Biochem. J. (1978) **171**, 665–674
Printed in Great Britain

Properties of Potato Lectin and the Nature of its Glycoprotein Linkages

By ANTHONY K. ALLEN, NILA N. DESAI and ALBERT NEUBERGER
Department of Biochemistry, Charing Cross Hospital Medical School, Fulham Palace Road, London W6 8RF, U.K.

and J. MICHAEL CREETH
Department of Biochemistry, University of Bristol Medical School, University Walk, Bristol BS8 1TD, U.K.

(*Received* 26 *September* 1977)

1. Potato lectin is a glycoprotein that contains about 47% (by weight) L-arabinose, 3% D-galactose and 11% hydroxyproline. It has a monomeric molecular weight of about 50000 and probably exists as a monomer–dimer system in aqueous solution, with the monomer predominating. It has a very high viscosity, which would indicate either that the molecule is very expanded or that it is an elongated ellipsoid. 2. After prolonged proteolytic digestion of a reduced and carboxymethylated derivative of the lectin, a glycopeptide was isolated (of mol.wt. 32000–34000) that included all the carbohydrate and hydroxyproline of the original glycoprotein but less than 30% of the total original amino acid residues. 3. The arabinose of the glycoprotein is present exclusively as the β-arabinofuranoside and this includes those residues that are directly linked to the hydroxyproline residues of the polypeptide chain. All the arabinose of the glycoprotein is linked to the polypeptide chain through the hydroxyproline residues; the ratio of arabinose to hydroxyproline is 3.4:1. Although α-arabinofuranosides are known to be present in arabinans and arabinogalactans, the natural occurrence of β-arabinofuranosides has not previously been reported. 4. Nine or ten serine residues of the polypeptide chain are substituted with single α-galactopyranoside residues that can be removed by the action of α-galactosidase from coffee beans but not by a β-galactosidase. This is the first report of an α-galactoside linkage to serine. The effect of α-galactosidase is much greater on a glycopeptide from which the arabinose has been already removed, which indicates a steric hindrance of the galactosidase action by adjacent chains of arabinosides. 5. In 0.5M-NaOH (pH 13.7), galactose residues were removed from the serine residues of the glycopeptide by a process of β-elimination. This reaction took place very slowly in the intact glycopeptide but much more rapidly when the arabinofuranoside residues had been removed. This inhibitory effect of the arabinofuranoside residues on the β-elimination reaction is likely to be due to a negative charge on the hydroxy groups of the adjacent arabinofuranoside residues, which would be ionized at this high pH value. 6. It is suggested that potato lectin may be representative of a class of soluble plant glycoproteins that would include precursors of the cell-wall glycoprotein extensin. If this is the case, extensin should also contain β-L-arabinofuranosides linked to hydroxyproline and α-D-galactopyranosides linked to serine residues of the polypeptide chain.

Lectins are proteins or glycoproteins that have the capacity to specifically bind particular mono- or oligo-saccharides and compounds containing such groups. They are found in many plants and in some animals and are usually recognized by their agglutinating effect on erythrocytes [see reviews by Lis & Sharon (1973) and by Nicolson (1974)].

We described previously a method for the purification of potato lectin (Allen & Neuberger, 1973) by conventional chromatography. Subsequently Delmotte *et al.* (1975) described methods for purification by affinity chromatography. Potato lectin shows a specificity directed towards oligosaccharides that contain *N*-acetylglucosamine residues (Pardoe *et al.*, 1969; Masumoto & Osawa, 1971; Allen & Neuberger, 1973). We showed (Allen & Neuberger, 1973) that it has an unusual composition, being a glycoprotein containing about 50% carbohydrate, of which over 90% is arabinose, with smaller amounts of galactose. The most abundant amino or imino acid is hydroxyproline, and we suggested that the attachment of the carbohydrate to the protein moiety

Abbreviation used: SP-Sephadex, sulphopropyl-Sephadex.

of the lectin would probably be between hydroxyproline and arabinose, since this had been shown to be the case for the plant cell-wall glycoprotein extensin (Lamport *et al.*, 1973; Heath & Northcote, 1971, 1973). Evidence is given in the present paper that the lectin has a glycosylated domain comprising only one-third of the polypeptide chain but containing all of the carbohydrate. It is also shown that the hydroxyproline residues are linked to β-L-arabinofuranoside residues, whereas the serine residues are linked to α-D-galactose residues. A preliminary account of a part of this work has already been published (Allen *et al.*, 1976*a*).

Experimental

Materials

DEAE-cellulose (Whatman grade DE23) and CM-cellulose (Whatman grade CM23) were obtained from H. Reeve Angel and Co., London E.C.4, U.K. SP-Sephadex (C-50), Sephadex G-25 and Sephadex G-150 were from Pharmacia, Uppsala, Sweden, and Bio-Gel P-2 was from Bio-Rad Laboratories, Richmond, CA, U.S.A. Arabinan, α-methyl β-L-arabinopyranoside and poly-L-hydroxyproline were from Koch–Light Laboratories Ltd., Colnbrook, Bucks., U.K. Pronase (proteinase from *Streptomyces griseus*) was obtained from Sigma Chemical Co., St. Louis, MO, U.S.A. α-Galactosidase (from green coffee beans) was from the Boehringer Corporation (London) Ltd., Lewes, Sussex, U.K. and β-galactosidase (from *Escherichia coli*) was a gift from Dr. P. M. Dey, Royal Holloway College, Egham, Surrey, U.K. The amino acids and sugars used as analytical standards were obtained from BDH Chemicals Ltd., Poole, Dorset, U.K. Other chemicals were of BDH AnalaR grade or of the highest purity available.

Methods

Analysis of amino acids and sugars. Amino acids were measured on a Locarte amino acid analyser after hydrolysis of the glycoprotein or glycopeptide in 3M-toluene-*p*-sulphonic acid or constant-boiling HCl under N_2 at 110°C for 24h. The elution systems, correction factors for serine and threonine, and methods for measuring cystine and tryptophan, were those used previously (Allen & Neuberger, 1973; Allen *et al.*, 1976*b*).

Neutral sugars were measured by g.l.c. after methanolysis and trimethylsilylation of the glycoprotein (Chambers & Clamp, 1971), with a Perkin–Elmer F.11 gas chromatograph fitted with dual columns. Sugar analyses were related to the amino acid analyses by adding internal standards (mannitol for neutral sugars and *p*-fluorophenylalanine for amino acids) to samples taken from the same stock solution of glycoprotein or glycopeptide.

Polarimetry. Optical rotations were determined on a Perkin–Elmer model 141 polarimeter at 25°C with a 10cm cell.

Isoelectric focusing. An ISCO apparatus was used (ISCO Instrument Specialties Co., 4700 Superior, Lincoln, NE 68504, U.S.A.) with Ampholine buffers of range pH3.5–10 and pH9–11 (LKB, Stockholm, Sweden).

Ultracentrifugation. Solutions were prepared in a phosphate/chloride buffer (16.7mM-NaH_2PO_4/16.7mM-Na_2HPO_4/33mM-NaCl/2mM-NaN_3), pH6.8, I 0.10. Sedimentation-velocity experiments were performed in the Centriscan ultracentrifuge (MSE Ltd., Crawley, Sussex, U.K.), at 54000rev./min ($r_{av.}$ 6.5cm) and 25°C, scans being made at 4min intervals by schlieren optics. Sedimentation coefficients were calculated by using a computer program that allowed for the finite time of scan; they were corrected to standard conditions.

Sedimentation-equilibrium determinations were made by the method of Creeth & Holt (see Gratzer *et al.*, 1972) by using a Beckman model E ultracentrifuge with interference optics; the general conditions described by Creeth *et al.* (1974) were followed.

Where non-linear logarithmic plots were observed, the data were processed to obtain the number-average and weight-average molecular weights for the whole distribution, denoted $\bar{M}_n$ and $\bar{M}_w$. Where the plot was linear within experimental error (ignoring any slight curvature near the base of the cell), a single value, denoted M, was determined from the slope.

Partial specific volumes ($\bar{v}$) for both lectin and glycopeptide were calculated from the respective amino acid and carbohydrate compositions by using the values for the $\bar{v}$ values for the individual amino acids quoted by Cohn & Edsall (1943), and the value for galactose of 0.613ml/g quoted by Gibbons (1972). No value for arabinose could be found; accordingly a series of three density determinations on a purified sample of arabinose (BDH Ltd., Poole, Dorset, U.K.) was made with an Anton Paar densimeter. These density measurements, at concentrations of about 10, 20 and 25mg/ml, gave a mean $\bar{v}$ value for arabinose of 0.622ml/g, and this value was incorporated in the calculations. A single determination of $\bar{v}$ for the lectin at 4.7mg/ml was also made; for consistency with the glycopeptide calculations, however, the computed values of $\bar{v}$ were used for both lectin and glycopeptide. It was assumed that $\bar{v}$ for the lectin was unchanged in 6M-guanidinium chloride.

The reduced viscosity ($\eta_{red.}$) of the lectin in aqueous buffer was determined in the folded Ostwald type of viscometer (Holt & Creeth, 1972) at 3.0mg/ml. The procedure of Creeth & Knight (1965) was followed to give limiting estimates of shape.

Assay of agglutination. The lectin was assayed by

observing its agglutinating effect on a 1% suspension of rabbit erythrocytes in phosphate-buffered saline (7.20g of NaCl, 1.48g of Na_2HPO_4 and 0.43g of KH_2PO_4/litre; final pH7.2). Details of the procedure are given elsewhere (Allen & Neuberger, 1973; Allen *et al.*, 1976*b*).

Purification of potato lectin. The purification scheme is basically that described by us previously (Allen & Neuberger, 1973), with the exception that a single passage through a column (60cm×3.7cm) of Sephadex G-150 was used in place of the two successive passages through Sephadex G-100. Therefore the complete procedure involved extraction and homogenization of potato tubers in sodium acetate buffer (pH3.8), precipitation of the supernatant by the addition of solid $(NH_4)_2SO_4$ to produce 60% saturation (361g/litre), chromatography on DEAE-cellulose, CM-cellulose, Sephadex G-150 and finally SP-Sephadex. The content of lectin in tubers of different varieties varies considerably (Krüpe, 1956). The variety King Edward VII, which is available in Britain, has a reasonably high content of lectin (Allen & Neuberger, 1978), and has been used as a source of lectin both for this preparation and for our previous investigation (Allen & Neuberger, 1973).

Preparation of a glycopeptide from potato lectin

The lectin (10mg) was reduced and alkylated by the method of Konigsberg (1972), which involved reduction by dithiothreitol in 6M-guanidinium chloride at 50°C, followed by carboxymethylation with iodoacetic acid and subsequent dialysis. The resultant *S*-carboxymethylcysteinyl-lectin was then digested with Pronase under the conditions used by Spiro & Bhoyroo (1974) for the digestion of fetuin. Digestion was at 37°C in 0.15M-Tris/acetate buffer, pH7.8, containing 1.5mM-$CaCl_2$ in the presence of toluene for a total of 6 days. Initially, Pronase equal to 1% of the weight of lectin was added, followed at 2 and 4 days by amounts equal to 0.05% of the weight of the lectin. The digest was then centrifuged at 600*g* for 2min and applied to a column (60cm×2cm) of Bio-Gel P-2 or Sephadex G-25 equilibrated with water. Fractions (3ml) were collected and portions were assayed for neutral sugar by the phenol/H_2SO_4 method (Dubois *et al.*, 1956) and for amino acids by ninhydrin. Material that contained carbohydrate (the glycopeptide) was eluted at the void volume (V_0) and the ninhydrin-positive material was eluted later in the position expected for free amino acids. A sample of the pooled free-amino acid fraction was analysed directly, and a sample of the glycopeptide fraction after hydrolysis in 3M-toluene-*p*-sulphonic acid for 24h at 110°C was also analysed. Since the analysis of the glycopeptide showed that it contained a number of amino acids that are not normally linked to sugar residues, the cycle of digestion and gel chromatography was repeated. The glycopeptide was then re-analysed for its amino acid content and for neutral sugars.

Partial acid hydrolysis of the glycopeptide

(*a*) *With oxalic acid.* The material was hydrolysed in a sealed tube under N_2 for 5h at 100°C in 12.5mM-oxalic acid (1ml), cooled, applied to a column (60cm ×2cm) of Bio-Gel P-2 and eluted with water.

(*b*) *With* 1M-*HCl.* The material was hydrolysed at 30°C in 1M-HCl, and the rate of reaction followed by the rate of change of optical rotation. Control hydrolyses were done on arabinan and methyl β-L-arabinopyranoside. From these readings, half-lives for the compounds were derived. The products of hydrolysis from both methods were purified on the column of Bio-Gel P-2 (see above).

Alkaline degradation of the lectin and glycopeptide

(*a*) 5M-*NaOH.* Samples of the lectin (2mg) were kept at 100°C for 24h in 1ml of 5M-NaOH in a sealed tube. The sample was then cooled, neutralized with 5M-HCl (with Bromothymol Blue as an indicator), and chromatographed on Bio-Gel P-2. A polypropylene centrifuge tube was used as a liner in later experiments to prevent formation of sodium silicate from the action of NaOH on the glass of the tube (Hugli & Moore, 1972).

(*b*) 0.5M-*NaOH at* 4°*C.* The lectin (1mg) was dissolved in 0.5ml of 0.5M-NaOH and dialysed against 0.5M-NaOH at 4°C for 2 days followed by dialysis against several changes of water. The lectin was then analysed for amino acids and sugars.

(*c*) 0.5M-*NaOH*+0.3M-$NaBH_4$ *at* 4°*C.* The glycopeptide and the acid (1M-HCl)-treated glycopeptide were dissolved in 1ml of solution that contained 0.5M-NaOH and 0.3M-$NaBH_4$, and were incubated at 4°C for 2 days. A small sample (100μl) of each was added to 100μl of conc. HCl and hydrolysed for 24h at 110°C under N_2. The samples were then analysed for their amino acid composition.

(*d*) 0.5M-*NaOH*+0.3M-$NaBH_4$ *at* 30°*C.* Both the glycopeptide and the acid-treated glycopeptide solutions were then incubated at 30°C, and samples (100μl) were taken after 1 and 7 days, hydrolysed in HCl as above, and analysed for their amino acid composition. A sample was also taken after 1 day for carbohydrate analysis by g.l.c.

α- and β-Galactosidase treatment of glycopeptides

α-*Galactosidase.* About 100μg of glycopeptide (either complete or acid-treated) was dissolved in 200μl of 0.05M-citrate/phosphate buffer, pH6.0, and incubated with 40μl (2 units) of coffee-bean α-galactosidase at 37°C for 7 days in the presence of

toluene. (One unit will hydrolyse 1.0 μmol of *p*-nitrophenyl α-D-galactoside/min at 25°C.) The incubation mixture was then dialysed against water and analysed.

β-Galactosidase. About 100 μg of glycopeptide was dissolved in 300 μl of phosphate-buffered saline, pH 7.2, and incubated with 2 μl (16 units) of *E. coli* β-galactosidase at 37°C for 24h in the presence of toluene. (One unit will hydrolyse 1.0 μmol of lactose/min at 37°C.) The incubation mixture was then dialysed against water and analysed.

Results and Discussion

Properties of potato lectin

The physicochemical properties of the native lectin are summarized in Table 1 (i). The sedimentation-velocity patterns showed single symmetrical peaks, but it is noteworthy that the sedimentation coefficient is independent of the concentration. The molecular weight, however, is distinctly concentration-dependent, the values increasing with concentration. These observations taken together suggest that the lectin is an aggregating system; if the degree of aggregation is to a dimer only, the corresponding monomer molecular weight would be about 50000. At 1 mg/ml, about 30% would be present as dimer. The viscosity is remarkably high for a protein; taking the reduced viscosity as identical with the intrinsic viscosity, one can find the extremes of expansion or asymmetry indicated. Clearly the molecule is either considerably expanded or very asymmetric; no distinction is possible on these results alone.

From the values of the molecular weight in denaturing solvents [Table 1 (ii)], it appears that the aggregation is not suppressed by guanidinium chloride alone; the lower value obtained after reduction of the disulphide bonds is in good agreement with that suggested for the native monomer and indicates that the molecule has only a single peptide chain.

The values found for the sedimentation coefficient are in good agreement with those previously obtained (Allen & Neuberger, 1973) for a different batch of lectin. The monomer molecular weight indicated, about 50000, is only slightly higher than the value of 46000 previously obtained by Allen & Neuberger (1973) for both the native lectin and the *S*-carboxymethylcysteinyl derivative on gel filtration in 6M-guanidinium chloride on Sepharose 4B. It is in contrast with the other values obtained previously (Allen & Neuberger, 1973) for the native lectin on gel filtration on Sephadex G-150 with non-denaturing buffers (120000) and on polyacrylamide-gel disc electrophoresis (105000).

Potato lectin is a very basic protein. On isoelectric focusing, three isolectins of pI 9.64±0.02, 9.52±0.03 and 9.42±0.02 were obtained. Other preparations have in addition yielded lesser quantities of other isolectins with pI values of 8.43±0.03 and 7.70±0.05.

The analysis of the whole lectin is given in Table 2(*a*); the values are calculated as molar proportions by assuming that the molecular weight of the monomer is about 50000. Preparations of potato lectin are mixtures of isolectins that are likely to differ slightly in composition and to vary in proportion between batches of potatoes. It is therefore not unexpected that, although the composition generally resembles that of our previous preparation (Allen & Neuberger, 1973), there are some differences; the proportion of hydroxyproline is higher in this preparation, phenylalanine is present in small (less than molar) amounts and histidine is absent. The proportion of carbohydrate is similar to that in our previous preparation, but we were not able to detect the 1 mol of glucosamine that was present in our previous preparation. We did, however, as previously, find 1 mol of a basic amino acid that behaves chromatographically like ornithine. It is noteworthy that Jermyn & Yeow (1975) also observed an unknown basic amino acid that was identical with ornithine in chromatographic behaviour in their 'all-β' lectins isolated from a number of different plants.

Isolation of a glycopeptide from potato lectin

After repeated digestion of the *S*-carboxymethylcysteinyl derivative of potato lectin with Pronase

---

Table 1. *Physicochemical parameters of native and denatured potato lectin and of the derived glycopeptide*

(i) Native lectin
- $s_{20,w}$ 3.18±0.02 S (*c* 4.7 mg/ml)
  - 3.18±0.03 S (*c* 1.5 mg/ml)
- $\bar{v}$ 0.655 ml/g (calculated)
  - 0.63 ml/g (observed)
- $\eta_{red.}$ 22.6 ml/g (*c* 3.0 mg/ml)
- $\bar{M}_n$ 58600, $\bar{M}_w$ 66100 (*c* 2 mg/ml); $\bar{M}_n$ 56900, $\bar{M}_w$ 63400 (*c* 1 mg/ml)
- M_1 if mono–dimer system (see the text): approx. 50000
- Shape estimations: (*a*) spherical-expanded, $\eta_{red.}/2.5\bar{v}=14$
  - (*b*) compact prolate ellipsoid f/f_0 (frictional ratio) = 2.4; b/a (axial ratio) ~ 30

(ii) Denatured lectin
- *M* 60000 (*c* 0.2 mg/ml in 6M-guanidinium chloride)
  - 48000 (*c* 0.2 mg/ml in 6M-guanidinium chloride/0.04M-dithiothreitol)

(iii) Glycopeptide
- $\bar{v}$ 0.641 ml/g (calculated)
- *M* 32500 (*c* 2 mg/ml in aqueous buffer)
  - 32000 (*c* 1 mg/ml in aqueous buffer)
  - 34000 (by calculation from composition given in Table 2*b*)

Table 2. *Amino acid and carbohydrate composition of potato lectin compared with fractions obtained after Pronase digestion, mild acid hydrolysis and 5M-NaOH digestion*

For further details see the text. Results are expressed as molar ratios relative to hydroxyproline (=50). The values for the native lectin are to the nearest whole residue, those for the glycopeptides to the nearest tenth of a residue. Abbreviation used: ND, not determined.

Amino acid	(*a*) Native lectin	(*b*) Glycopeptide isolated after Pronase digestion	(*c*) Glycopeptide after Pronase and 1M-HCl treatments	(*d*) Carbohydrate-containing fraction isolated after digestion with 5M-NaOH
Hyp	50	50	50	50†
Asp	12	2.3	1.2	0.6
Thr	14*	1.6*	0.6*	ND
Ser	31*	16.8*	14.5*	0.4
Glu	17	4.1	0.9	0.8
Pro	17	3.7	0.5	0.0
Gly	30	4.9	2.2	1.0
Ala	10	4.4	2.7	0.9
½Cys	26	<0.1	<0.1	0.0
Val	1	<0.1	<0.1	0.0
Met	1	0.0	0.0	0.0
Ile	4	0.5	<0.2	0.2
Leu	3	3.0	2.1	0.5
Tyr	8	0.3	<0.2	0.0
Phe	0.5	0.5	<0.2	0.0
His	0	0	ND	ND
Lys	9	1.4	ND	ND
Orn	1	ND	ND	ND
Arg	3	0.6	ND	ND
Trp	8	<0.5	ND	ND
Total residues	245.5	94.1	74.7	
Ara	172	172	0.0	172
Gal	9.6	9.1	9.4	0.0
Mol.wt.	50300	34100	9400	580 (average)

* Corrected for destruction during hydrolysis.
† Sum of L-hydroxyproline and D-allohydroxyproline.

(see under 'Methods'), a glycopeptide of high molecular weight was separated by chromatography on Bio-Gel P-2 or Sephadex G-25. It contained all the hydroxyproline, arabinose and galactose of the original molecule, lesser amounts of serine and other amino acids, but lacked *S*-carboxymethylcysteine (Table 2*b*). This was confirmed by analysis of the fraction that included the free amino acids released by the action of Pronase, as this fraction lacked hydroxyproline, arabinose and galactose. It was found in a control experiment that poly-L-hydroxyproline was not digested by Pronase.

The glycopeptide did not pass through Visking 8/32 dialysis tubing when dialysed at 4°C in water for 24h, and from this it seemed likely that its molecular weight was at least 10000. The determination of molecular weight by sedimentation equilibrium gave a value of 32000 [Table 1(iii)]. This is in reasonable agreement with the value of 34000 obtained by calculation if one assumes that all 50 hydroxyproline residues of the original glycoprotein are in one peptide chain after the Pronase digestion.

The optical rotatory dispersion of the glycopeptide is shown in Fig. 1; it had an $[\alpha]_D$ of +59°. The values for the rotations are shown for the glycopeptide (10mg/ml) in water in a 10cm cell.

Partial hydrolysis of the glycopeptide

We found that arabinose residues of the glycopeptide could be removed by mild acid hydrolysis. Thus hydrolysis in 12.5mM-oxalic acid at 100°C for 5h (or 1M-HCl at 30°C for 10 days; see below) resulted in a glycopeptide that lacked arabinose but still contained all the galactose and almost all the amino acids of the original glycopeptide (Table 2*b*). The small differences in amino acid composition between the intact and the acid-treated glycopeptide could be explained at least partially by cleavage of a peripheral acid-labile peptide bond involving an aspartic acid residue (Harris *et al.*, 1956). The arabinose that was released by this treatment had an $[\alpha]_D$ of +100° at equilibrium, showing that it was the natural L-stereoisomer (literature value +104.5°).

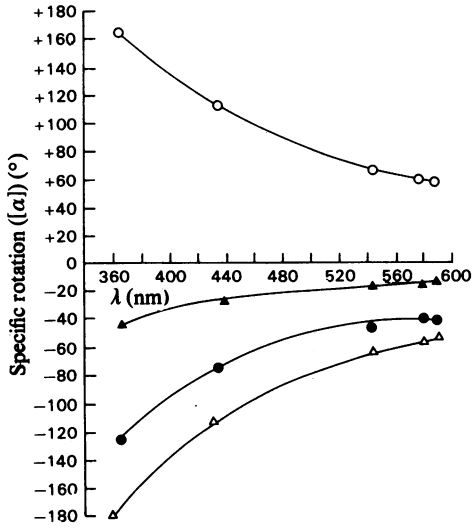

Fig. 1. *Optical-rotatory-dispersion curves of glycopeptides and hydroxyproline*

Complete arabinoglycopeptide from potato lectin (○), the glycopeptide from which arabinose has been removed by mild acid hydrolysis (●), free hydroxy-L-proline (▲) and polyhydroxy-L-proline (△) were examined. The curve for the specific rotation of the complete arabinoglycopeptide is shown; the rotations for the other compounds are adjusted to a concentration of hydroxyproline that is equimolar to that of the complete glycopeptide. The composition of the glycopeptides is given in Table 2. For further details, see the text.

The optical rotatory dispersion of the acid-treated glycopeptide is shown in Fig. 1. It is adjusted to the same value for hydroxyproline as the optical-rotatory-dispersion traces of the untreated glycopeptide, as are those of poly-L-hydroxyproline and free hydroxyproline, which are also shown. The optical rotation of the acid-treated glycopeptide is about three times greater than would be expected (in terms of the rotation of free hydroxyproline) from its hydroxyproline content (see Fig. 1). It is, however, almost as laevorotatory as an equivalent quantity of polyhydroxyproline. This indicates that the hydroxyproline is present in the glycopeptide as contiguous units of three or more residues that could form an open rigid left-handed helix in the conformation of polyproline (form II) as was suggested by Heath & Northcote (1971) for glycopeptides from extensin.

It is apparent that the removal of the arabinose from the glycopeptide had the effect of changing the optical rotation from positive to negative. The values for the specific molar rotation of the arabinose residues were calculated from this difference by using analytical values for the glycopeptides obtained by amino acid analysis and g.l.c. When compared with molar rotation of the α- and β-methyl arabino-furanosides and -pyranosides (Table 3) it can be seen that the molar rotation of the arabinose in the glycopeptide (+198°) is very close to that of the β-methyl arabinofuranoside (+194°). This rotation makes it unlikely that the α-furanoside is present in significant amounts, since it has a strong laevorotation ($[M]_D$ −209°). The rotation of the arabinose in the glycopeptide is also significantly different from those of the arabinopyranosides (Table 3).

Table 3. *Molar rotation $[M]_D$ for the arabinose as bound in the potato lectin glycopeptide compared with the literature values of $[M]_D$ for methylarabinosides*

For further details, see the text.

Arabinose in glycopeptide	+198° (calculated)
Methyl α-L-arabinofuranoside	−209°*
Methyl β-L-arabinofuranoside	+194°*
Methyl α-L-arabinopyranoside	+25°†
Methyl β-L-arabinopyranoside	+401°†

* Augestad & Berner (1956).
† Recalculated from Overend *et al.* (1962).

The diagnosis of a furanoside linkage for the arabinose was confirmed by the acid-lability of the linkage. It had already been found (see above) that the conditions that were used by Fincher *et al.* (1974) to hydrolyse the α-L-arabinofuranoside linkages of arabinogalactan, namely 12.5mM-oxalic acid at 100°C for 5h, were sufficient to release free arabinose from the lectin glycopeptide. We have also found that these conditions were sufficient to release arabinose from arabinan (an α-L-arabinofuranoside polymer), but not from β-methyl L-arabinopyranoside. To obtain a more accurate value for the hydrolysis for comparison with literature values, 1M-HCl was used at 30°C. At this temperature the half-life of the arabinose of the glycopeptide is 2.0 days, compared with 2.1 days for the α-L-arabinofuranoside linkage of arabinan and the 20-fold greater stability of β-methyl L-arabinopyranoside (half-life 41 days). By comparison, Augestad & Berner (1956) on hydrolysing methyl L-arabinofuranosides in 1M-HCl at 20°C observed half-lives of 56h 5min for the α- and 8h 15min for the β-furanoside.

Effect of alkaline treatments on whole lectin and glycopeptide

Since galactose remained attached to the peptide chain after mild acid hydrolysis (see Table 2), it was concluded that it was present as a pyranoside and was linked directly to an amino acid of the peptide chain, presumably by an ether linkage to hydroxyproline or serine. It should be possible to distinguish between these possibilities by alkaline

treatment, as hydroxyproline glycosides are known to be stable to alkali (Vercellotti & Just, 1967; Fincher *et al.*, 1974), but serine or threonine glycosides are usually degraded in alkali by a process of β-elimination (see review by Marshall & Neuberger, 1970).

Treatment with 5M-*NaOH*. Native potato lectin (2.0mg) was treated with 5M-NaOH (see under 'Methods') at 100°C for 24h. After neutralization the product was chromatographed on Bio-Gel P-2 with water as the eluent. A peak of material that contained carbohydrate was eluted in three tubes, before the position of the free amino acids. Analysis after appropriate acid hydrolysis or methanolysis indicated that this fraction was a glycopeptide containing mainly hydroxyproline (an equimolar mixture of L-hydroxyproline and D-allohydroxyproline) attached to arabinose oligosaccharides (Table 2*d*). The ratio of arabinose to the sum of the hydroxyproline residues varied from 3.7:1 to 2.9:1 according to the fraction. There were only about 3mol.equiv. of free hydroxyproline present in the free amino acid fractions compared with 47mol.equiv. in the glycosylated fraction. This indicated that almost all the hydroxyproline of the potato lectin is glycosylated by arabinose, and furthermore that galactose is not attached to hydroxyproline.

Treatment with 0.5M-*NaOH*. As previously reported (Allen & Neuberger, 1973), treatment of the complete potato lectin with 0.5M-NaOH at 4°C for 48h has no effect on the carbohydrate composition. The effect of similar conditions (apart from the addition of 0.3M-$NaBH_4$) on either the complete glycopeptide (with arabinose still attached) or on the acid-treated glycopeptide (which lacks arabinose) caused a slight (20–30%) loss of serine (see Table 4). On continuing treatment for 24h with the same reagent (after raising the temperature from 4 to 30°C), only 1.7 galactose residues and 5.5 residues of serine were lost from the whole glycopeptide. By contrast, all the galactose (9.4 residues) and 10.7 residues of serine were eliminated from the acid-treated glycopeptide. This showed firstly that the serine residues were substituted with only one galactose residue and secondly that the arabinose chains were protecting at least seven of these serine galactosides from β-elimination.

Galactosidase treatment (*Table* 5). The whole glycopeptide and the acid-treated glycopeptide were incubated with both α- and β-galactosidases (see under 'Methods'). β-Galactosidase had no effect on the composition of either glycopeptide. α-Galactosidase completely removed the galactose from the acid-treated glycopeptide, but under the same con-

Table 4. *β-Elimination of glycopeptides*

The effects of treatment with 0.5M-NaOH+0.3M-$NaBH_4$ on the amino acid and carbohydrate composition of (*a*) the glycopeptide derived from potato lectin by Pronase digestion (contains arabinose) and (*b*) the same glycopeptide after treatment with 1M-HCl at 30°C for 10 days (does not contain arabinose) are shown. Results are expressed as molar ratios relative to hydroxyproline (=50). For further details, see the text. Abbreviation used: ND, not determined.

	(*a*)				(*b*)			
	Untreated	96h at 4°C	24h at 30°C	144h at 30°C	Untreated	96h at 4°C	24h at 30°C	144h at 30°C
Hyp†	50	50	50	50†	50	50	50	50†
Ser*	16.8	13.5	11.3	10.2	14.5	11.3	3.8	2.5
Ala	4.4	3.2	3.0	3.3	2.7	2.3	3.9	4.2
Ara	172	ND	172	ND	0	ND	0	ND
Gal	9.1	ND	7.4	ND	9.4	ND	0	ND

* Corrected for destruction during hydrolysis.
† Sum of L-hydroxyproline and D-allohydroxyproline.

Table 5. *Galactosidase treatment of glycopeptides*

The effects of treatment with α- and β-galactosidases on the carbohydrate composition of (*a*) the glycopeptide derived from potato lectin by Pronase digestion and (*b*) the same glycopeptide after treatment with 1M-HCl at 30°C for 10 days are shown. Results are expressed as molar ratios relative to hydroxyproline (=50). For further details, see the text; the full composition of the glycopeptides is given in Table 2.

	(*a*)			(*b*)		
	Untreated	α-Galactosidase	β-Galactosidase	Untreated	α-Galactosidase	β-Galactosidase
Ara	172	172	172	0	0	0
Gal	9.1	6.1	9.1	9.4	<0.05	9.4

ditions removed only one-third of the galactose from the native glycopeptide. It is therefore concluded that all the galactose residues of the glycoprotein are α-linked to serine, and that the presence of arabinose in the glycopeptide inhibits the action of the galactosidase, presumably by steric hindrance.

General discussion

The monomer of potato lectin has a mol.wt. of about 50000. In aqueous solution the lectin exists as a monomer–dimer system with the monomer predominating. This 50000-mol.wt. species is a true monomer, since the molecular-weight values obtained by ultracentrifugation in guanidinium chloride in the presence of dithiothreitol are also in the region of 50000. This is the only example of which we are aware of a lectin in which a monomer is present in aqueous solution. Since lectins are presumed to agglutinate cells by forming cross-bridges between the glycoproteins of adjacent cells, one must conclude either that the monomoer has two binding sites or that the dimer is the active species. The frictional ratio (f/f_0) of 2.4 is compatible with either a very long cylindrical molecule or a very expanded sphere, or perhaps a combination of these. If the molecule is cylindrical or ellipsoidal this would probably make it a more effective bridge with less steric hindrance for agglutinating cells than an expanded sphere. This may be of particular importance for lectins such as potato lectin or wheat-germ agglutinin, which have a specificity directed towards *NN'*-diacetylchitobiose, which is found in the core region of many glycopeptides.

In common with other glycoproteins such as glycophorin (Marchesi *et al.*, 1976), blood-group glycoproteins (Donald, 1973) and gastric mucus (Scawen & Allen, 1977), potato lectin consists of two or more very dissimilar domains. The glycopeptide (mol.wt. 32000–34000) that was isolated by Pronase digestion of the *S*-carboxymethylcysteinyl-lectin represents one domain that contains all the hydroxyproline and all of the carbohydrate of the original lectin as well as half of the serine residues but none of the cystine. The remaining two-thirds of the polypeptide chain could be either continuous or in two parts separated by the glycopeptide portion. This remaining portion must be highly cross-linked, since about one residue in six is a half-cystine and it would include most of the charged groups of this very basic protein. This cystine-rich region is also likely to contain the lectin-binding site, because the glycopeptide, which is unlikely to have been affected by reduction and alkylation, showed neither agglutinating activity nor inhibitory effect on the agglutinating action of the whole lectin.

From the evidence given in the present paper there is little doubt that the arabinose of potato lectin is linked to the 4-hydroxyproline residues of the polypeptide chain and that almost all the hydroxyproline residues are substituted with L-arabinose with an average hydroxyproline/arabinose ratio of 1:3.4. The arabinose is present as the furanoside; from the polarimetric observations, at least 98% of the arabinofuranoside must be present as the β-anomer (Table 3) and this therefore means that the arabinose that is involved in direct linkage with hydroxyproline (27% of the total) must also be of the β-configuration. (If the linkage of arabinose to the hydroxyproline residues were of the α-configuration and the rest of the arabinose chain were of the β-configuration, the average molar rotation of the bound arabinose would be +84° rather than the observed value of +198°.)

As far as we are aware this is the first report of β-arabinofuranosides in a natural product. Arabinans (Whistler & Smart, 1953) and arabinogalactans (Fincher *et al.*, 1974) have only α-arabinofuranoside residues.

We have shown by the use of a specific α-galactosidase that galactose is present in the glycoprotein as a pyranoside linked to serine residues by an α-linkage. The ratio of galactose to serine lost on β-elimination of the acid-treated glycopeptide is close to unity, making it probable that there is a single galactose residue linked to each of nine or ten serine residues. Since all the galactose of the whole lectin is present in the glycopeptide produced by Pronase digestion of the whole lectin, these serine residues must be in the same region of the glycoprotein as the hydroxyproline arabinosides.

The anomeric nature of the linkage compound between galactose and serine in a glycoprotein has not been previously reported. The linkage between *N*-acetylgalactosamine and serine or threonine has been shown to be of the α-configuration in a variety of glycoproteins, and β-linkages between galactose and hydroxylysine are well-documented in collagens and basement membranes from higher animals (see reviews by Marshall, 1972, 1974).

The anomeric nature of the linkage compound between galactose and serine in extensin or its precursors (Heath & Northcote, 1971; Lamport *et al.*, 1973; Cho & Chrispeels, 1976) is not known, nor is the linkage between galactose and serine or threonine in earthworm cuticle collagen (Lee & Lang, 1968). Linkages between hydroxyproline and galactose have been reported from wheat endosperm arabinogalactan-peptide (Fincher *et al.*, 1974) and from cell walls of the motile green alga *Chlamydomonas reinhardii* (Miller *et al.*, 1972), but the anomeric configuration is not known in either case.

The conditions necessary for β-elimination (see review by Marshall & Neuberger, 1970) of the serine-to-galactose linkage require further comment, since these linkages in the potato lectin are unusually

stable to alkali, particularly those in the glycopeptide with arabinose still attached. We think it unlikely that the effect is due to steric hindrance, and, since there are nine or ten glycosylated serine residues in the glycopeptide, this cannot be due to the inhibitory effect on an ionized carboxy group in a terminal position. However, in 0.5M-NaOH, in which the pH is about 13.7, the 2′- and 3′-hydroxy groups of arabinofuranosides may be presumed to be mostly ionized [pK values of 13.3 have been obtained by Remin *et al.* (1976) for the hydroxy groups of arabinofuranosyl nucleosides]. These negatively charged groups, of which there must be a large number at pH13.7, are likely to be adjacent to some, if not all, of the glycosylated serine residues and would have an inhibitory effect on β-elimination. When these arabinofuranoside groups are removed by mild acid treatment, the free hydroxy groups of the hydroxyproline residues may be weakly ionized, but they are not likely to have as low a pK value as have arabinofuranoside hydroxy groups and thus their inhibitory effect on the β-elimination reaction would be much less. Although the hydroxy groups of pyranose sugars are known to have higher pK values than those of furanose sugars, pK values of near 14 have been observed for pyranosides, and therefore similar but smaller effects might be expected in other heavily glycosylated glycoproteins. In many animal glycoproteins, such as mucins or glycophorin, a more important factor may be that many of the side chains terminate in sialic acids, which would be likely to have some inhibitory effect on β-elimination. It is likely that prior treatment of such glycoproteins with neuraminidase would increase the rate of β-elimination.

The glycopeptide domain of potato lectin, with its high content of hydroxyproline arabinosides and serine galactosides, is very similar to that of the plant cell-wall glycoprotein extensin (Heath & Northcote, 1971; Lamport *et al.*, 1973; Chrispeels *et al.*, 1974), and possibly to other material isolated from plant cell walls (Selvendran *et al.*, 1975; Selvendran, 1975; Pope, 1977). Hydroxyproline arabinosides have also been shown to be present in soluble plant glycoproteins (Pusztai & Watt, 1969; Mani & Radnakrishnan, 1974; Knee, 1975; Pope, 1977).

Among the soluble glycoproteins that contain arabinose and hydroxyproline are a widely distributed class of lectins that, although they may not agglutinate erythrocytes, have a general specificity directed towards β-D-glycopyranosyl residues (Jermyn & Yeow, 1975). They also differ from potato lectin in having a much higher carbohydrate/protein ratio (about 8:1). Yamagishi *et al.* (1976) have isolated a proteoglycan from rice bran and established the presence of O-α-L-arabinofuranosyl–hydroxyproline linkages. This preparation probably represents a class of compound different from our β-arabinose-containing lectin, since it contains in addition much higher proportions of galactose and xylose.

It is therefore apparent that hydroxyproline residues in plant glycoproteins can be linked to either β-arabinofuranosides, as in potato lectin, α-arabinofuranosides, as in rice bran proteoglycan (Yamagishi *et al.*, 1976) or galactopyranosides, as in the arabinogalactans (Fincher *et al.*, 1974) or in *Chlamydomonas* cell wall (Miller *et al.*, 1972). Perhaps the 'sequon' (sequence of amino acids; Marshall, 1974) required for glycosylation of hydroxyproline by a β-L-arabinofuranoside is a series of adjacent hydroxyproline residues. If this is the case, potato lectin could be regarded as a representative of a class of soluble plant glycoproteins that would include the extensin precursors. We therefore consider it important to establish the anomeric type of the glycosidic linkages of extensin, which is known to contain hydroxyproline residues linked to arabinofuranosides and serine residues linked to galactopyranosides (Heath & Northcote, 1973; Lamport *et al.*, 1973; Cho & Chrispeels, 1976).

Note Added in Proof (Received 15 March 1978)

Since submission of this paper, Akiyama & Katō (1977) have reported that hydroxyprolyl tri- and tetra-arabinosides that were isolated from cell-wall preparations by cultured tobacco cells by alkaline digestion contain β-L-arabinofuranoside residues.

We thank the Medical Research Council and the Agricultural Research Council for financial support, and Dr. P. M. Dey for the gift of β-galactosidase.

References

Akiyama, Y. & Katō, K. (1977) *Agric. Biol. Chem.* **41**, 79–81

Allen, A. K. & Neuberger, A. (1973) *Biochem. J.* **135**, 307–314

Allen, A. K. & Neuberger, A. (1978) *Methods Enzymol.* **50C**, 340–345

Allen, A. K., Desai, N. N. & Neuberger, A. (1976*a*) *Proc. Int. Congr. Biochem. 10th*, Abstr. 10-6-302

Allen, A. K., Desai, N. N. & Neuberger, A. (1976*b*) *Biochem. J.* **155**, 127–135

Augestad, I. & Berner, E. (1956) *Acta Chem. Scand.* **10**, 911–916

Chambers, R. E. & Clamp, J. R. (1971) *Biochem. J.* **125**, 1009–1018

Cho, Y. P. & Chrispeels, M. J. (1976) *Phytochemistry* **15**, 165–169

Chrispeels, M. J., Sadava, D. & Cho, Y. P. (1974) *J. Exp. Bot.* **25**, 1157–1166

Cohn, E. J. & Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides*, p. 370, Reinhold Publishing Corp., New York

Creeth, J. M. & Knight, C. G. (1965) *Biochim. Biophys. Acta* **102**, 549–558

Creeth, J. M., Bhaskar, K. R., Donald, A. S. R. & Morgan, W. T. J. (1974) *Biochem. J.* **143**, 159–170
Delmotte, F., Kieda, C. & Monsigny, M. (1975) *FEBS Lett.* **53**, 324–330
Donald, A. S. R. (1973) *Biochim. Biophys. Acta* **317**, 420–436
Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356
Fincher, G. B., Sawyer, W. H. & Stone, B. A. (1974) *Biochem. J.* **139**, 535–545
Gibbons, R. A. (1972) in *Glycoproteins, their Composition, Structure and Function* (Gottschalk, A., ed.), pp. 31–140, Elsevier, Amsterdam
Gratzer, W. B., Creeth, J. M. & Beaven, G. H. (1972) *Eur. J. Biochem.* **31**, 505–509
Harris, J. I., Cole, R. D. & Pon, N. G. (1956) *Biochem. J.* **62**, 154–159
Heath, M. F. & Northcote, D. H. (1971) *Biochem. J.* **125**, 953–961
Heath, M. F. & Northcote, D. H. (1973) *Biochem. J.* **135**, 327–329
Holt, J. C. & Creeth, J. M. (1972) *Biochem. J.* **129**, 665–676
Hugli, T. E. & Moore, S. (1972) *J. Biol. Chem.* **247**, 2828–2834
Jermyn, M. A. & Yeow, Y. M. (1975) *Aust. J. Plant Physiol.* **2**, 501–531
Knee, M. (1975) *Phytochemistry* **14**, 2181–2188
Konigsberg, W. (1972) *Methods Enzymol.* **25**, 185–188
Krüpe, M. (1956) *Blutgruppenspezifische Pflanzliche Eiweisskörper*, pp. 11–12, F. Enke Verlag, Stuttgart
Lamport, D. T. A., Katona, L. & Roerig, S. (1973) *Biochem. J.* **133**, 125–131
Lee, Y. C. & Lang, D. (1968) *J. Biol. Chem.* **243**, 677–680
Lis, H. & Sharon, N. (1973) *Annu. Rev. Biochem.* **42**, 541–574
Mani, U. V. & Radnakrishnan, A. N. (1974) *Biochem. J.* **141**, 147–153
Marchesi, V. T., Furthmayer, H. & Tomita, M. (1976) *Annu. Rev. Biochem.* **45**, 667–698
Marshall, R. D. (1972) *Annu. Rev. Biochem.* **41**, 673–702
Marshall, R. D. (1974) *Biochem. Soc. Symp.* **40**, 17–26
Marshall, R. D. & Neuberger, A. (1970) *Adv. Carbohydr. Chem. Biochem.* **25**, 407–478
Masumoto, I. & Osawa, T. (1971) *Vox Sang.* **21**, 548–557
Miller, D. H., Lamport, D. T. A. & Miller, M. (1972) *Science* **176**, 918–920
Nicolson, G. L. (1974) *Int. Rev. Cytol.* **39**, 89–190
Overend, W. G., Rees, C. W. & Sequeira, J. S. (1962) *J. Chem. Soc.* 3429–3440
Pardoe, G. I., Bird, G. W. G. & Uhlenbruck, G. (1969) *Z. Immunitätsforch. Allerg. Klin. Immunol.* **137**, 442–457
Pope, D. G. (1977) *Plant Physiol.* **59**, 894–900
Pusztai, A. & Watt, W. B. (1969) *Eur. J. Biochem.* **10**, 523–532
Remin, M., Darzynkiewicz, E., Dworak, A. & Shugar, D. (1976) *J. Am. Chem. Soc.* **98**, 367–376
Scawen, M. & Allen, A. (1977) *Biochem. J.* **163**, 363–368
Selvendran, R. R. (1975) *Phytochemistry* **14**, 2175–2180
Selvendran, R. R., Davies, A. M. C. & Tidder, E. (1975) *Phytochemistry* **14**, 2169–2174
Spiro, R. G. & Bhoyroo, V. D. (1974) *J. Biol. Chem.* **249**, 5704–5717
Vercellotti, J. R. & Just, E. K. (1967) *Carbohydr. Res.* **5**, 102–106
Whistler, R. L. & Smart, C. L. (1953) *Polysaccharide Chemistry*, pp. 296–299, Academic Press, New York and London
Yamagishi, T., Matsuda, K. & Watanabe, T. (1976) *Carbohydr. Res.* **50**, 63–74