Biochem. J. (1992) 283, 105–112 (Printed in Great Britain)

Analysis of the five glycosylation sites of human α_1-acid glycoprotein

Michael J. TREUHEIT,* Catherine E. COSTELLO† and H. Brian HALSALL*‡

*Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221-0172, U.S.A., and †Department of Chemistry, Massachusetts Institute of Technology Mass Spectrometry Facility, Cambridge, MA 02139, U.S.A.

Orosomucoid (OMD) contains complex bi-, tri- and tetra-antennary glycan chains. Subfractionation of OMD into three molecular variants using concanavalin A lectin chromatography is based on variations in these complex structures. Standard h.p.l.c. profiles have been developed to analyse the percentage and distribution of the glycoforms present at each glycosylation site in OMD and its molecular variants. The ability to quantify the glycoforms present at each site allows us to extend the earlier results of others and resolve the remaining questions concerning the glycan structures of these variants. Most significantly, the proportions of bi-, tri- and tetra-antennary chains differ at each site for the three molecular variants. The most strongly retained variant from concanavalin A is uniquely capable of possessing biantennary chains at all five sites, whereas the unretained variant is completely devoid of biantennary chains. Only glycosylation site II of the five present is 100 % biantennary in the retained and weakly retained variants. In addition, the two gene products of OMD were differentially glycosylated. Molecular masses of the glycoforms were verified by matrix-assisted u.v. laser desorption mass spectrometry. On the basis of the site distribution of oligosaccharides in the variants, efforts were made to understand the factors that control the processing of the carbohydrate chains in OMD. The results indicate that the 'site-directed' model of processing offers the most consistent explanation for the structures seen at the individual glycosylation sites of OMD.

INTRODUCTION

Orosomucoid (OMD; α_1-acid glycoprotein) is one of the best characterized serum glycoproteins. It is generally considered to be a member of the lipocalin family [1,2] which function either in the transport of lipophilic compounds or as receptors for them [2]. Although a few reports exist that directly demonstrate that drugs or other compounds are bound to OMD in serum [3,4], it should as yet be considered as belonging to the group of plasma proteins the physiological functions of which remain unknown.

OMD is classified as one of the positive acute phase reactants, and the plasma concentrations of this protein rise by as much as 4-fold in some disease states [5–7]. OMD contains five *N*-asparaginyl-linked complex glycan chains making it one of the most heavily glycosylated of the serum glycoproteins [8]. The glycan structures have been extensively studied [9–11], and it is clear that in OMD an individual polypeptide carries several different types of glycan structure with considerable heterogeneity at the same glycosylation site. Since some of the proposed functions for OMD may be due to the nature of the glycan chains on this protein [12–15], it would appear that the extent of processing of the glycans can serve to create discrete glycoforms possibly leading to functional diversity. The role of the *N*-linked oligosaccharides in mediating protein-specific biological activities in OMD, as well as other glycoproteins, remains an intense research question.

OMD can be subfractionated into three fractions on the basis of variations in the structures of the glycan chains by using concanavalin A (Con A) affinity chromatography. The fractions are referred to as Con A unretained (UR), Con A weakly retained (WR) and Con A retained (R), and populate normal plasma at 46.0 %, 39.1 % and 14.9 % respectively [16,17]. In early work, Bayard & Kerckaert [16] suggested that within a fraction each molecule contains only a single glycan form, but more recently Bierhuizen *et al.* [17] have concluded that the R fraction contains the biantennary form only at sites I and II, in agreement with earlier n.m.r. studies [9,10]. They further suggest that site II is the biantennary site common to the WR and R fractions that is involved in the interaction with Con A. However, since the calculated number of biantennary chains in the R fraction was higher than two, these authors could not exclude the occurrence of OMD forms containing more than two biantennary chains, leaving a lack of concensus about the types of structure present on the individual forms of the OMD separable by Con A [9,10,16,17].

During an acute phase, it is observed that the relative proportions of the Con A variants change [18–22], a phenomenon that may be related to the physiological function of OMD [12–15]. Little, however, is known about the alterations that occur at the glycosylation sites which are responsible for the concentration shifts.

The work described here details the glycan structures present at the individual glycosylation sites of the molecular variants of OMD, and permits interpretation of the extent of processing at individual sites. This information appears to resolve the remaining questions about the glycan structures of the variants obtained by using Con A, and may be useful in drawing conclusions about which factors control branching in complex glycan biosynthesis in the molecular variants of OMD [23–29].

MATERIALS AND METHODS

Chemicals

All reagent-grade chemicals and h.p.l.c.-grade solvents were

Abbreviations used: OMD, orosomucoid; Con A, concanavalin A; UR, orosomucoid fraction not retained by concanavalin A; WR, orosomucoid fraction weakly retained by concanavalin A; R, orosomucoid fraction retained by concanavalin A; m.a.l.d.-t.o.f./m.s., matrix-assisted laser-desorption time-of-flight m.s.; f.a.b./m.s., fast-atom-bombardment m.s.; GP I, GP II, gene products I and II.

‡ To whom correspondence should be addressed.

purchased from Fisher Scientific, Pittsburgh, PA, U.S.A. H.p.l.c.-grade water was purified in our laboratory.

Purification of OMD from pooled plasma

Plasma was obtained from University Hospital, Hoxworth Division, Cincinnati, OH, U.S.A. Initial protein precipitations were produced by stirring poly(ethylene glycol) (3350; Sigma Chemical Co., St. Louis, MO, U.S.A.) into 3 litres of plasma to 23% weight of poly(ethylene glycol)/volume of plasma, stirring continuously at 4 °C overnight. The precipitated proteins were centrifuged at 1500 **g** for 10 min and the supernatant was collected. The remaining steps in the isolation were as outlined by Halsall *et al.* [30].

Con A affinity chromatography

Con A–Sepharose 4B (10–15 mg of lectin/ml of gel; Sigma) was equilibrated with 0.1 M-Tris/HCl containing 1.0 M-NaCl, 5 mM-$MgCl_2$, 5 mM-$CaCl_2$ and 5 mM-$MnCl_2$, pH 7.4. The purified pooled plasma OMD was chromatographed using this buffer as reported [16,17], yielding the U and WR fractions. The bound peak (R) was eluted using the equilibration buffer containing 0.15 M-methyl α-mannoside. The results of SDS/PAGE of these purified Con A fractions were indistinguishable from those obtained by others [17] (results not shown).

Reduction and alkylation

The procedure was essentially that of Ikenaka *et al.* [31]. The reaction was stopped by dialysis against water.

Endoproteinase Glu-C cleavage

Endoproteinase Glu-C (V-8 proteinase; Boehringer-Mannheim Biochemicals, Indianapolis, IN, U.S.A.) was used at an enzyme to reduced and alkylated protein concentration ratio of 1:20 (0.25 mg of enzyme and 5 mg of protein/400 μl of buffer). The buffer was 0.1 M-phosphate, pH 7.8, and treatment was for 24 h at room temperature.

Portions (150 μl) were subjected to h.p.l.c. and the individual peaks collected. The fractions were separated initially on a Vydac (218-TP-5415 C_{18}) column (The Separations Group, Hesperia, CA, U.S.A.). The column eluate was monitored at 276 nm. The gradient used was 5 min 100% h.p.l.c.-grade water containing 0.1% trifluoroacetic acid, 45 min to 70% acetonitrile containing 0.1% trifluoroacetic acid, with an absorbance unit full scale (AUFS) of 0.25 at a flow rate of 1 ml/min.

Peaks containing hexose were determined using the phenol/H_2SO_4 test [32]. Samples were taken at 3, 7, 10, 15 and 24 h to measure the extent of enzymic cleavage as a function of the number of peaks containing sugar. Cleavage was essentially complete after 10 h but 24 h digests were routinely done for convenience.

Peptide analysis

Dansyl chloride labelling followed the procedure of Winter [33] with standards from Sigma, and amino acid analyses were performed by the University of Cincinnati Protein Chemistry Core Facility.

H.p.l.c.

Reverse-phase h.p.l.c. gradients were generated by dual Kratos Spectroflow 400 pumps controlled by a Kratos model 783 (Ramsey, NJ, U.S.A.). Chromatograms were recorded using a 7290 SpectraPhysics Integrator (San Jose, CA, U.S.A.). After the initial separation of the endoproteinase Glu-C cleavage products, further purification of the glycopeptides positive to phenol/H_2SO_4 used detection at 214 nm and a Vydac (218-TP-510 C_{18}) column with various gradients of acetonitrile with 0.1% trifluoroacetic acid [34]. This column was brought to the starting gradient composition by using a reverse gradient, and was allowed to equilibrate at that composition for 45 min.

Desialylation

The non-specific sialidase from *Clostridium perfringens*, insolubilized on beaded agarose (approx. 25 units/g of agarose; Sigma) was used to desialylate the glycopeptides. Approx. 200 μg of the isolated glycopeptide was incubated with 0.25 unit of the enzyme in 400 μl of 0.1 M-acetate, pH 5.0, for 24 h at 37 °C. The glycopeptides were separated from the beads by centrifugation at 300 **g** for 5 min in Spin-X centrifuge filter units (Costar, Cambridge, MA, U.S.A.). The desialylated glycopeptides were concentrated by elution from C_{18} SepPaks (Waters Chromatography Division, Milford, MA, U.S.A.) using acetonitrile/water (1:1, v/v), and dried under N_2. The dried glycopeptides were tested for sialic acid using the Warren [35] assay, or rerun on the C_{18} column (218-TP-510) to determine changes in the h.p.l.c. profiles.

Warren assay

The sialic acid content of a glycopeptide was determined using the thiobarbituric acid assay as developed by Warren [35] and modified by O'Kennedy [36]. All reagents and standards were heated to 70 °C before use, and 200 μl samples of glycopeptide solution were used.

Treatment with α-L-fucosidase

The procedure of Abe *et al.* [37] was followed, but using a 48 h digest with fucosidase from Boehringer-Mannheim Biochemicals. The digest was loaded on to the Vydac (218-TP-510) column and eluted.

Mild acid treatment

All glycopeptides were treated with 150 μl of 0.05 M-H_2SO_4 for 30 min at 80 °C, cooled, and rerun on the Vydac (218-TP-510) column to determine the changes in the h.p.l.c. profile. Additional carbohydrate cleavage could be detected with a 45 min hydrolysis.

Electrophoresis

SDS/PAGE was performed under reducing conditions according to Laemmli [38] using an 11% slab gel (Hoefer Scientific Instruments, San Francisco, CA, U.S.A.). SDS/PAGE was used to assess purity of the OMD preparations and the migration of the Con A variants. Permanent gel records were obtained by using Gelbond (FMC Bioproducts, Rockland, ME, U.S.A.).

Mass spectrometry

The desialylated glycoforms within an individual glycosylation site were isolated by conservative peak cutting, and identified by matrix-assisted laser-desorption time-of-flight mass spectrometry (m.a.l.d.-t.o.f./m.s.) or fast-atom-bombardment mass spectrometry (f.a.b./m.s.) (Table 2) at the Mass Spectrometry Facility at the Massachusetts Institute of Technology. The glycopeptides were dissolved in water/acetonitrile (3:1, v/v) and a portion of this solution was mixed with the matrix. Sinapinic acid (3,5-dimethoxy-4-hydroxy-*trans*-cinnamic acid) [39] was used as the matrix for m.a.l.d.-t.o.f./m.s. measurements, which were carried out using frequency-tripled 355 nm light from a Nd:YAG laser on the Vestec 2000 time-of-flight mass spectrometer, and

accelerating voltage +20 to +30 kV. The f.a.b. matrix was a 1:1 mixture of glycerol/thioglycerol. F.a.b. mass spectra were acquired with the first two sectors (m.s.-1) of a JEOL HX110/HX110 tandem mass spectrometer, using 25 kV Cs^+ ions as the primary beam and +10 kV accelerating voltage.

RESULTS

V-8 proteinase gave an extremely reproducible enzymic cleavage pattern, from which h.p.l.c. profiles of the individual glycosylation sites were obtained. This proteinase had very little effect on native OMD and therefore all V-8 proteinase cleavages were done on the reduced and alkylated form. Cleavage primarily at the glutamic acid residues gave unique glycopeptides for each of the five glycosylation sites and the V-8 proteinase profiles for all forms of OMD were very similar. The same h.p.l.c. conditions were used throughout for the initial separation of the peptides from the V-8 proteinase digest.

The six hexose-containing peaks from the h.p.l.c. profile were purified further using shallow gradients, and the component peptides identified by *N*-terminal and amino acid analysis (Table 1). Each glycosylation site was eluted as a group of overlapping peaks (Fig. 1*a*), the complexity of which could be decreased by desialylation. This is to be expected, and is due to the removal of

Table 1. Amino acid sequences of the glycosylated V-8 proteinase peptides

* Denotes glycosylated asparagine. Upper level residues are substitutions in GP II.

Site	Peptide	Amino acid sequence
II	1	Y-N*-K-S-V-Q-E 37 … 43
V (GP II)	2	N*-G-T-V-S-R-Y-E 85 … 92
V (GP I)	3	N*-G-T-I-S-R-Y-V-G-G-Q-E 85 … 96
IV	6	N F * S-S (upper level) Y-Q-T-R-Q-D-Q-C-I-Y-N-T-T-Y-L-N-V-Q-R-E 65 … 84
III	11	I-Q-A-T-F-F-Y-F-T-P-N*-K-T-E 44 … 57
I	14	E-I-P-L-C-A-N-L-V-P-V-P-I-T-N*-A-T-L-D 1 … 19

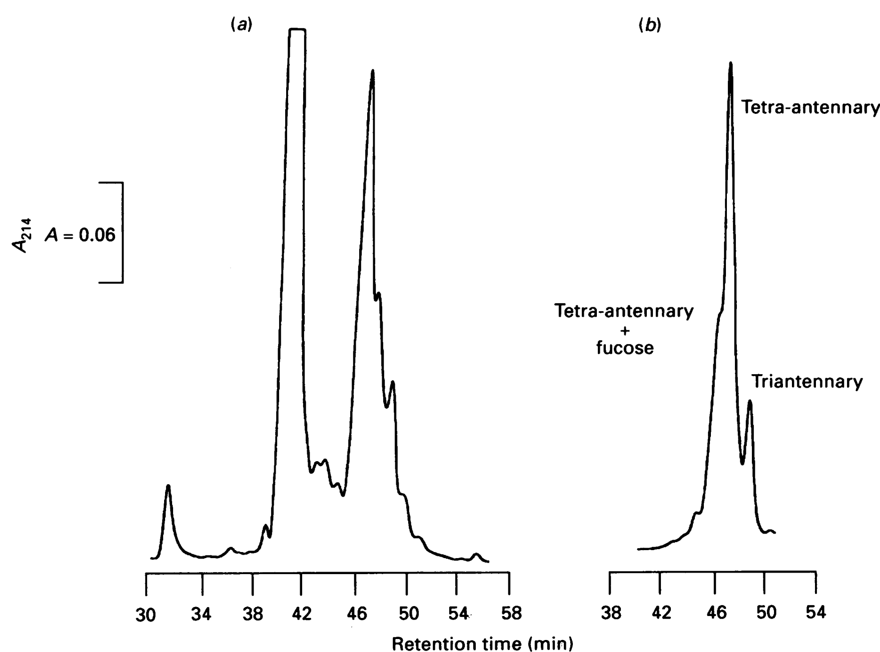

Fig. 1. Reversed-phase h.p.l.c. profiles for site V GP I eluted as a group of overlapping peaks (*a*) before desialylation or (*b*) after desialylation

Chromatographic conditions: 10–20% acetonitrile with 0.1% trifluoroacetic acid over 90 min. Detection was at 214 nm, flow rate 1 ml/min. See the text for peak assignments.

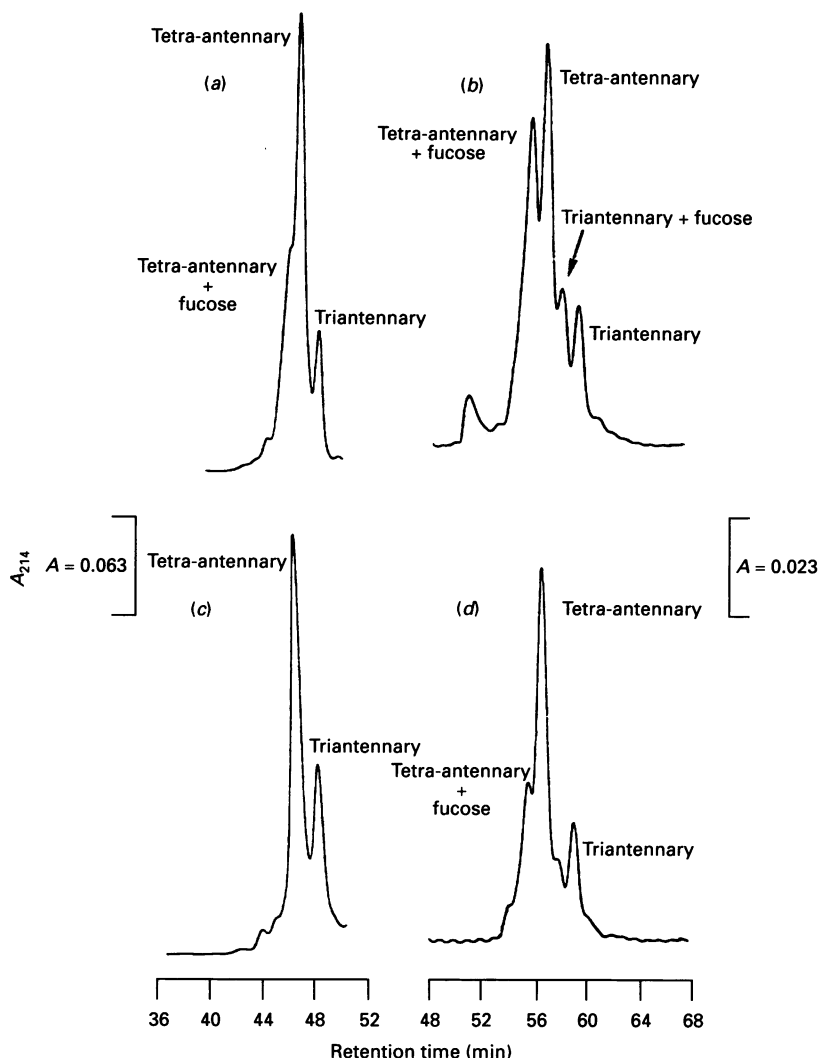

Fig. 2. Reverse-phase h.p.l.c. profiles for site V

(*a*) Desialylated GP I. (*b*) Desialylated GP II. (*c*) Desialylated GP I treated for 30 min with 0.05 M-H_2SO_4 at 80 °C. (*d*) Desialylated GP II treated for 30 min with 0.05 M-H_2SO_4 at 80 °C. Chromatographic conditions were: (*a*) and (*c*) 10–20 % acetonitrile with 0.1 % trifluoroacetic acid over 90 min, (*b*) and (*d*) 5–15 % acetonitrile with 0.1 % trifluoroacetic acid over 90 min. Detection was at 214 nm at a flow rate of 1 ml/min.

the minor site heterogeneity created by the known under-sialylation of the antennae in OMD [17]. Glycopeptides digested for 24 h with neuraminidase had no residual *N*-acetylneuraminic acid as measured by the Warren assay.

This single treatment provided the greatest enhancement of the chromatographic resolution of the glycan structures within a site (Fig. 1*b*). After rechromatography, the number of major peaks corresponded to the number of glycan forms present for that site as determined by Schmid *et al.* [9,10].

Dente *et al.* [40] determined that OMD consists of two genetic variants which code for two slightly different proteins. The amino acid substitution of glutamic acid for valine at position 92 [40] in gene product II (GP II) permitted the two gene-product forms of site V to be isolated. The numerous amino acid substitutions between residues 65 and 84 also allowed the identification of GP II for site IV. The h.p.l.c. profiles for the desialylated glycopeptides of site V indicated that GP I and GP II were differentially glycosylated (Figs. 2*a* and 2*b*). In both cases, the differences extend to fucosylation, since its removal by α-L-fucosidase or mild acid shows that at site V, GP I contains less fucose than does GP II (Figs. 2*c* and 2*d*). The remaining glycosylation sites were treated similarly to determine which sites and glycoforms contain fucose.

The positions of the major glycoforms in the chromatogram were definitively established using m.a.l.d.-t.o.f./m.s. to give the $(M+H)^+$ or $(M+Na)^+$ determinations for the parent ion. The m.a.l.d. measurement for the biantennary glycan at site II gave a weak signal for the parent ion, and therefore this sample was measured independently by f.a.b./m.s. in the positive mode to give the $(M+H)^{+\cdot}$ value for the parent ion. The expected average molecular masses were calculated by using the peptide backbone for the individual site plus the appropriate glycan chain. The m.a.l.d.-t.o.f./m.s. data were obtained at relatively low resolution, i.e. the isotopic multiplet of the molecular ion region is collapsed into a single broad peak, and therefore the data in Table 2 are reported as average molecular masses. The data confirm not only the glycan structure but also the peptide molecular mass and thus the assignment of the respective glycosylation site (Table 1). The sensitivity of m.a.l.d.-t.o.f./m.s. [41,42] and its relative tolerance to salts and buffers [43] offer significant advantages for the molecular mass determinations of glycopeptides in this mass range, even though the technique has been used to date almost

Table 2. m.a.l.d.-t.o.f./m.s. data obtained for the individual peaks from each glycosylation site

Molecular mass is the calculated average molecular mass (Da) summed from the known peptide sequence and the corresponding complex sugar. m.a.l.d. m/z is the average molecular mass of the $(M+H)^+$ or $(M+Na)^+$ value of the parent ion. NR, no results were obtained.

Site	Glycan structure	Molecular mass	m.a.l.d. m/z (± 2)
I	Biantennary	3656.9	3678.8 $(M+Na)^+$
	Triantennary	4022.2	4044.6 $(M+Na)^+$
	Triantennary + fucose	4168.4	4187.9 $(M+Na)^+$
II	Biantennary	2489.0*	2490.4 $(M+H)^+$
	Triantennary	2855.8	2856.8 $(M+H)^+$
	Tetra-antennary	3221.1	NR
III	Biantennary	3330.4	3330.1 $(M+H)^+$
	Triantennary	3695.7	3695.9 $(M+H)^+$
	Triantennary + fucose	3840.8	3841.9 $(M+H)^+$
	Tetra-antennary	4061.1	4060.9 $(M+H)^+$
	Tetra-antennary + fucose	4206.2	4206.9 $(M+H)^+$
IV	Triantennary (GP I)	4583.6	4586.9 $(M+H)^+$
	Triantennary (GP II)	4588.7	4591.0 $(M+H)^+$
	Triantennary + fucose (GP I)	4728.8	4729.6 $(M+H)^+$
	Triantennary + fucose (GP II)	4734.8	4737.1 $(M+H)^+$
	Tetra-antennary (GP I)	4949.0	4948.7 $(M+H)^+$
	Tetra-antennary (GP II)	4954.1	4954.7 $(M+H)^+$
	Tetra-antennary + fucose (GP I)	5095.2	5094.8 $(M+H)^+$
	Tetra-antennary + fucose (GP II)	5100.2	5101.5 $(M+H)^+$
	Tetra-antennary + 2 fucose (GP I)	5239.3	5239.8 $(M+H)^+$
	Tetra-antennary + 2 fucose (GP II)	5246.2	5246.7 $(M+H)^+$
	Tetra-antennary + GalGlcNAc	5314.3	5317.0 $(M+H)^+$
V (GP I)	Biantennary	2903.9	2905.2 $(M+H)^+$
	Triantennary	3269.2	3269.1 $(M+H)^+$
	Triantennary + fucose	3414.3	3414.7 $(M+H)^+$
	Tetra-antennary	3634.5	3633.7 $(M+H)^+$
	Tetra-antennary + fucose	3780.7	3781.8 $(M+H)^+$
	Tetra-antennary + 2 fucose	3926.8	3928.9 $(M+H)^+$

* The molecular mass obtained from f.a.b./m.s. in the positive ion mode to give the $(M+H)^+$ value of the parent ion.

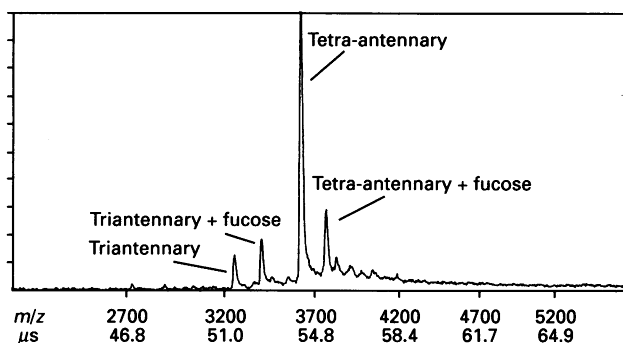

Fig. 3. Molecular ion region in the positive-ion m.a.l.d.m.s. of a fraction containing several glycoforms, predominantly the tetra-antennary component, of OMD site V GP I

The spectrum shown is the sum of 50 laser shots. Sinapinic acid matrix, 355 nm irradiation, accelerating voltage 28 kV. Total glycopeptide applied to the probe about 26 ng (6.6 pmol). For mass assignments, see Table 2.

exclusively for molecular mass determinations of much larger proteins and glycoproteins. The presence of various glycoforms is apparent in the spectra, and the molecular masses of even minor components could be determined. Fig. 3 shows the m.a.l.d.-t.o.f. mass spectrum obtained for the fraction containing primarily the site V (GP I) tetra-antennary glycoform, as well as lesser amounts of three others.

These data permitted the assembly of a composite elution order for the major glycoforms with a common polypeptide backbone [34], which was then used as a basis for analysing the glycoform subpopulations of OMD in the UR, WR and R fraction. The h.p.l.c. profiles of the individual glycosylation site for the UR, WR and R fractions were generated, and those for glycosylation site V GP I are illustrated in Fig. 4, as a typical example. The percentage of each glycoform present at the individual sites was calculated from the profiles, including those for unfractionated OMD. These results are found in Tables 3–6. It should be noted that although some of the fucose-containing peaks were identified, the chromatographic resolution was not great enough to permit their quantitative determination. Additionally, the resolution of the m.s. exceeded that of the chromatography such that other minor structural forms that were not visible chromatographically were identified (Table 2). The (small) contribution of these fucosylated and other minor forms has been included with that of the appropriate non-fucosylated glycoform. This simplification does not bear upon the conclusions reached, however. The individual site data revealed that the WR and R fractions contained only a single glycoform at site II, and that the remaining four sites in the R fraction also contained a considerable amount of this glycoform.

The nature of the glycoforms at sites I and II figure prominently in arguments concerning the differences between U, WR and R molecular variants [9,10,17], and where needed, confirmation of the assignment as biantennary was sought beyond retention-time matching with the established biantennary peaks from unfractionated OMD and the m.s. data. This was provided by the finding that all peaks assigned as biantennary bound quantitatively to a Con A affinity column, and were released quantitatively with methyl α-mannoside. Fig. 5 represents the bound and unbound fractions of site V GP I for the R fraction.

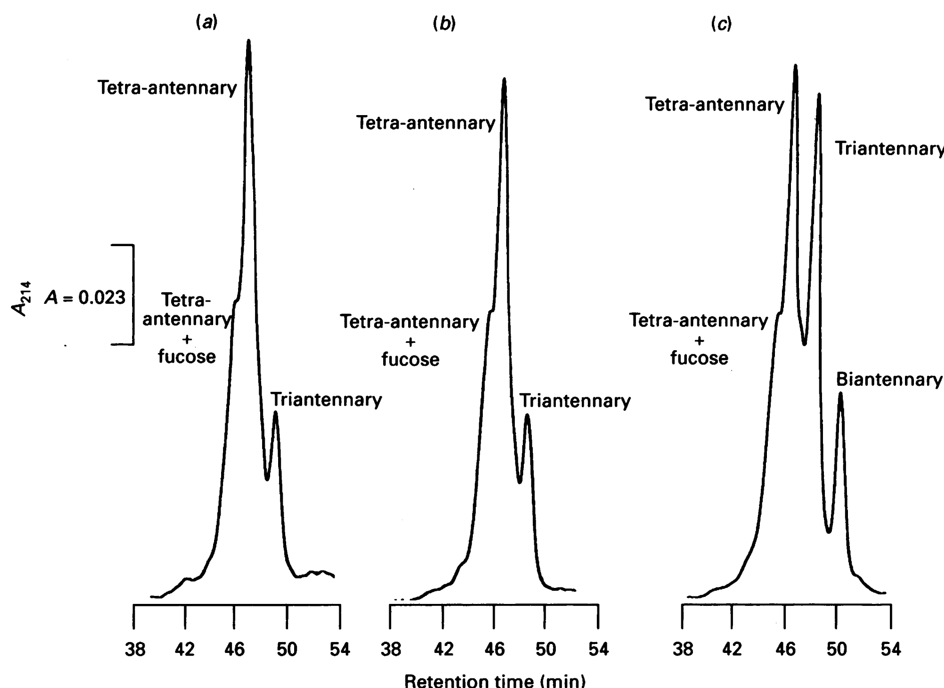

Fig. 4. Reverse-phase h.p.l.c. profiles for desialylated site V GP I

(*a*) From the UR fraction. (*b*) From the WR fraction. (*c*) From the R fraction. All chromatographic conditions were: 10–20% acetonitrile with 0.1% trifluoroacetic acid over 90 min. Detection was at 214 nm at a flow rate of 1 ml/min.

Table 3. Percentage of the complex glycan types at each site for unfractionated OMD

Site (*n*)	Percentage of tetra-antennary	Percentage of triantennary	Percentage of biantennary
I (10)	0.0	91.4±0.3	8.6±0.3
II (5)	14.6±0.3	24.8±0.6	60.6±0.4
III (9)	58.9±0.5	32.0±0.3	9.0±0.2
IV (9)	79.1±0.8	20.9±0.8	0.0
V GP I (12)	84.8±0.4	15.2±0.4	0.0
GP II (3)	77.2±0.1	22.8±0.2	0.0
Total all 5 sites	47.7±1.1	37.9±1.2	14.4±0.5

Table 4. Percentage of the complex glycan types at each site for the U fraction

Site (*n*)	Percentage of tetra-antennary	Percentage of triantennary	Percentage of biantennary
I (1)	0.0	100.0	0.0
II (3)	38.6±0.2	61.4±0.2	0.0
III (2)	71.1±0.1	28.9±0.1	0.0
IV (2)	81.3±0.1	18.7±0.1	0.0
V GP I (2)	84.7±0.1	15.3±0.1	0.0
GP II (1)	90.6	9.4	
Total all 5 sites	52.6±0.3	47.4±0.3	0.0

For site I, R and WR fractions were 58% and 9% biantennary respectively, whereas at site II, both variants were 100% biantennary. Additionally, the biantennary glycoform was present at all five sites of the R fraction, but only I–III of the WR fraction. The UR fraction contained no biantennary glycans at all, as would be expected from its lack of reactivity towards Con A.

Table 5. Percentage of the complex glycan types at each site for the WR fraction

Site (*n*)	Percentage of tetra-antennary	Percentage of triantennary	Percentage of biantennary
I (1)	0.0	91.3	8.7
II (3)	0.0	0.0	100.0
III (2)	57.9±0.1	32.0±0.0	10.0±0.1
IV (2)	71.0±0.1	29.0±0.1	0.0
V GP I (2)	83.8±0.1	16.2±0.1	0.0
GP II (1)	81.0	19.0	0.0
Total all 5 sites	42.4±0.1	33.9±0.1	23.7±0.1

Table 6. Percentage of the complex glycan types at each site for the R fraction

Site (*n*)	Percentage of tetra-antennary	Percentage of triantennary	Percentage of biantennary
I (1)	0.0	42.2	57.8
II (3)	0.0	0.0	100.0
III (2)	34.7±0.1	30.0±0.0	35.3±0.1
IV (2)	66.8±0.1	26.0±0.0	7.2±0.1
V GPI (2)	57.0±0.1	30.3±0.1	12.7±0.1
GP II (1)	7.0	93.0	0.0
Total all 5 sites	29.2±0.2	28.8±0.1	42.0±0.2

These results coupled with the m.a.l.d.-t.o.f./m.s. data necessitate that the glycans to be found at each glycosylation site [9,11] be modified according to Table 2.

DISCUSSION

As the methodologies for analysing single-site glycosylation

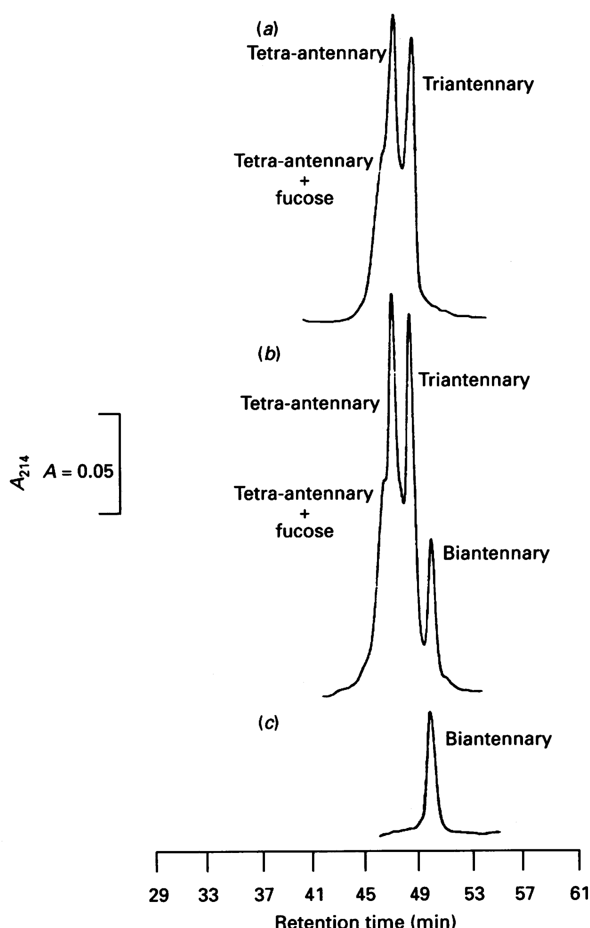

Fig. 5. Reverse-phase h.p.l.c. profiles of desialylated site V GP I from the R fraction

(*a*) The unbound fraction after rerunning on Con A. (*b*) Before Con A fractionation. (*c*) The bound fraction eluted from Con A. All chromatographic conditions were: 10–20 % acetonitrile with 0.1 % trifluoroacetic acid over 90 min. Detection was at 214 nm at a flow rate of 1 ml/min.

patterns in multiglycosylated glycoproteins have evolved, so has the realization of the remarkable diversity accessible to glycoproteins in both their surface glycan patterns and biological activity, whether mediated through the glycans or the protein core. This is no less true for OMD in its reactivity towards the lectin Con A. In early work, Bayard & Kerckaert [16] suggested that the Con A variants were each unique glycoforms, i.e. that the R-fraction glycans are all biantennary and the UR-fraction glycans all tetra-antennary. This could not be confirmed by other workers [17,44,45], and most recently, Bierhuizen *et al.* [17] concluded that the WR and R fractions contain one and two biantennary glycans respectively on the basis of the percentage content of biantennary glycans that they found in these species.

In the present work, being able to quantify the glycoforms present at each site allows a number of conclusions to be drawn which extend the earlier results [9,10,17] concerning the placement of the biantennary glycans. If the percentages of the glycoforms present in the unfractionated OMD and the three Con A variants (Tables 3–6) are compared with the results of Bierhuizen *et al.* [17], it is found that the two data sets are quite similar. Individual site analysis, however, reveals that the R fraction is capable of possessing biantennary glycans at all five sites in GP I with only site II being exclusively biantennary. The apparent lack of biantennary glycans at sites IV and V in the unfractionated OMD is readily attributed to the low percentages of biantennary glycans present in the unfractionated OMD (about 0.3 % for sites IV and V), being below the detection limit of the method.

The WR fraction contains biantennary glycans at sites I, II and III, with site II being exclusively biantennary, in common with the R fraction. This means that 10–19 % of the WR glycoforms contain two or three biantennary chains, and these at least would be expected to bind to Con A [46]. That they do not, and the observation that the WR fraction disappears on reduction and alkylation of unfractionated OMD [16] suggest that access to the mannosyl 3,4,6-OH groups in the WR fraction is restricted by either the glycan chains from other sites or interaction of the glycan with the protein core.

The results confirm the fact that the UR fraction contains no biantennary glycans [17], with site I being exclusively triantennary. The overall ratio of triantennary to tetra-antennary glycans was nearly 1:1 which is in good agreement with the results of Bierhuizen *et al.* [17].

The populations produced by fractionation of multiglycosylated proteins using lectins may reveal something of the interdependency of processing at different sites [47]. The detailed profile of the U fraction (Table 4) indicates that there exists a gradient of processing from site I to site V, and site V being more highly processed to the tetra-antennary chains. The other detailed profiles (Tables 5 and 6) basically maintain the trend even though there is a bias to biantennary glycans at site II. In the U fraction, of course, there is a reversed situation in the triantennary species, which is also maintained in the WR fraction except at site II. This increased processing further from the *N*-terminus is in contrast with the observations of Pollack & Atkinson [48], and has also been seen in Thy-1 [49]. The significance of this, however, is unknown and must await comparison with the detailed structures of other glycoproteins.

In a multiglycosylated glycoprotein synthesized in the liver such as OMD, these data should reflect the temporal history of the processing control that has occurred, and much of what is seen in the U fraction is apparently a result of the control of *N*-acetylglucosaminyltransferase IV, which is responsible for the initiation to the triantennary form [29,30]. Control of this transferase at site II is severe but not absolute, since the U fraction contains no biantennary glycans. The lack of processing to triantennary chains at site II in the R and WR fractions, even when other sites are mostly converted, indicates that the action of *N*-acetylglucosaminyltransferase IV at site II is independent of its prior action at all the other sites. The most distinctive feature of the data, however, is the complete lack of conversion of the triantennary glycan at site I to tetra-antennary glycan, i.e. the inability of *N*-acetylglucosaminyltransferase V to utilize the site I triantennary glycan as a substrate.

Although the relative activities of the glycosyltransferases may account for some of the inter- and intra-site heterogeneity in glycoproteins [50–55], the recognition of other influential factors has led to the development of two principal models. The 'accessibility' model [27–29] proposes that the extent of processing at individual sites is determined primarily by the ability of the processing enzymes to approach their oligosaccharide substrates, and that this may be further modulated by the incorporation of key glycosyl residues, which convert substrates into non-substrates and vice versa. There is no clear evidence that these mechanisms are operating here. For example, at site I the model would require that the addition of *N*-acetylglucosamine by *N*-acetylglucosaminyltransferase IV at one or more of sites III, IV and V provides the steric hindrance necessary to block site I (Tables 5 and 6). Although this cannot be ruled out, it seems unlikely in view of the limited extent of the chain elongations at this point in the processing and the fact that site I is still accessible to enzymes that operate later in the pathway.

The 'site-directed' processing model [25] recognizes that the interaction of the glycan substrate with the neighbouring protein structure may stabilize three-dimensional glycan structures or their subsets which are then no longer substrates for the next process in the sequential pathway but can remain so for later ones. The model is derived from the observation that *N*-acetylglucosaminyltransferase III acts only on the glycan substrate in which the mannose-α-1,6 arm is folded back over the chitobiose core [25,26,56] and tentatively predicts that a number of processing decisions are based on conformational restrictions imposed on this branching mannose. This model then emphasizes the role of the conformation of the glycan and its interaction with the protein in determining the pathway of processing and could account for the processing restrictions seen at sites I and II. Thus at site I, *N*-acetylglucosaminyltransferase V cannot use the mannose-α-1,6 arm as a substrate. In this case the unrestricted state is highly unfavourable since no processing is seen. The processing by *N*-acetylglucosaminyltransferase IV at site II might be similarly controlled, the distributions finally seen in Tables 4–6 reflecting the populations of the mannose-α-1,3 arm that are accessible (Table 4) and inaccessible (Tables 5 and 6) at this point in the processing pathway.

The presence of a minor second gene product is a potential complication, and we have therefore restricted the discussion above to those features of the data about which conclusions can be drawn independently of GP II. To these we can add, however, that sites I and II are under very similar if not identical processing control for *N*-acetylglucosaminyltransferases V and IV respectively, implying that the local polypeptide environments are also very similar and that the differences seen elsewhere are probably not the result of the differential expression of GP I and II in different cells [45,57–60]. At site V (Asn-85), the two GPs are differentially glycosylated in both fucosylation and the extents of processing seen in the different Con A reactions. In terms of the site-directed processing model, this may speculatively result from the amino acid substitutions that are proximal to site V.

REFERENCES

1. Pervaiz, S. & Brew, K. (1987) FASEB J. **1**, 209–214
2. Pevsner, J., Reed, R. R., Feinstein, P. G. & Snyder, S. H. (1988) Science **241**, 336–339
3. Treuheit, M. & Halsall, H. B. (1991) Chromatographia **31**, 63–67
4. Treuheit, M. & Halsall, H. B. (1991) Biomed. Chromatogr., in the Press
5. Bacchus, H. (1965) Cancer **18**, 1285–1291
6. Martyn, J. A., Abernethy, D. R. & Greenblatt, D. J. (1984) Clin. Pharmacol. Ther. **35**, 535–540
7. Kremer, J. M., Wilting, J. & Janssen, L. H. (1988) Pharmacol. Rev. **40**, 1–47
8. Schmid, K., Nimberg, R. B., Kimura, A., Yamaguchi, H. & Binette, J. P. (1977) Biochim. Biophys. Acta **492**, 291–302
9. Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J., Binette, J. & Schmid, K. (1978) Biochemistry **17**, 5206–5219
10. Schmid, K., Binette, J. P., Dorland, K., Vliegenthart, J., Fournet, B. & Montreuil, J. (1979) Biochim. Biophys. Acta **581**, 356–359
11. Yoshima, H., Matsumoto, A., Mizuochi, T., Kawasaki, T. & Kobata, A. (1981) J. Biol. Chem. **256**, 8476–8484
12. Bennett, M. & Schmid, K. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 6109–6113
13. Laine, E., Couderc, R., Roch-Arveiller, M., Vasson, M. P., Giroud, J. P. & Raichvarg, D. (1990) Inflammation **14**, 1–9
14. Bories, P. N., Guenounou, M., Kodari, E., Agneray, J. & Durand, G. (1987) Biochem. Biophys. Res. Commun. **147**, 710–715
15. Lejeune, P., Mallet, B., Farnarier, C. & Kaplanski, S. (1989) Biochim. Biophys. Acta **990**, 122–127
16. Bayard, B. & Kerckaert, J.-P. (1980) Biochem. Biophys. Res. Commun. **95**, 777–784
17. Bierhuizen, M., Govers, C., Ferwerda, W., Kueleman, C., Pos, O. & van Dijk, W. (1988) Eur. J. Biochem. **175**, 387–394
18. Nicollet, I., Lebreton, J. P., Fontaine, M. & Hiron, M. (1981) Biochim. Biophys. Acta **668**, 235–245
19. Hansen, J. E. S., Larsen, V. A. & Bog-Hansen, T. C. (1984) Clin. Chim. Acta **138**, 41–47
20. Pos, O., Oostendorp, R., Stelt, M., Scheper, R. & van Dijk, W. (1990) Inflammation **14**, 133–141
21. Pos, O., van Dijk, W., Ladiges, N., Linthorst, C., Sala, M., Tiel, D. & Boers, W. (1988) Eur. J. Cell Biol. **46**, 121–128
22. Wells, C., Bog-Hansen, T. C., Cooper, E. H. & Class, M. R. (1981) Clin. Chim. Acta **109**, 59–67
23. Yet, M.-G., Shao, M.-C. & Wold, F. (1988) FASEB J. **2**, 22–31
24. Dahms, N. M. & Hart, G. W. (1986) J. Biol. Chem. **261**, 13186–13196
25. Carver, J. P. & Cummings, D. A. (1987) Pure Appl. Chem. **59**, 1465–1476
26. Wooten, E. W., Bazzo, R., Edge, C. J., Dwek, R. A. & Rademacher, T. W. (1990) Eur. Biophys. J. **18**, 139–148
27. Paulson, J. C. & Colley, K. J. (1989) J. Biol. Chem. **264**, 17615–17618
28. Schachter, H., Narasimham, S., Gleeson, P. & Vella, G. (1983) Can. J. Biochem. Cell Biol. **61**, 1049–1066
29. Brockhausen, I., Carver, J. P. & Schachter, H. (1988) Can. J. Biochem. Cell Biol. **66**, 1134–1151
30. Halsall, H. B., Kirley, T. & Friedman, M. (1982) Prep. Biochem. **12**, 111–120
31. Ikenaka, T., Ishiguro, M., Emura, J., Kaufmann, H., Isemura, S., Bauer, W. & Schmid, K. (1972) Biochemistry **11**, 3817–3829
32. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) Anal. Chem. **28**, 350–356
33. Winter, G. (1986) in Practical Protein Chemistry: A Handbook (Darbre, A., ed.), pp. 367–374, John Wiley and Sons, Chichester, New York, Brisbane, Toronto and Singapore
34. Treuheit, M. & Halsall, H. B. (1991) Chromatographia **31**, 478–480
35. Warren, L. (1959) J. Biol. Chem. **234**, 1971–1975
36. O'Kennedy, R. (1980) Ir. J. Med. Sci. **101**, 92–96
37. Abe, K., McKibbin, J. & Hakamori, S. (1983) J. Biol. Chem. **258**, 11793–11797
38. Laemmli, U. K. (1970) Nature (London) **227**, 680–685
39. Beavis, R. C. & Chait, B. T. (1989) Rapid Commun. Mass Spectrom. **3**, 432–435
40. Dente, L., Pizza, M. G., Metspalu, A. & Cortese, R. (1987) EMBO J. **6**, 2289–2296
41. Karas, M. & Hillenkamp, F. (1988) Anal. Chem. **60**, 2299–2301
42. Karas, M., Ingendoh, A., Bahr, U. & Hillenkamp, F. (1989) Biomed. Environ. Mass Spectrom. **18**, 841–843
43. Beavis, R. C. & Chait, B. T. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6873–6877
44. Chandrasekaran, E. V., Davila, M., Nixon, D. & Mendicino, J. (1984) Cancer Res. **44**, 1557–1567
45. Berman, E. & Allerhand, A. (1982) Proc. Int. Carbohydr. Symp. 11th, Vancouver, abstr. IV-8
46. Krusius, T., Finne, J. & Rauvala, H. (1976) FEBS Lett. **71**, 117–120
47. Rademacher, T. W. & Dwek, R. A. (1989) CIBA Found. Symp. **145**, 241–256
48. Pollack, L. & Atkinson, P. H. (1983) J. Cell Biol. **97**, 293–300
49. Parekh, R. B., Tse, A. G. D., Dwek, R. A., Williams, A. F. & Rademacher, T. W. (1987) EMBO J. **6**, 1233–1244
50. Yamashita, K., Tachibana, Y., Ohkura, T. & Kobata, A. (1985) J. Biol. Chem. **260**, 3963–3969
51. Kim, J. S., Perdomo, J., Whitehead, J. S. & Curtis, K. J. (1972) J. Clin. Invest. **51**, 2033–2039
52. Lammers, G. & Jamieson, J. C. (1989) Biochem. J. **261**, 389–393
53. Nakaishi, H., Sanai, Y., Shibuya, M. & Nagai, Y. (1988) Biochem. Biophys. Res. Commun. **150**, 766–774
54. Matsuura, H., Greene, T. & Hakomori, S. (1989) J. Biol. Chem. **264**, 10472–10476
55. Iwase, H., Kato, Y. & Hotta, K. (1987) Proc. Int. Symp. Glycoconjugates 9th, 176
56. Brisson, J.-R. & Carver, J. P. (1983) Biochemistry **22**, 3680–3686
57. Parekh, R. B., Dwek, R. A., Thomas, J. R., Opdenakker, G. & Rademacher, T. W. (1989) Biochemistry **28**, 7644–7662
58. Parekh, R. B., Dwek, R. A., Rudd, P. M., Thomas, J. R. & Rademacher, T. W. (1989) Biochemistry **28**, 7670–7679
59. Swiedler, S. J., Freed, H. H., Tarentino, A. L., Plummer, T. H. & Hart, G. W. (1985) J. Biol. Chem. **260**, 4046–4054
60. Neel, D., Merlu, B., Turpin, E., Rabourdin-Combe, C., Mach, B., Goussault, Y. & Charron, D. J. (1987) Biochem. J. **244**, 433–442

Received 8 May 1991/8 August 1991; accepted 25 September 1991