Carbohydrate Research, 40 (1975) 171–182

IMMUNOCHEMICAL AND CHEMICAL INVESTIGATIONS OF THE STRUCTURE OF GLYCOPROTEIN FRAGMENTS OBTAINED FROM EPIGLYCANIN, A GLYCOPROTEIN AT THE SURFACE OF THE TA3-Ha CANCER CELL*†

JOHN F. CODINGTON, KEYES B. LINSLEY, ROGER W. JEANLOZ§,

Laboratory for Carbohydrate Research, Departments of Biological Chemistry and Medicine, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114 (U. S. A.)

TATSURO IRIMURA, AND TOSHIAKI OSAWA

Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113 (Japan)

(Received December 9th, 1974, accepted December 18th, 1974)

ABSTRACT

The structures of the carbohydrate chains present in fragments of a large-molecular-weight glycoprotein, epiglycanin, cleaved from the surface of viable TA3-Ha murine mammary carcinoma ascites cells and purified by gel filtration, were studied by immunochemical and chemical methods. Inhibitory activities for neuraminidase-treated and untreated glycoprotein material in the hemagglutination of NN-specific human erythrocytes by eight purified lectins were determined. Excellent inhibition was obtained in the *Bauhinia purpurea*, *Arachis hypogaea*, *Iberis amara*, and *Wistaria floribunda* systems, and weak inhibition against the *Ricinus communis* and *Glycine max* lectins. No activity against hemagglutination by the *Phaseolus vulgaris* and *Phaseolus limensis* lectins was observed. These results, when compared with those obtained by periodate oxidation, alkaline borohydride reduction, and partial methylation, suggest the possible presence of six different carbohydrate chains of 1 to 5 components in length, having as terminal groups *N*-acetylneuraminic acid, galactose, and 2-acetamido-2-deoxygalactose. All chains are attached to a single polypeptide chain by *O*-glycosyl bonds involving a 2-acetamido-2-deoxygalactose residue and a serine or threonine residue. It is suggested that the native molecule of epiglycanin of molecular weight 500,000 contains more than 500 carbohydrate chains attached to a single polypeptide chain of ~1,300 amino acid units.

*Dedicated to Professor Michael Heidelberger, in honor of his 87th birthday.

†This is publication No. 659 of the Robert W. Lovett Memorial Group for the Study of Diseases Causing Deformities, Harvard Medical School and Massachusetts General Hospital. This investigation was supported by research grants (to R.W.J.) from the National Cancer Institute (CA-08418), National Institutes of Health, U. S. Public Health Service, and from the National Science Foundation (the Japan–U. S. Cooperative Science Program GF-39782), and by a research grant (to T.O.) from the Japan Society for the Promotion of Science (the Japan–U. S. Cooperative Science Program 5RO50).

§To whom communications should be addressed.

INTRODUCTION

Current interest in glycoproteins at the surface of tumor or transformed cells is based upon evidence that such molecules may play vital roles in cellular function[1,2], and that certain fundamental differences between tumor and normal cells may be expressed in differences in the structure of cell-surface glycoproteins[3,4]. For these reasons, an investigation of the chemical and biological properties of a large, mucin-type glycoprotein (epiglycanin) at the surface of the murine TA3-Ha mammary carcinoma ascites cell was undertaken. This glycoprotein appears to be unique to the ascites subline (TA3-Ha) of the tumor that had lost strain-specificity[5], as it could not be detected in the subline (TA3-St) that had retained specificity[6]. Some physico-chemical[7–9] and biological[5,10] properties of this glycoprotein have been reported.

Inhibition of hemagglutination by lectins or by antisera can be accomplished by glycoproteins bearing the appropriate, carbohydrate-haptenic groups[11,12]. The major glycoprotein of the erythrocyte surface, which possesses a number of different receptor sites[12–14], is an effective inhibitor of hemagglutination by a large number of lectins, including that from the seeds of *Vicia graminea*[15]. Epiglycanin, however, was found to be as effective in inhibiting hemagglutination by the *Vicia graminea* lectin as the most active fragments from the erythrocyte glycoprotein[16]. This finding suggests that both glycoproteins contain some identical carbohydrate structures that are presumably similar to the disaccharide chain proposed by Uhlenbruck and Dahr[17] as the active receptor site for the *Vicia graminea* lectin.

As part of an investigation of the structure of epiglycanin, we have extended the study of hemagglutination inhibition to other lectins. In this study, three glycoprotein fragments isolated from the TA3-Ha ascites cell[7,8] have been studied as inhibitors of agglutination of human erythrocytes by eight lectins of various specificities. In addition, periodate oxidation–alkaline borohydride reduction–partial hydrolysis and partial methylation methods were applied to epiglycanin in an attempt to elucidate the structures of the carbohydrate chains.

EXPERIMENTAL

Glycoprotein fractions. — Glycoprotein Fraction-I (GPF-I), eluted in the void volume from a Bio-Gel P-100 column[9], represented the total amount of the native glycoprotein material (epiglycanin) released from viable cells by incubation with TPCK-trypsin (18 μg per ml at 4°), 1 to 2 mg per 10^9 cells. Fractionation of GPF-I on a column of Sepharose 4B resin gave Fractions A, B, and C.

Sialic acid-free glycoproteins were prepared by incubation of a sample of glycoprotein (100–150 μg) in balanced salt solution (200 μl) with neuraminidase (5 units, from *Vibrio cholerae*, 500 units/ml, Behring Diagnostics, Somerville, N.J. 08876) for 18 h at 24°. The solutions were lyophilized.

Lectins and antisera. — The hemagglutinating lectins of *Bauhinia purpurea* and *Iberis amara* seeds were purified according to the procedure of Irimura and Osawa[18].

The *Arachis hypogaea* lectin was extracted from peanuts, commercially obtained in Japan, and was purified according to the procedure of Terao and Osawa[19]. The hemagglutinating lectin of *Wistaria floribunda* was separated from the mitogen by the procedure of Toyoshima *et al.*[20]. The *Ricinus communis* lectin was obtained from commercial castor beans, and purified by the method of Tomita *et al.*[21]. The *Glycine max* lectin was purified according to the procedure of Gordon *et al.*[22], and that of *Phaseolus limensis*, according to the procedure of Gould and Scheinberg[23]. The lectin of *Phaseolus vulgaris* was a product of Difco Laboratories, Detroit, Mich. 48232. Rabbit anti-M and anti N sera were prepared according to the procedure of Race and Sanger[24].

Hemagglutination inhibition assays. — The titration and inhibition assays were performed with human erythrocytes freshly obtained from a donor, according to the methods previously described[25]. The cells used for the inhibition assays on *B. purpurea*, *I. amara*, *G. max*, and *A. hypogaea* lectins were neuraminidase-treated cells. The neuraminidase treatment of human erythrocytes was performed as described previously[18].

Gas–liquid chromatography. — Three different procedures were employed in the analysis of carbohydrate components.

Method A. The samples were analyzed by the method of Reinhold[26].

Method B. In this method, a modification of Method A, more-vigorous methanolysis conditions, with M hydrogen chloride in anhydrous methanol for 20 h at 85°, were used.

Method C. The samples were heated with 2M aqueous trifluoroacetic acid for 2 h at 120°. The mixture was dried, acetylated as in Method A, and heated with 0.5M hydrogen chloride in anhydrous methanol for 16 h at 65°. The carbohydrate components were converted into the per-*O*-(trimethylsilyl)ated methyl glycosides and analyzed as in Method A.

The carbohydrate compositions of Fractions A, B, and C were determined by Method B. Amino acid analyses were performed by the method of Roach and Gehrke[27].

Alkaline borohydride reduction. — *Method A.* A solution (9.5 ml) of GPF-I (0.71 mg) containing 0.3M sodium borohydride and 0.1M sodium hydroxide was kept for 5 days at 24°. After evaporation, the residue was dissolved in 6M hydrochloric acid, and the solution flushed with nitrogen and heated for 20 h at 110°. The amino acids were determined with a Beckman Automatic Amino Acid Analyzer. An analysis of GPF-I that had not been treated with alkaline borohydride was performed concurrently.

Method B. A solution (0.4 ml) of Fraction C (0.24 mg) in 2.0M sodium borohydride and 0.1M sodium hydroxide was heated for 19 h at 45°. The mixture was made neutral with hydrochloric acid, and filtered through a calibrated column (2.5 × 53 cm) of Bio-Gel P-2 resin. Two fractions, (*a*) 90–126 ml, and (*b*) 150–220 ml, were lyophilized. Fraction *b* was passed through a column of AG-50W X8 (200–400 mesh,

H^+) cation-exchange resin to remove the salts. Both Fractions *a* and *b* were analyzed by gas–liquid chromatography (Method B).

Periodate oxidation. — To a sample of GPF-I (185 μg) in a balanced salt solution[28] (0.40 ml) was added *Vibrio cholerae* neuraminidase (25 μl, 12.5 units). An identical sample of GPF-I was treated in parallel, but without addition of neuraminidase. After incubation for 16 h at 37°, each sample was treated with sodium metaperiodate (0.5 mg) in water (0.50 ml), and the solution was kept in the dark for 21 h at 24°. To remove the excess of periodate, 0.10 ml of 1,2-ethanediol was added, and, after 3 h at 24°, the solution was dialyzed against water for 20 h at 4°. The nondialyzable solution was lyophilized, and the residue was analyzed for carbohydrate components by gas–liquid chromatography (g.l.c.; Method B).

Partial methanolysis. — A solution of GPF-I (0.21 mg) in 0.5M hydrogen chloride in anhydrous methanol (1.0 ml) was heated in a sealed tube for 60 min at 60°. The solvent was removed by passage of a stream of dry nitrogen, and the residue was dried *in vacuo*. Trimethylsilylation was performed by the addition of a mixture of hexamethyldisilazane, chlorotrimethylsilane, and pyridine (100 μl) (Sylon HTP, Supelco, Inc., Bellefonte, Pa. 16823). After 60 min at 24°, all liquid was removed with a stream of nitrogen. The residue was dissolved in hexane, and injected into a gas chromatograph equipped with a stainless-steel column containing 0.1% of OV-17 on GLC-110 (Applied Science Lab. Inc., State College, Pa. 16802).

RESULTS

Hemagglutination inhibitory activities of Fractions A, B, and C with eight lectins and with antisera to blood-group substances M and N are presented in Table I. Inhibition by *Bauhinia purpurea*, *Iberis amara*, *Arachis hypogaea*, and *Wistaria floribunda* was excellent. The three glycoprotein fractions exhibited marked differences in activity, Fraction A being the most active in the *Arachis hypogaea* system, and Fraction C the least effective with the *Bauhinia purpurea* hemagglutinin. Fraction B was more active than either Fraction A or Fraction C against *Bauhinia purpurea*. By contrast, very little difference was observed between Fractions A, B, or C in the inhibition of agglutination by *Wistaria floribunda*.

Very weak, yet significant, inhibitory activity was exhibited against *Ricinus communis* and *Glycine max* hemagglutinins by some fractions, but no detectable activity was observed in the inhibition of agglutination by the two other lectins tested, namely, *Phaseolus vulgaris* and *Phaseolus limensis*, the blood-group A specific lectin. As expected[16], the three fractions exhibited no activity against antisera to blood-group specificities M or N. In addition, no inhibitory activity was observed for GPF-I against antisera to blood-group substances A or B. Removal of *N*-acetylneuraminic acid from the penultimate D-galactopyranosyl residues in the glycoprotein fractions generally resulted in an increase in activity with lectins (see Table I).

The compositions of Fractions A, B, and C (see Table II) demonstrate significant differences for the carbohydrate components of the three fractions. In Fraction C,

TABLE I

HEMAGGLUTINATION INHIBITION ACTIVITIES OF PURIFIED GLYCOPROTEIN FRACTIONS ISOLATED FROM THE TA3-Ha MAMMARY CARCINOMA ASCITES CELL BEFORE AND AFTER TREATMENT WITH NEURAMINIDASE

Lectin	*Hemagglutination inhibition activity*[a]					
	Fraction A		*Fraction B*		*Fraction C*	
	Nontreated	*Treated*	*Nontreated*	*Treated*	*Nontreated*	*Treated*
Bauhinia purpurea[b]	28	28	18	13	73	73
Iberis amara[b]			4	4		
Arachis hypogaea (Jap.)[b]	2	2	3	1	3	5
Wistaria floribunda (Fr. A)	55	28	56	18	76	36
Glycine max[b]	>440	>440	320	160	290	290
Ricinus communis	>440	220	320	80	>580	>580
Phaseolus limensis (Anti-A)			>125	>125		
Phaseolus vulgaris	>440	>440	>640	>640	>580	>580
Absorbed rabbit Anti-M			>125	>125		
Absorbed rabbit Anti-N			>125	>125		

[a]The values are expressed as μg of inhibitor per ml required to give complete inhibition of agglutination in 30 min. [b]Tested against neuraminidase-treated, human erythrocytes.

TABLE II

CARBOHYDRATE COMPOSITIONS[a] AND MOLECULAR WEIGHTS[b] OF GLYCOPROTEIN FRACTIONS OBTAINED FROM THE SURFACE OF THE TA3-Ha ASCITES CELL

Carbohydrate component	*Fraction A*		*Fraction B*		*Fraction C*	
	%	*Rel. mol. pr.*[c]	%	*Rel. mol. pr.*[c]	%	*Rel. mol. pr.*[c]
Gal	24	4.1	27	5.0	30.5	5.2
GalNAc	18	2.4	24	3.5	28	3.7
GlcNAc	7.5	1.0	6.8	1.0	7.4	1.0
NeuNAc	14	1.3	12.5	1.3	7.3	0.7
Man	0.4	0.07	0.4	0.07	0.4	0.07
Fuc	0		0		0	
Total (%)	63		71		73	
Mol. wt. ($\times 10^{-4}$)	46		22		10	

[a]Carbohydrate compositions, which have been reported, in part, previously[8], were determined by gas–liquid chromatography[26]. [b]Molecular weights were determined by sedimentation equilibrium, as previously reported[7]. [c]Molar proportion relative to 2-acetamido-2-deoxy-D-glucose as unity.

the relative proportions of galactose and 2-acetamido-2-deoxygalactose residues are markedly higher than those in Fraction A, but the proportion of sialic acid is considerably lower. All fractions contain the same four major carbohydrate components and the same, perhaps insignificant, proportion of mannose. As previously reported[8,9], the three fractions contain similar amino acid compositions.

As shown in Table III, alkaline borohydride reduction (Method A) resulted in the significant loss (73%) of only two amino acids, serine and threonine. Recovery of

TABLE III

AMINO ACID COMPOSITION OF GLYCOPROTEIN FRACTION I BEFORE AND AFTER ALKALINE BOROHYDRIDE TREATMENT[a]

Amino acid	*Untreated material*	*Treated material*	*Net change*
Aspartic acid	9	7	−2
Threonine	367	83	−284
Serine	281	81	−200
Glutamic acid	13	19	+6
Proline	114	114	0
Glycine	84	80	−4
Alanine	106	162	+56
2-Aminobutyric acid	0	105	+105
Valine	7	8	+1
Isoleucine	3	17	+14
Leucine	18	19	+1

[a]Residues per 1,000 residues.

the reduction products from the alkali-induced, β-elimination reaction was 28% for alanine (from serine) and 35% for 2-aminobutyric acid (from threonine). The carbohydrate components in this reaction mixture included galactitol, in addition to 2-acetamido-2-deoxygalactitol; no 2-acetamido-2-deoxygalactose was detected. At a higher concentration of sodium borohydride (Method B), 2-acetamido-2-deoxygalactitol was again detected (37% yield), but no galactitol. This result suggests that the reduction of galactose, which occurred in Method A, resulted from a "peeling" reaction due to the high concentration of alkali, and that galactosyl groups are not involved in glycopeptide linkages in epiglycanin. It is noteworthy that the number of serine and threonine residues, which would have been destroyed in a molecule of 500,000 molecular weight, namely, 540 (see Table III), is similar to the total number of 2-acetamido-2-deoxygalactose residue groups (590) that are present, based upon a content of 24% (see Table II). The small proportion of aspartic acid residues, less than one residue per 100 amino acid residues, and of mannose residues, approximately one residue per 60 galactose residues, present in epiglycanin suggests that asparagine-linked carbohydrate chains, if present at all, exist in extremely minor proportions.

The *N*-acetylneuraminic acid residues of GPF-I appear to be completely released by neuraminidase, as the amount of this component liberated by the enzyme

is similar to that cleaved by dilute acid (0.05M sulfuric acid for 60 min at 80°) and determined by the thiobarbituric acid method[29] and to that determined by g.l.c. (Methods A or B). The presence of *N*-glycolylneuraminic acid was not detected in GPF-I, although it was present in small proportion (10%) in the total sialic acid released from viable cells by neuraminidase[30].

The results of periodate oxidation of GPF-I are presented in Table IV. G.l.c. by Method C, used for this determination, does not show the sialic acid content. Based upon the amount of 2-acetamido-2-deoxyglucose estimated, the recovery of macromolecular material after dialysis was 34% for the neuraminidase-treated GPF-I, and 58% for the material which had not been treated with the enzyme. Periodate oxidation destroyed 63% of the galactose residues in the untreated material, but, after removal of *N*-acetylneuraminic acid, the galactose residues were almost completely eliminated (92%). More 2-acetamido-2-deoxygalactose residues (9%) were oxidized after removal of the sialic acid groups than with this component present in the terminal position (5%).

TABLE IV

PERIODATE OXIDATION OF GLYCOPROTEIN FRACTION[a] I

Glycoprotein Fraction I	*Carbohydrate components*			
	Gal	*GalNAc*	*GlcNAc*	*NeuNAc*
Unoxidized GPF-I	3.8	2.2	1.0	1.0
Untreated with neuraminidase				
Recovered (%)[b]	21	55	58	
(Molar proportion)	1.4	2.1	1.0	
Oxidized (%)[c]	63	5	0	
(Molar proportion)	2.4	0.1	0.0	
Treated with neuraminidase				
Recovered (%)[b]	2	29	34	
(Molar proportion)	0.3	2.0	1.0	
Oxidized (%)[c]	92	9	0	
(Molar proportion)	3.5	0.2	0.0	

[a]The samples were oxidized and dialyzed, and the undialyzable products were analyzed by g.l.c. (Method C). [b]Based upon a theoretical recovery of untreated and unoxidized GPF-I of 100%. [c]Calculations based upon the assumption that GlcNAc was not oxidized.

The results of limited methanolysis of GPF-I are presented in Table V. Values for the methyl glycosides formed in this reaction are compared with values obtained after complete hydrolysis of the carbohydrate moiety. As expected, *N*-acetylneuraminic acid residues were completely cleaved off (95%), and only a small proportion of the galactose residues (15%) remained bound. A large proportion (72%) of the 2-acetamido-2-deoxyglucose residues was released, but no 2-acetamido-2-deoxygalactose residues were detected.

TABLE V

PARTIAL METHANOLYSIS[a] OF GLYCOPROTEIN FRACTION[b] I

Carbohydrate component	*Hydrolysis* (%)	*Partial methanolysis* (%)	*Proportion of component cleaved* (%)
Galactose	26	22	85
2-Acetamido-2-deoxygalactose	19	0	0
2-Acetamido-2-deoxyglucose	8.5	6	72
N-Acetylneuraminic acid	12[c]	11.5	95

[a]Samples were heated with 0.5M hydrogen chloride in anhydrous methanol for 60 min at 65°. [b]Values determined by g.l.c. (Method C). [c]Value determined by the thiobarbituric acid method[29] after cleavage with neuraminidase or dilute sulfuric acid.

DISCUSSION

The chemical structure of the carbohydrate chains of epiglycanin, a high-molecular-weight glycoprotein present at the surface of the rapidly proliferating, non-strain-specific Ha subline of the TA3 mammary carcinoma ascites cell of the strain A mouse, but absent from the slower-growing, strain-specific St subline[6], was investigated by chemical and hemagglutination inhibition methods. The use of various methods for this investigation was necessitated by the microheterogeneity of glycoprotein chains in mucin-type material and by the difficulty of collecting sufficient amounts of material to permit investigation of minor structures by chemical means. All structural studies were performed on fractions that had been cleaved from the surfaces of viable cells by the action of TPCK-trypsin, and purified by gel filtration[7,9]. Glycoprotein Fraction I (a fraction excluded from a Bio-Gel P-100 column[9]) was studied by chemical methods, whereas Fractions A, B, and C (obtained by fractionation[7] of GPF-I) were studied by immunochemical methods. Fractions A, B, and C showed similar amino acid compositions[8,9] and identical carbohydrate components (see Table II), but the latter were present in significantly different proportions.

Fractions A, B, and C showed the strongest inhibition against hemagglutination of red blood-cells (see Table I) by *Bauhinia purpurea*, *Iberis amara*, *Arachis hypogaea*, and *Wistaria floribunda*. These lectins are specific for structures that include a carbohydrate chain attached to a peptide by an *O*-glycosyl linkage involving an *O*-(2-acetamido-2-deoxy-α-D-galactopyranosyl)-L-serine (or -L-threonine) group[12,17,31]. The activities of these lectins are probably closely related to that of the *Vicia graminea* lectin[12,17], which involves, as determinant, an *O*-β-D-galactopyranosyl-(1→3)-*O*-(2-acetamido-2-deoxy-α-D-galactopyranosyl)-L-serine or -L-threonine group, as proposed by Uhlenbruck and Dahr[17]. Epiglycanin has been shown to inhibit, strongly, the agglutination of red blood-cells by an extract of *Vicia graminea* seeds[10,16], and this lectin served for the detection of epiglycanin in the body fluids of mice bearing the TA3-Ha ascites tumor[10]. The observation of variations in the level of inhibitory

activity exhibited by Fractions A, B, and C against hemagglutination by the four lectins offers substantial evidence not only that these lectins exhibit significantly different specificities, but that not one is identical in activity to the *Vicia graminea* lectin. These variations (see Table I) may be due to several different factors, the most obvious being the relative proportion and availability of active haptenic groups. Other important factors may include molecular size and shape of the fractions, and the presence of adjacent or nearby structures that could exert either enhancing or inhibiting effects.

The inhibitory activity against the hemagglutination of the *Wistaria floribunda* lectin was significantly enhanced by the removal of the terminal *N*-acetylneuraminic acid group of Fractions A, B, and C with neuraminidase (see Table I). Neuraminidase treatment produced an even greater enhancement of activity in the *Vicia graminea* system[33]. These results suggest the probable attachment of an *N*-acetylneuraminic acid group to a subterminal D-galactose residue, since D-galactose has been shown to be the terminal group of the *Vicia graminea* receptor[17]. This proposed structure is in agreement with numerous reports that, in glycoproteins, the terminal *N*-acetylneuraminic acid group is linked either to a *β*-D-galactopyranosyl or to a 2-acetamido-2-deoxy-D-galactopyranosyl residue[34].

Weak, yet significant, activity was observed in the inhibition of agglutination by the two lectins *Ricinus communis* (castor bean) and *Glycine max* (soybean). *Ricinus communis* has been reported to be specific for a 2-acetamido-2-deoxy-*O*-*β*-D-galactopyranosyl-D-glucopyranosyl determinant structure[12], and the activity against this lectin was markedly increased after removal of the *N*-acetylneuraminic acid groups, an observation that supports an *O*-(*N*-acetylneuraminosyl)-D-galactose sequence. It may be significant that Fraction C, which contains a smaller proportion of 2-acetamido-2-deoxy-D-glucose residues, had no detectable activity against the *Ricinus communis* lectin. *Glycine max* lectin has been reported to recognize either a 2-acetamido-2-deoxy-*O*-*β*-D-galactopyranosyl-D-glucose or -D-galactose sequence[31], but it is not clear whether the same haptenic sequence is involved in the *Glycine max* as in the *Ricinus communis* system.

The results of alkaline borohydride reduction established the 3-*O*-(2-acetamido-2-deoxy-*β*-D-galactopyranosyl)-L-serine (or -L-threonine) group as the major, if not the only, type of glycopeptide linkage in epiglycanin. The frequency of mucin-type carbohydrate chains suggested by these studies is, to the best of our knowledge, greater than in any other previously reported glycoprotein. 2-Acetamido-2-deoxy-D-galactitol, but no 2-acetamido-2-deoxy-D-galactose or D-galactitol, was detected after alkaline borohydride reduction under conditions where no "peeling effect" was expected. Further support for the suggestion that all 2-acetamido-2-deoxy-D-galactose residues are involved in the protein–carbohydrate linkages was furnished by the results of partial methanolysis (see Table V). No cleavage of this hexosamine component was observed, although major proportions of D-galactose and 2-acetamido-2-deoxy-D-glucose were liberated. Based upon the proportion of 2-acetamido-2-deoxy-D-galactose present in GPF-I, ~590 residues of this component are present in a

molecule of epiglycanin of mol. wt. 500,000 (a value estimated for the native molecule on the basis of physical measurements). A number of the same order of magnitude (510) was found for the L-serine and L-threonine residues degraded in the alkaline borohydride reduction. These values suggest a total of at least 500 carbohydrate chains linked to a single polypeptide chain[7] of ~1,300 amino acid residues, or an average of one carbohydrate chain attached to every second or third amino acid in the peptide backbone. Amino acid analyses indicated the presence of ~800 L-serine and L-threonine residues in a molecule of mol. wt. 500,000, which suggests that over 60% of the total number of hydroxyamino acid residues are involved in the glycopeptide linkages. The recovery of the L-serine and L-threonine components after alkaline sodium borohydride reduction as alanine and 2-aminobutyric acid was low, but the conditions used were not optimal for this recovery. The observation (see Table IV) that removal of the *N*-acetylneuraminic acid residues with neuraminidase increased, from 63 to 92%, the proportion of D-galactose destroyed by periodate oxidation indicates an *O*-(*N*-acetylneuraminosyl)-(2→3)-D-galactosyl terminal group and is consistent with the result obtained by immunochemical methods. Removal of the terminal D-galactose group in GPF-I with a β-D-galactosidase from *Clostridium perfringens*[35,36] established that this component is β-D-linked.

The chemical structures suggested on the basis of chemical or immunochemical evidence are presented in Fig. 1. A small proportion (5%) of the 2-acetamido-2-deoxy-D-galactose residues is destroyed by the periodate oxidation, which suggests the

Long-chain type

α-NeuNAc-(2→3)-β-D-Gal-(1→3 or 4)-D-GlcNAc-(1→2,4, or 6)-D-Gal-(1→3 or 4)-α-D-GalNAc-Ser(Thr)

1

β-D-Gal-(1→3 or 4)-D-GlcNAc-(1→2,4, or 6)-D-Gal-(1→3 or 4)-α-D-GalNAc-Ser(Thr)

2

Short-chain type

α-NeuNAc-(2→3)-β-D-Gal-(1→3 or 4)-α-D-GalNAc-Ser(Thr)

3

β-D-Gal-(1→3 or 4)-α-D-GalNAc-Ser(Thr)

4

α-D-GalNAc-Ser(Thr)

5

Fig. 1. Possible structures of carbohydrate chains in epiglycanin.

presence of chains composed of only one carbohydrate unit, as shown by **5**. A structure identical to that suggested for the one recognized by the *Vicia graminea* lectin[17] is presented as **4**, and with an additional, terminal *N*-acetylneuraminic acid group as structure **3**. The protection from periodate oxidation of ~95% of the 2-acetamido-2-deoxy-D-galactose residues suggests that the β-D-galactopyranosyl

groups are linked to the 2-acetamido-2-deoxy-α-D-galactopyranosyl residues by either a (1→3) or a (1→4) bond. The α-D configuration of the 2-acetamido-2-deoxy-D-galactopyranosyl residue shown in structures **1**–**5** is based solely on analogy with the structures of other mucin-type glycoproteins, and the pyranose ring, on the resistance to periodate oxidation and stability towards methanolysis.

The presence of two major types of carbohydrate chains in epiglycanin has been previously suggested[34]. Structures **3**, **4**, and **5** just discussed represent the short-chain type. Structures **1** and **2**, both containing 2-acetamido-2-deoxy-D-glucopyranose residues and composed of 5 and 4 carbohydrate units, respectively, represent the long-chain type, and are based on the composition of the fractions (see Table II). 2-Acetamido-2-deoxy-D-glucose generally constitutes 7 to 8% of GPF-I. Variations in the composition of GPF-I have occurred in certain preparations, and the molar proportion of *N*-acetylneuraminic acid has reached twice that of 2-acetamido-2-deoxy-D-glucose. These proportions suggest the presence of a chain containing both of these components (as in **1**). Results of the periodate oxidation experiments indicate that the terminal *N*-acetylneuraminic acid group in **1** is attached to a D-galactose residue by an α-D-(2→3) linkage, similar to that suggested for **3**. The results of hemagglutination inhibition studies with *Ricinus communis* and *Glycine max* lectins support the presence of the 2-acetamido-2-deoxy-*O*-β-D-galactopyranosyl-D-glucose structure, and the resistance to periodate oxidation suggests a (1→3) or a (1→4) linkage between the two components. The location of the 2-acetamido-2-deoxy-D-glucose residue in the middle of the chain is suggested by the absence of 2-acetamido-2-deoxy-D-glucitol after the β elimination under strongly alkaline conditions.

No chemical evidence is available to suggest the presence of asparagine-linked carbohydrate chains in epiglycanin. As discussed earlier, the high frequency of mucin-type chains and the extremely low levels of aspartic acid and D-mannose residues militate against the presence of a 2-acetamido-1-*N*-(L-aspart-4-oyl)-2-deoxy-β-D-glucopyranosylamine linkage, except in extremely minute proportion. This conclusion is supported by the observation that no inhibitory activity was observed for Fractions A, B, or C in hemagglutination experiments involving the lectins *Phaseolus vulgaris* and *Phaseolus limensis*. *Phaseolus vulgaris* lectin has been reported to bind to a 2-acetamido-2-deoxy-*O*-D-mannopyranosyl-D-glucopyranosyl sequence, which is found only in asparagine-linked chains. No inhibition of the action of *Phaseolus limensis* lectin, which is specific for the blood-group A antigen, was detected in GPF-I.

The distinction between long chains **1** and **2** and short chains **3**, **4**, and **5** rests solely upon the presence of a 2-acetamido-2-deoxy-D-glucose residue in the long chains. It may, indeed, be the case that only one chain type, namely **1**, is present in epiglycanin, and that all of the other structures proposed, **2**–**5**, represent an incomplete structure of **1**. This suggestion depends on the structures of the carbohydrate–peptide linkage and of the first two carbohydrate residues. More-complicated, branched structures are possible; for example, those having the *N*-acetylneuraminic acid group linked to the penultimate D-galactose residue in **1**. Until enough material

is available for the complete separation and investigation of the structure of each chain, this problem will probably not be resolved.

ACKNOWLEDGMENT

The authors thank Mrs. Cyla Silber for technical assistance in this investigation.

REFERENCES

1 G. M. W. COOK AND R. W. STODDART, *Surface Carbohydrates of the Eukaryotic Cell*, Academic Press, New York, 1973, pp. 257–293.
2 V. GINSBURG AND A. KOBATA, in L. I. ROTHFIELD (Ed.), *Structure and Function of Biological Biomembranes*, Academic Press, New York, 1971, pp. 439–459.
3 T. O. FOX, J. R. SHEPPARD, AND M. M. BURGER, *Proc. Nat. Acad. Sci. U.S.*, 68 (1971) 244–247.
4 P. GOLD, M. GOLD, AND S. O. FREEDMAN, *Cancer Res.*, 28 (1968) 1331–1334.
5 B. H. SANFORD, J. F. CODINGTON, R. W. JEANLOZ, AND P. D. PALMER, *J. Immunol.*, 110 (1973) 1233–1237.
6 J. F. CODINGTON, B. H. SANFORD, AND R. W. JEANLOZ, *J. Nat. Cancer Inst.*, 51 (1963) 585–591.
7 H. S. SLAYTER AND J. F. CODINGTON, *J. Biol. Chem.*, 248 (1973) 3405–3410.
8 J. F. CODINGTON, B. TUTTLE, AND R. W. JEANLOZ, in J. MONTREUIL AND G. SPIK (Eds.), *Colloquium on the Glycoconjugates*, Colloques Internationaux du C.N.R.S. (in press).
9 J. F. CODINGTON, B. H. SANFORD, AND R. W. JEANLOZ, *Biochemistry*, 11 (1972) 2559–2564.
10 A. G. COOPER, J. F. CODINGTON, AND M. C. BROWN, *Proc. Nat. Acad. Sci. U.S.*, 71 (1974) 1224–1228.
11 N. SHARON AND H. LIS, *Science*, 177 (1972) 949–959.
12 T. OSAWA, T. TERAO, T. KAWAGUCHI, M. FUKUDA, S. TOYOSHIMA, AND T. IRIMURA, in J. MONTREUIL AND G. SPIK (Eds.), *Colloquium on the Glycoconjugates*, Colloques Internationaux du C.N.R.S. (in press).
13 V. T. MARCHESI, R. L. JACKSON, J. P. SEGREST, AND I. KAHANE, *Fed. Proc.*, 32 (1973) 1833–1837.
14 M. FUKADA AND T. OSAWA, *J. Biol. Chem.*, 248 (1973) 5100–5105.
15 E. LISOWSKA, *Nature* (London), 198 (1963) 865–866.
16 G. F. SPRINGER, J. F. CODINGTON, AND R. W. JEANLOZ, *J. Nat. Cancer Inst.*, 49 (1972) 1469–1470.
17 G. UHLENBRUCK AND W. DAHR, *Vox Sang.*, 21 (1971) 338–351.
18 T. IRIMURA AND T. OSAWA, *Arch. Biochem. Biophys.*, 151 (1972) 475–482.
19 T. TERAO AND T. OSAWA, (submitted for publication).
20 S. TOYOSHIMA, Y. ALIYAMA, K. NAKANO, A. TONOMURA, AND T. OSAWA, *Biochemistry*, 10 (1971) 4457–4463.
21 T. TOMITA, T. KUROKAWA, K. ONOZAKI, H. ICHIKI, T. OSAWA, AND T. UKITA, *Experientia*, 28 (1972) 84–85.
22 J. A. GORDON, S. BLUMBERG, H. LIS, AND N. SHARON, *FEBS Lett.*, 24 (1972) 193–196.
23 N. R. GOULD AND S. SCHEINBERG, *Arch. Biochem. Biophys.*, 137 (1970) 1–11.
24 R. R. RACE AND R. SANGER, *Blood Groups in Man*, Blackford, Oxford, 1954, p. 78.
25 I. MATSUMOTO AND T. OSAWA, *Arch. Biochem. Biophys.*, 140 (1970) 484–491.
26 V. REINHOLD, *Methods Enzymol.*, 25 (1972) 244–249.
27 D. ROACH AND C. W. GEHRKE, *J. Chromatogr.*, 44 (1969) 269–278.
28 R. DULBECCO AND M. VOGT, *J. Exp. Med.*, 99 (1954) 167–182.
29 L. WARREN, *J. Biol. Chem.*, 234 (1959) 1971–1975.
30 J. F. CODINGTON, B. H. SANFORD, AND R. W. JEANLOZ, *J. Nat. Cancer Inst.*, 45 (1970) 637–647.
31 T. IRIMURA, T. KAWAGUCHI, G. TERAO, AND T. OSAWA, (submitted for publication).
32 E. LISOWSKA AND A. MORAWIECKI, *Eur. J. Biochem.*, 3 (1967) 237–241.
33 J. F. CODINGTON, A. G. COOPER, M. C. BROWN, AND R. W. JEANLOZ, *Biochemistry*, (in press).
34 H. TUPPY AND A. GOTTSCHALK, in A. GOTTSCHALK (Ed.), *Glycoproteins*, 2nd edn., Elsevier, Amsterdam, 1972, pp. 403–449.
35 S. CHIPOWSKY AND E. J. MCGUIRE, *Fed. Proc.*, 28 (1969) 606.
36 J. F. CODINGTON AND R. W. JEANLOZ, unpublished results.
37 R. W. JEANLOZ AND J. F. CODINGTON, in E. Y. C. LEE AND E. E. SMITH (Eds.), *Biology and Chemistry of Eukaryotic Cell Surfaces*, Academic Press, New York, 1974, pp. 241–257.