Molecular-mass heterogeneity of *Griffonia simplicifolia* lectin IV subunits

Differences in the oligosaccharide moieties in the *N*-terminal region

Pandurang V. NIKRAD,* Joyce R. PEARLSTONE,† Michael R. CARPENTER,† Raymond U. LEMIEUX* and Lawrence B. SMILLIE†‡

*Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2,

and †MRC of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Lectin IV of *Griffonia simplicifolia* ($M_r \sim 56000$), which has a strong affinity for both the Lewis b and Y blood-group determinants, is a dimeric protein of two subunits, α (29 kDa) and β (27 kDa), separable by SDS/PAGE and containing covalently linked oligosaccharide. After digestion with *N*-glycanase, the protein migrates as a single band with a mobility identical with that of the β -subunit. After cleavage with hydroxylamine of ^3H -labelled, but otherwise intact, lectin, the radioactively labelled oligosaccharide was found to be associated with two blocked *N*-terminal peptides separable by h.p.l.c. and having identical amino acid compositions. One of these had three or four glucosamine residues per molecule, whereas the other had only one or two. Sequence analyses of these, as well as of a 21 kDa hydroxylamine-cleaved fragment and of the intact lectin pretreated with pyroglutamate aminopeptidase, have provided a unique sequence for residues 1–62 of the two subunits. Evidence is presented for two sites of *N*-linked oligosaccharide attachment at Asn-5 and Asn-18. Whereas the α -subunit has oligosaccharide linked to both sites, the β -subunit has carbohydrate associated with only one (Asn-18). Sugar analyses of the whole lectin reveal a monosaccharide composition of $(\text{Xyl})_3(\text{Fuc})_2(\text{Man})_{10}(\text{GlcNAc})_6$, representing 6.4 % of the mass of the molecule. Taken together with the susceptibility of the Asn-5 linkage (but not of Asn-18) to *N*-glycanase digestion, the observations indicate that the structures of the oligosaccharides at residues 5 and 18 are different.

INTRODUCTION

A fourth lectin (GSIV) from the seeds of the plant *Griffonia simplicifolia* was first isolated and characterized in 1982 [1]. With a molecular mass of 56 kDa the glycoprotein was reported to consist of two subunits separable by SDS/PAGE and which we now designate as α (~ 29 kDa) and β (~ 27 kDa). The carbohydrate portion was reported to be close to 3 % and to consist of residues of D-GlcNAc, D-Man, D-Xyl and L-Fuc [1]. The lectin became of exceptional interest in studies of molecular recognition [2,3] when it was discovered [1,3] that it had a strong affinity for both the Lewis b (Le^b) and Y human blood-group determinants in the form of synthetic tetrasaccharides. Systematic studies of a wide number of chemically modified structures then led to the conclusion that these determinants are complexed by GSIV through the recognition of a topographical feature common to both tetrasaccharides and involving the two fucose units and the terminal end of the galactose [2]. Modifications of the β -D-GlcNAc units in both the Le^b -OMe and Y-OMe structures had essentially no effect on binding.

A knowledge of the three-dimensional structures of both the protein and its complex with the oligosaccharide is necessary for the achievement of a detailed understanding of these observations. Fortunately, in this regard, both the lectin and its complexes with Le^b -OMe and Y-OMe and several derivatives of Le^b -OMe have been crystallized [4] in excellent condition for single-crystal X-ray crystallography.

The interpretation of the electron-density maps generated by X-ray diffraction analysis requires a knowledge of the amino acid sequence of the protein; consequently an elucidation of the latter was undertaken in our laboratories. In the present paper we

demonstrate that amino acid residues 1–62 are identical in both subunits of the dimeric protein. However, whereas the α -subunit has two sites of *N*-linked oligosaccharide attachment at asparagine residues 5 and 18, the β -subunit has carbohydrate associated with only one of these sites (Asn-18). These observations account for the difference in apparent M_r of the two subunits as estimated by SDS/PAGE. The *N*-linked oligosaccharide at Asn-5 was susceptible to *N*-glycanase, whereas that at Asn-18 was not. Sugar analyses have given a carbohydrate content of $(\text{Xyl})_3(\text{Fuc})_2(\text{Man})_{10}(\text{GlcNAc})_6$, representing 6.4 % of the mass of the whole molecule. Taken together these observations indicate that the structures of the oligosaccharides at residues 5 and 18 are different.

MATERIALS AND METHODS

The GSIV protein was isolated from seeds by affinity chromatography as described previously [1]. Digestion of the protein (0.5 mg, 9 nmol) with *N*-glycanase (EC 3.2.2.18) (purchased from Genzyme Corp., Boston, MA, U.S.A.) was performed in 0.7 ml of 0.25 M-sodium phosphate buffer, pH 8.6, containing 1 % SDS at 37 °C for 20 h (500 units/mg of GSIV).

For digestion with calf liver pyroglutamate aminopeptidase (EC 3.4.19.3; Boehringer-Mannheim), 1 mg of heat-denatured (5 min in a boiling-water bath) GSIV was incubated at a final enzyme/substrate weight ratio of 1:12 in 1 ml of 0.1 M-potassium phosphate buffer, pH 7.8, containing 10 mM-EDTA, 5 mM-dithiothreitol, 0.05 % glycerol, 0.5 % sodium cholate at 4 °C for 48 h. Addition of enzyme was in two equal portions at 0 and 24 h.

Abbreviations used: GSIV, *Griffonia simplicifolia* lectin IV; PTH, 3-phenyl-2-thiohydantoin; -OMe, methyl ester.

‡ To whom correspondence and reprint requests should be sent.

Radiolabelling of the carbohydrate of GSIV was as follows: to the protein (12.2 mg, 218 nmol) dissolved in 1.65 ml of 0.2 M-sodium acetate buffer, pH 4.5 was added 122 μ l of 0.1 M-sodium metaperiodate (12.2 μ mol) and the reaction mixture stirred in the dark at 4 °C for 4 h. After addition of ethylene glycol (0.82 ml, 15 mmol) and stirring at 0 °C for 30 min, the solution was adjusted to pH 7.5 with 2 M-NaOH and concentrated to 1 ml in a 5 ml Amicon cell. The solution was transferred to a vial containing NaB^3H_4 (378 μ g; 500 mCi/mmol) and stirred overnight at 4 °C. After acidification to pH 4.5 with 0.1 M-acetic acid, containing 0.9% (w/v) NaCl, and incubation at room temperature for 30 min, the ^3H -labelled GSIV was dialysed thoroughly against water and freeze-dried [5].

Unlabelled GSIV or ^3H -labelled GSIV was dissolved in freshly prepared 2 M-hydroxylamine/0.1 M K_2CO_3 , pH 10.0, at a concentration of 1–2 mg/ml, incubated at 45 °C for 4 h, dialysed against water and freeze-dried.

SDS/PAGE of labelled and unlabelled protein was performed with the discontinuous buffer system of Laemmli [6], using either a constant 13% or a 7–20% (w/v) gradient polyacrylamide concentration. Protein bands were detected by staining with Coomassie Brilliant Blue R-250 or by autoradiography. Carbohydrate staining was with the periodate/Schiff reagent [7].

Neutral sugar analyses of GSIV were based on the method of Albersheim *et al.* [8]. In the present method the amounts of alditol acetates present in the reduced acid hydrolysates of the glycoprotein were determined by isotopic-dilution analyses. The procedure was based on calibration plots established by g.c.–m.s. analyses of mixtures of the acetates of xylose, fucose and mannose with near-equimolar amounts of their $1\text{-}^3\text{H}$ derivatives. The acetates were well separated, and the mass spectra using single-ion monitoring [9,10] provided separate lines for the acetylated alditol- NH_4^+ and $[1\text{-}^3\text{H}]\text{alditol-}\text{NH}_4^+$ molecular ions. Plots of the ratios of the heights of the lines for a given alditol acetate against the known ratios of the alditol acetates in the mixture were found, as expected, to be linear. Slopes of the lines for the three alditols were 0.98 ± 0.02 . In the determination of the sugar content of GSIV, the $[1\text{-}^3\text{H}]\text{alditol acetates}$ were added to the lectin in amounts three times that expected [1] to be formed using Albersheim's [8] procedure (hydrolysis of the glycoproteins in 2 M-trifluoroacetic acid for 2 h at 110 °C, followed by borohydride reduction and acetylation). The molar amounts of sugar relative to GSIV were based on the dry weight of protein.

Methods for obtaining amino acid compositions using a Durrum model D500 analyser, manual dansyl determinations and assignment of acids and amides from electrophoretic mobilities were all described previously [11–17]. Amino acid sequences were obtained by automated Edman degradation with an Applied Biosystems gas-phase sequencer (model 470A), using the manufacturer's standard 03RPTH procedure, coupled with an Applied Biosystems model 120A h.p.l.c. apparatus for on-line identification of 3-phenylthio-2-hydantoin (PTH) derivatives. An Applied Biosystems model 130A separation system equipped with a model 420A derivatizer and a 920A data analysis module superseded the Durrum D500 amino acid analyser part of the way into this work, and some amino acid compositions and glucosamine analyses were obtained using this instrument system. Samples for glucosamine analyses were hydrolysed for 24 h at 100 °C. These were corrected for destructive losses by subjecting standard samples of GlcNAc to identical hydrolytic treatment.

Trypsin (EC 3.4.21.4) and chymotrypsin (EC 3.4.21.1) digestions were performed in 50 mM- NH_4HCO_3 , pH 8.5, for 24 h at 37 °C using a final enzyme/substrate molar ratio of 1:40. In the latter case, chymotrypsin was added in five aliquots at 4 h intervals. Digestion with thermolysin (EC 3.4.24.4) was in 50 mM- NH_4HCO_3 , pH 7.8, containing 1 mM- CaCl_2 , using an enzyme/

substrate molar ratio of 1:50. Endoproteinase Asp-N cleavages in 50 mM-sodium phosphate, pH 7.9, were at an enzyme/substrate weight ratio of 1:63 for 18 h at 37 °C. Bovine pancreatic trypsin [tosylphenylalanylchloromethane ('TPCK')-treated] was from Worthington Biochemical Corp. Thermolysin (*Bacillus thermoproteolyticus*) and α -chymotrypsin [tosyl-lysylchloromethane ('TLCK')-treated; bovine pancreatic] were purchased from Sigma Chemical Co. Endoproteinase Asp-N (from a mutant of *Pseudomonas fragi*) was a product of Boehringer-Mannheim.

Peptides were isolated by high-voltage paper electrophoresis at pH 6.5 and/or h.p.l.c. on a Spectra-Physics model SP4270 instrument, as described previously [11,12]. A Vydac 218TP54 C_{18} reverse-phase column (4.6 mm $\times$ 25 cm), and linear gradients of acetonitrile in 0.1% trifluoroacetic acid, were used.

RESULTS AND DISCUSSION

Native GSIV consists of two subunits, GSIV α and GSIV β , having molecular masses of ~ 29 and ~ 27 kDa respectively (Fig. 1a, lane 3). Various attempts to separate these on chromatographic columns using a variety of matrices in both non-denaturing and denaturing conditions were unsuccessful. When the α - and β -subunits were isolated by electroelution from the SDS/PAGE gel and analysed, their amino acid compositions were found to be very similar (results not shown). The *N*-termini of both subunits were blocked, since no sequence was obtained upon application of the whole protein to automated sequencing. In agreement with a previous report [1], both subunits contain carbohydrate, as indicated by autoradiography of ^3H -labelled GSIV (Fig. 1b, lane 1) and by staining with periodate/Schiff reagent (Fig. 1c, lane 1). Upon treatment with *N*-glycanase, an enzyme capable of cleaving *N*-linked oligosaccharide polypeptides to produce free oligosaccharide and polypeptide, the two bands corresponding to the α - and β -subunits collapsed into one (Fig. 1a, lane 1) with a mobility identical with that of the β -subunit (27 kDa). The difference in molecular mass between

Fig. 1. SDS/PAGE of GSIV before and after treatment with *N*-glycanase and hydroxylamine

(a) (7–20% Gradient gel) stained with Coomassie Blue; (b) (13% gel) autoradiograph; (c) (7–20% gradient gel) stained with periodate/Schiff reagent. Lanes (a)3 and (c)1 are whole unlabelled GSIV; lane (b)1 is ^3H -labelled GSIV. Lanes (a)1 and (c)2 are after treatment with *N*-glycanase; lanes (a)2 and (b)2 are after hydroxylamine treatment.

them is thus presumably due to a difference in their carbohydrate content. However, the data indicate that the enzyme has not removed all of the carbohydrate, since both the original GSIV β -band and the single band obtained after enzymic treatment stain for carbohydrate (Fig. 1c, lanes 1 and 2).

In order to determine the site(s) where the carbohydrate is attached on the α - and β -subunits, the oligosaccharides of the native lectin were radiolabelled with ^3H (Fig. 1b, lane 1) and then cleaved with hydroxylamine (see the Materials and methods section). Native lectin, without radiolabelling, was also treated with this reagent. Analysis of the products by SDS/PAGE (Fig. 1b, lane 2, and Fig. 1a, lane 2) indicated that cleavage was incomplete, leaving a significant proportion of intact α - and β -chains. However, cleavage products included a single band at ~ 21 kDa (designated fragment $\text{H2}_{\alpha\beta}$) which was only weakly radioactive. In addition, two strongly radioactive bands of molecular mass approx. 7.5 kDa (fragment H1_{α}) and 6 kDa (fragment H1_{β}) were also produced (see Fig. 1b, lane 2). The hydroxylamine digest of the native lectin was then applied to the C_{18} reverse-phase h.p.l.c. column (Fig. 2). Peaks 1 and 2 had identical amino acid compositions (Table 1) and blocked N -termini, since they failed to sequence in the automated sequencer. They were therefore derived from the N -terminal ends of the protein's polypeptide chains. However, they did differ in their glucosamine content, with H1_{α} having approximately twice that of H1_{β} (Table 1). The differences in their molecular masses (Fig. 1b, lane 2) and their elution positions on the C_{18} reverse-phase column (Fig. 2) may therefore be attributed to differences in their carbohydrate content. Peak 3 contained a band of molecular mass ≈ 10 kDa shown after further tryptic digestion to originate from the 21 kDa fragment as a result of non-specific chemical cleavage (results not shown). The 21 kDa fragment, $\text{H2}_{\alpha\beta}$, was eluted under the same peak as uncleaved lectin (Fig. 2, peak 4). Since attempts to isolate the components of peak 4 were unsuccessful, an amino acid composition for the 21 kDa fragment was not obtained. However, application of this mixture (peak 4) to the automated sequencer resulted in a unique sequence (Table 2; fragment $\text{H2}_{\alpha\beta}$). This result was obtained because the intact lectin has a blocked N -terminus. The N -terminal glycine residue

Table 1. Amino acid and glucosamine composition of N -terminal hydroxylamine fragments*

Amino acid or sugar	Composition (mol/mol)	
	H1_{α}	H1_{β}
Asx	7.0 (7)	7.0 (7)
Thr	7.0 (7)	7.0 (7)
Ser	2.0 (2)	2.0 (2)
Glx	2.4 (3)	2.7 (3)
Gly	3.0 (3)	3.3 (3)
Ala	3.0 (3)	3.4 (3)
Val	1.0 (1)	1.0 (1)
Ile	2.0 (2)	2.0 (2)
Leu	4.0 (4)	4.0 (4)
Tyr	1.8 (2)	1.8 (2)
Phe	3.4 (3)	3.0 (3)
Lys	2.0 (2)	1.8 (2)
Arg	1.0 (1)	0.9 (1)
Pro	2.0 (2)	1.8 (2)
Trp†	— (1)	— (1)
Total	(43)	(43)
Glucosamine	3.3	1.5

* Values in parentheses are those deduced from sequence analyses (Fig. 3).

† Trp was not determined, but one residue was identified by sequence analysis (Fig. 3).

Fig. 2. Fractionation of hydroxylamine-treated GSIV by reverse-phase h.p.l.c.

The sample (100 μl , 8.9 nmol) was applied to a Vydac 218 TP54 column equilibrated with 0.1% trifluoroacetic acid (1 ml/min). A gradient of 0–30% acetonitrile in 0.1% trifluoroacetic acid (1.5%/min), followed by a gradient of 31–65% acetonitrile (1.0%/min), was applied (.....). Absorbance was monitored at 220 nm (—) and 280 nm (---) and appropriate fractions were collected manually.

of this sequence is consistent with the cleavage by hydroxylamine of an Asn–Gly bond.

Since the above data indicated that the sites of carbohydrate attachment were limited to the N -terminal 43 residues of the subunits of the protein, represented by fragments H1_{α} and H1_{β} , further efforts were directed towards the identification of these in the sequences of the fragments. Treatment of intact GSIV with pyroglutamate aminopeptidase followed by application of the product (PA of Table 2) to automated sequencing provided a unique sequence of 14 residues. It is to be noted that PTH-Asn was recovered in low yield at cycle 4 and that this is a potential site of oligosaccharide attachment since PTH-Thr was identified in cycle 6.

For purposes other than those described here portions of fractions H1_{α} and H1_{β} were combined to give fraction $\text{H1}_{\alpha\beta}$. A sample of this was digested with trypsin and fractionated by h.p.l.c. Three peptides, $\text{H1}_{\alpha\beta}$ -T1, -T2, and -T3 were recovered. Peptide $\text{H1}_{\alpha\beta}$ -T1 had a blocked N -terminus, but an amino acid composition (Table 3) consistent with amino acid residues 1–17 in the final alignment of residues 1–62 (Fig. 3). Automated sequencing of $\text{H1}_{\alpha\beta}$ -T2 gave a blank in cycle 1, but sequenced cleanly from cycles 2 to 13. Manual dansyl analysis of this peptide identified Asx as its N -terminus. This residue is a potential site of N -linked oligosaccharide, since PTH-Thr was identified in cycle 3 of the Edman degradation. This peptide corresponds to residues 18–30 in the final alignment (Fig. 3). Peptide $\text{H1}_{\alpha\beta}$ -T3 sequenced cleanly and corresponded to residues 31–39 in the final sequence (Fig. 3). These three peptides, $\text{H1}_{\alpha\beta}$ -T1 to $\text{H1}_{\alpha\beta}$ -T3, were not analysed for glucosamine.

When the N -terminal peptide $\text{H1}_{\alpha\beta}$ -T1 was further cleaved with endoproteinase Asp-N and the products fractionated by h.p.l.c., peptides $\text{H1}_{\alpha\beta}$ -T1-Ala, $\text{H1}_{\alpha\beta}$ -T1-Alb and $\text{H1}_{\alpha\beta}$ -T1-A2 were recovered. The first two of these had blocked N -termini and identical amino acid compositions corresponding to amino acid residues 1–9 in the final aligned sequence (Fig. 3). However, whereas $\text{H1}_{\alpha\beta}$ -T1-Ala contained either one or two residues of glucosamine, $\text{H1}_{\alpha\beta}$ -T1-Alb contained none (Table 3). Peptide

Table 2. Automated sequence analyses of GSIV fragments and peptides

Edman cycle no.	PA (defined in Fig. 3)	H ₂ αβ*																
		PTH yield (pmol)		H1 _{αβ} -T2		H1 _{αβ} -T3		H1 _{αβ} -T1-A2		H1 _α -C4		H1 _β -C4		H1 _β -C2		H1 _β -C1-TH2		
		Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	Residue identified	PTH yield (pmol)	
1	N	270	G	360	-	-	I	810	D	240	S	1390	S	1550	T	540	V	61
2	T	340	N	180	-	2550	P	680	F	290	L	1390	L	1910	Y	1080	N	45
3	V	500	P	75	-	1220	G	460	W	100	K	565	K	981	P	474	F	36
4	N	130	V	46	-	2920	A	600	S	44	-	-	-	-	D	421	-	-
5	F	540	R	38	-	1440	L	340	Y	57	G	623	G	888	F	279	-	-
6	T	340	S	14	-	902	Q	270	S	17	T	417	T	687	W	116	-	-
7	Y	370	S	19	-	1180	L	240	L	11	E	359	E	535	-	-	-	-
8	P	420	A	20	-	901	T	100	T	1	I	444	I	572	-	-	-	-
9	D	240	G	13	-	644	K	110	-	-	T	223	T	339	-	-	-	-
10	F	350	Q	27	680	974	-	-	-	-	F	217	F	259	-	-	-	-
11	W	240	A	13	540	387	-	-	-	-	-	-	-	-	-	-	-	-
12	S	100	S	10	500	174	-	-	-	-	-	-	-	-	-	-	-	-
13	Y	200	Y	20	580	104	-	-	-	-	-	-	-	-	-	-	-	-
14	S	110	S	-	140	-	-	-	-	-	-	-	-	-	-	-	-	-
15	-	-	E	13	240	-	-	-	-	-	-	-	-	-	-	-	-	-
16	-	-	P	7	300	-	-	-	-	-	-	-	-	-	-	-	-	-
17	-	-	V	3	200	-	-	-	-	-	-	-	-	-	-	-	-	-
18	-	-	F	-	66	-	-	-	-	-	-	-	-	-	-	-	-	-
19	-	-	L	-	23	-	-	-	-	-	-	-	-	-	-	-	-	-

* Peptide H₂αβ* was subjected to automated sequence analysis twice. In run i, using 10% of the total sample, 17 cycles (except for cycle 14) were identified. In run ii, using the remainder of the sample, cycles 1-9 were lost, but cycles 10-19 identified, verifying and extending the sequence of run i.

Fig. 3. Summary of evidence for glycosylation sites and amino acid sequence of residues 1-62 of GSIV

The blocked *N*-terminus of both the α - and the β -subunit is pyrrolidone carboxylate. The evidence indicates that the oligosaccharide at Asn-5 is present only on the α -subunit and that at Asn-18 is present on both subunits. No evidence for differences in the amino acid sequences of the two subunits was obtained. The nomenclature for fragments and peptides generated by chemical or proteolytic cleavage are: PA, whole GSIV digested with pyroglutamate aminopeptidase; H, hydroxylamine; T, tryptic; C, chymotryptic; TH, thermolytic (i.e. thermolysin-cleaved); A, endoproteinase Asp-N. The notations for sequence determination are: $\rightarrow$, dansyl determination; $\blacktriangleright$, automated Edman degradation.

H1 $_{\alpha\beta}$ -T1-A2 had a composition corresponding to residues 10-17 and sequenced cleanly to its end (Fig. 3 and Table 3).

Since these results suggested that the differences in the α - and β -subunits were due to a difference in the presence of a carbohydrate moiety within residues 1-9, the separated H1 $_{\alpha}$ and H1 $_{\beta}$ fragments were digested separately with chymotrypsin. Resulting fragments designated H1 $_{\alpha}$ -C1 to H1 $_{\alpha}$ -C5 and H1 $_{\beta}$ -C1 to H1 $_{\beta}$ -C5 were isolated by high-voltage paper electrophoresis (pH 6.5) and h.p.l.c. Their amino acid and amino sugar analyses and sequence data are presented in Tables 2 and 3. Their alignment in the final sequence is shown in Fig. 3. Significantly, peptides H1 $_{\alpha}$ -C1 and H1 $_{\beta}$ -C1 were identical in that their *N*-termini were blocked and their amino acid compositions were the same and consistent with residues 1-6 in the aligned sequence. However, H1 $_{\alpha}$ -C1 contained two residues of glucosamine, whereas H1 $_{\beta}$ -C1 had none. This result is consistent with the presence of *N*-linked oligosaccharide at Asn-5 in the α -subunit and its absence in the β -subunit. In the case of peptides H1 $_{\alpha}$ -C4 and H1 $_{\beta}$ -C4, which had identical compositions and sequences and which provided overlaps for the previously sequenced H1 $_{\alpha\beta}$ -T1-A2 and H1 $_{\alpha\beta}$ -T2 (Fig. 3), two molecules of glucosamine were found in both peptides (Table 3). No PTH derivative was recovered in either peptide at cycle 4 (Table 2) corresponding to Asn-18 in the final sequence (Fig. 3). These data are consistent with the conclusion that *N*-linked

oligosaccharide is present at Asn-18 in both the α - and the β -subunits. Additional evidence for the absence of *N*-linked oligosaccharide at Asn-5 of the β -subunit was obtained from analysis of the products of thermolytic digestion of peptide H1 $_{\beta}$ -C1. Two peptides were recovered, H1 $_{\beta}$ -C1-TH1 and H1 $_{\beta}$ -C1-TH2, corresponding in composition to residues 1-3 and 4-6 respectively of the final sequence. Automated sequence analysis of the latter peptide (Table 2) identified Asn at cycle 2, confirming that the β -subunit is lacking carbohydrate at this position. Evidence for the overlapping of peptide H1 $_{\alpha\beta}$ -T3 and fragment H2 $_{\alpha\beta}$ was obtained by the isolation and sequence analysis of a chymotryptic peptide, cC6, derived from whole lectin. It spanned residues 36-56, completing the sequence of the *N*-terminus of GSIV through to residue 62.

Since the above data provided evidence for three oligosaccharide-attachment sites in whole GSIV, it was of considerable interest to re-examine the total neutral-sugar content of the lectin. By using the procedures described in the Materials and methods section, these were estimated to be 2.84 ± 0.05 for Xyl, 2.01 ± 0.09 for Fuc and 10.63 ± 0.37 for Man. Therefore, assuming that the acid hydrolyses were complete, a carbohydrate content of (Xyl) $_3$ (Fuc) $_2$ (Man) $_{10}$ (GlcNAc) $_6$ appears minimal. This carbohydrate content corresponds to 6.4% of the GSIV dimer as compared with the previously reported 3% [1] and is in good

Table 3. (cont.)

Nature of cleavage	Peptide	Yield (%)	M_r	Mobility (pH 6.5)	Net charge	Position of elution on C ₁₈ column (% acetonitrile) [¶]	Position in sequence	Amino acid composition† or sequence‡	Determination of acid or amide†
Chymotryptic digestion of H1 _α	H1 _α -C1	35	723§	-0.2	-1	22.1	1-6	(Z ₁ , B ₂ , T ₁ , V ₁ , F ₁ , GlcNAc) 1.0 2.0 1.2 0.7 1.0 2.0	HVE
	H1 _α -C2	19	829	-0.29	-1	34.0	7-12	(T ₁ , Y ₁ , P ₁ , B ₁ , F ₁ , W ₁) 1.1 1.0 1.0 0.8 1.1 -	HVE
	H1 _α -C3	85	268	0	0	11.7	13-14	(S ₁ , Y ₁) 0.8 1.2	-
	H1 _α -C4	53	1108§	0	0	26.3	15-24	S-L-K-N-G-T-E-I-T-F (GlcNAc) 1.0 1.0 1.1 1.0 0.9 1.0 1.1 1.0 1.0 2.0	HVE, AS
	H1 _α -C5	38	1082	0	0	27.9	25-35	(B ₁ , G ₂ , T ₁ , R ₁ , A ₂ , P ₁ , I ₁ , L ₂) 1.0 1.6 0.9 1.1 1.6 1.0 1.2 2.4	HVE
Chymotryptic digestion of H1 _β	H1 _β -C1	12	723	-0.35	-1	24.1	1-6	(Z ₁ , B ₂ , T ₁ , V ₁ , F ₁) 1.0 1.6 1.1 1.1 1.1	HVE
	H1 _β -C2	30	829	-0.29	-1	34.0	7-12	T-Y-P-D-F-W 1.1 1.1 1.0 0.9 1.0 -	HVE, AS
	H1 _β -C3	69	268	0	0	11.8	13-14	(S ₁ , Y ₁) 0.9 1.1	-
	H1 _β -C4	61	1108§	0	0	26.3	15-24	S-L-K-N-G-T-E-I-T-F (GlcNAc) 1.0 1.0 1.0 1.0 1.0 0.9 0.9 2.0	HVE, AS
	H1 _β -C5	24	1082	0	0	27.8	25-35	(B ₁ , G ₂ , T ₁ , R ₁ , A ₂ , P ₁ , I ₁ , L ₂) 1.0 2.0 1.0 1.2 2.0 0.9 1.0 2.0	HVE
Peptide from chymotryptic digestion of citraconylated GSIV	cC6	19	2163	+0.12	+1	21.9	36-56	Q-L-T-K-T-D-A-N-G-N-P-V 1.0 1.2 1.0 1.2 1.0 0.9 1.0 1.2 1.0 0.8 0.8 R-S-S-A-G-Q-A-S-Y 0.8 1.0 1.0 0.9 1.2 1.0 0.9 1.0 1.0	HVE, AS

* The yields are based on nmol of peptides recovered with respect to nmol of starting material. Lower yields for some peptides were due either to incomplete cleavage or to further purification at pH 1.8 and h.p.l.c.

† The amino acid analyses are molar ratios.

‡ the notation for sequence determination is: →, dansyl determination; →, automated Edman degradation; →, not identified during sequence analysis, but known from amino acid composition and/or analogous peptides. Peptides from the N-terminus were blocked. Determination of amides was by automated sequencing (AS) or from calculated net charge from electrophoretic mobilities (HVE).

§ The contribution of the carbohydrate moiety was not included in the determination of M_r .

|| Electrophoretic mobilities are relative to aspartic acid at pH 6.5. '0' indicates that the peptides are neutral at pH 6.5.

¶ The peptide eluted off the paper was subsequently purified by h.p.l.c. using a Vydac 218TP54 C₁₈ reverse-phase column.

agreement with three sites of oligosaccharide attachment as established in the present work.

The evidence detailed above shows that the molecular-mass difference in the two subunits of GSIV is accounted for by a difference in the oligosaccharide content of the α - and β -chains. Thus, whereas the α -subunit has *N*-linked oligosaccharide at both Asn-5 and Asn-18, the β -subunit has covalently bound carbohydrate only at Asn-18. The observation that treatment with *N*-glycanase leads to a decrease in the apparent molecular mass of the α -band, which then has a mobility on SDS/PAGE identical with that of the β -band, indicates that only the *N*-linked oligosaccharide at Asn-5 is cleaved by this enzyme. This difference in susceptibility to enzymic hydrolysis between the Asn-5 and Asn-18-linked oligosaccharides suggests a difference in their structures, a possibility further supported by the finding of only two Fuc units per molecule of GSIV. These oligosaccharide structures require further investigation, but preliminary results suggest that the fucose units are present on *N*-acetylated chitobiose units at Asn residues at position 18 of both subunits.

The elucidation of the complete amino acid sequence of GSIV and the covalent structures of the oligosaccharides is of obvious importance for the interpretation and refinement of the electron-density data generated in the X-ray-diffraction analyses of GSIV. In particular these studies [18] have recently shown that the organization of the dimers in the GSIV molecule is different from that in other lectins such as concanavalin A [19–21], favin [22] and pea lectin [23]. In the present case it appears that the *N*-linked oligosaccharide of Asn-5, and perhaps of Asn-18, resides at least partially between the subunits and that this carbohydrate is involved at the interactive face of the monomers. The complete covalent structure of this interesting molecule awaits further elucidation.

We are indebted to Mr. M. Nattriss for many amino acid and glucosamine analyses. This study was supported by grants from the Medical Research Council of Canada to L.B.S. and from the National Sciences and Engineering Research Council of Canada to R.U.L.

REFERENCES

1. Shibata, S., Goldstein, I. J. & Baker, D. A. (1982) *J. Biol. Chem.* **257**, 9324–9329
2. Spohr, U., Hindsgaul, O. & Lemieux, R. U. (1985) *Can. J. Chem.* **63**, 2644–2652
3. Spohr, U. & Lemieux, R. U. (1988) *Carbohydr. Res.* **174**, 211–237
4. Vandonselaar, M., Delbaere, L. T. J., Spohr, U. & Lemieux, R. U. (1987) *J. Biol. Chem.* **262**, 10848–10849
5. Lotan, R., Debray, H., Cacan, M., Cacan, R. & Sharon, N. (1975) *J. Biol. Chem.* **250**, 1955–1957
6. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
7. Konat, G., Offner, H. & Mellah, J. (1984) *Experientia* **40**, 303–304
8. Albersheim, P., Nevins, D. J., English, P. D. & Karr, A. (1967) *Carbohydr. Res.* **5**, 640–645
9. Rose, M. E. & Johnstone, R. A. W. (1982) *Mass Spectrometry for Chemist and Biochemist*, pp. 98–100, Cambridge University Press, Cambridge
10. Message, G. M. (1984) *Practical Aspects of Gas Chromatography/Mass Spectroscopy*, John Wiley and Sons, New York
11. Pearlstone, J. R., Carpenter, M. R. & Smillie, L. B. (1977) *J. Biol. Chem.* **252**, 971–977
12. Pearlstone, J. R., Carpenter, M. R. & Smillie, L. B. (1986) *J. Biol. Chem.* **261**, 16795–16810
13. Ryle, A. P., Sanger, F., Smith, L. F. & Kitai, R. (1955) *Biochem. J.* **60**, 541–556
14. Simpson, R. J., Neuberger, M. R. & Lin, T.-Y. (1976) *J. Biol. Chem.* **251**, 1936–1940
15. Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475
16. Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822
17. Offord, R. E. (1966) *Nature (London)* **211**, 591–593
18. Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Nikrad, P. V., Pearlstone, J. R., Carpenter, M. R., Smillie, L. B., Spohr, U. & Lemieux, R. U. (1989) *Trans. Am. Cryst. Assoc.*, in the press
19. Hardman, K. D. & Ainsworth, C. F. (1972) *Biochemistry* **11**, 4910–4919
20. Becker, J. W., Reeke, G. N., Jr., Wang, J. L., Cunningham, B. A. & Edelman, G. M. (1975) *J. Biol. Chem.* **250**, 1513–1524
21. Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb (Gilboa), A. J., Dodson, E. J., Papiz, M. Z., Wan, T. & Campbell, J. (1989) *EMBO J.* **8**, 2189–2193
22. Reeke, G. N., Jr., and Beeker, J. W. (1986) *Science* **234**, 1108–1111
23. Einspahr, H., Parks, E. H., Seguna, K., Subramanian, E. & Suddath, F. L. (1986) *J. Biol. Chem.* **261**, 16518–16527

Received 1 June 1990; accepted 5 July 1990