2**. Most of the components

are fully sialylated, although minor amounts of structures containing one non-sialylated antenna are also present. There is no evidence for tetra-antennary structures or for oligosaccharides containing antennae with fucose substitution or *N*-acetyl-lactosamine repeats. Our previous work has shown that these structures are readily detectable at the 3–5% level in mixtures of complex oligosaccharides[16]. Therefore, the structures described above are likely to constitute more than 95% of the N-glycan content of GPIIbL.

ACKNOWLEDGMENTS

We thank the Medical Research Council for Programme Grant Support. J.J.C. and J.G.-R. are supported by the S.E.U.I. (PB-86-0629 and PM-88-0022), and A.J.R. is a recipient of an S.E.R.C. studentship.

REFERENCES

1 N. Kieffer, and D. R. Phillips, *Annu. Rev. Cell Biol.*, 6 (1990) 329–358.
2 M. T. Eirin, J. J. Calvete, and J. Gonzalez-Rodriguez, *Biochem. J.*, 240 (1986) 147–153.
3 J. J. Calvete and J. Gonzalez-Rodriguez, *Biochem. J.*, 240 (1986) 155–161.
4 P. Usobiaga, J. J. Calvete, J. L. Saiz, M. T. Eirin, and J. Gonzalez-Rodriguez, *Eur. Biophys. J.*, 14 (1987) 211–218.
5 M. Poncz, P. Eisman, R. Heidenreich, S. Silver, G. Vilane, S. Surrey, E. Schwartz, and J. S. Bennett, *J. Biol. Chem.*, 262 (1987) 8476–8482.
6 A. Duperray, R. Berthier, E. Chagnon, J.-J. Ryckewaert, M. H. Ginsberg, E. F. Plow, and G. Marguerie, *J. Cell Biol.*, 104 (1987) 1665–1673.
7 J. P. Rosa and R. P. McEver, *J. Biol. Chem.*, 264 (1989) 12596–12603.
8 P. F. Bray, J. P. Rosa, V. R. Lingappa, Y. W. Kan, R. P. McEver, and M. A. Shuman, *Proc. Natl. Acad. Sci. U.S.A.*, 83 (1986) 1480–1484.
9 J. J. Calvete, W. Schafer, A. Henschen, and J. Gonzalez-Rodriguez, *FEBS Lett.*, 272 (1990) 37–40.
10 A. Troesch, A. Dupeway, B. Polack, and G. Marguerie, *Biochem. J.*, 268 (1990) 129–133.
11 J. J. Calvete, A. Henschen, and J. Gonzalez-Rodriguez, *Biochem. J.*, 261 (1989) 561–568.
12 A. C. Tarentino, C. M. Ganez, and T. H. Plummer, Jr., *Biochemistry*, 24 (1985) 4665–4671.
13 R. A. McDowell, unpublished work.
14 A. Dell, *Adv. Carbohydr. Chem. Biochem.*, 45 (1987) 19–72.
15 H. Sasaki, N. Ochi, A. Dell, and M. Fukuda, *Biochemistry*, 27 (1988) 8618–8626.
16 A. L. Chan, H. R. Morris, M. Panico, A. T. Etienne, M. E. Rogers, P. Gaffney, L. Creighton-Kempsford, and A. Dell, *Glycobiol.*, 1 (1991) 173–185.