*Proc. Natl. Acad. Sci. USA*
Vol. 95, pp. 7921–7926, July 1998
Biochemistry

# Arabinogalactan-proteins from *Nicotiana alata* and *Pyrus communis* contain glycosylphosphatidylinositol membrane anchors

JOELIAN J. YOUL, ANTONY BACIC†, AND DAVID OXLEY

Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville, Victoria 3052, Australia

*Communicated by Adrienne Clarke, University of Melbourne, Parkville, Australia, May 5, 1998 (received for review February 17, 1998)*

**ABSTRACT** **Arabinogalactan-proteins (AGPs) are a class of proteoglycans found in cell secretions and plasma membranes of plants. Attention is currently focused on their structure and their potential role in growth and development. We present evidence that two members of a major class of AGPs, the classical AGPs, AGP*Na*1 from styles of *Nicotiana alata* and AGP*Pc*1 from cell suspension cultures of *Pyrus communis*, undergo C-terminal processing involving glycosylphosphatidylinositol membrane anchors. The evidence is that (*i*) the transmembrane helix at the C terminus predicted from the cDNA encoding these proteins is not present—the C-terminal amino acid is $Asn_{87}$ and $Ser_{97}$ for AGP*Na*1 and AGP*Pc*1, respectively; (*ii*) both AGP protein backbones are substituted with ethanolamine at the C-terminal amino acid; and (*iii*) inositol, glucosamine, and mannose are present in the native AGPs. An examination of the deduced amino acid sequences of other classical AGP protein backbones shows that glycosylphosphatidylinositol-anchors may be a common feature of this class of AGPs.**

Arabinogalactan-proteins (AGPs) are a family of high molecular mass proteoglycans typically consisting of <10% protein and >90% carbohydrate. AGPs belong to a family of glycoproteins that also includes extensins, proline/hydroxyproline-rich glycoproteins, and solanaceous lectins (1, 2). The protein component of AGPs is typically rich in Hyp/Pro, Ala, and Ser, and the carbohydrate consists primarily of arabinose and galactose, usually in the form of a type II arabinogalactan with a β-(1→3)-galactan backbone, branched through C(O)6 to 6-linked galactosyl side chains that usually terminate with arabinosyl and/or glucuronosyl residues. AGPs have been found in all tissues of higher plants as components of the extracellular matrix, the plasma membrane, and the cell wall (for reviews see refs. 3–6).

The location of AGPs at the cell surface (7–9) indicates possible functional similarities with classes of animal proteoglycans. Although the function of any one AGP has yet to be determined, there is evidence for roles in plant development, including cell fate, cell proliferation, and cell expansion (10). The importance of these proposed functions has led to a major effort to identify individual AGPs and to determine the structures of their carbohydrate and protein moieties.

The recent cloning of cDNAs encoding the protein backbones of several AGPs (11–14) has revealed that they are a diverse family that can be classified as "classical" (Hyp/Pro, Ser, Ala-rich) and "nonclassical" (4, 5). The cDNAs for the classical AGPs encode proteins with a characteristic domain structure comprising an N-terminal secretion signal that is absent from the mature protein, a central Pro, Ser, Ala-rich domain predicted to be *O*-glycosylated and a C-terminal hydrophobic domain (see Fig. 1). This C-terminal domain is predicted to be a transmembrane helix, which suggests that classical AGPs may be associated with the plasma membrane. However, two classical AGPs (AGP*Na*1 from styles of *Nicotiana alata* and AGP*Pc*1 from cell suspension cultures of *Pyrus communis*) are buffer-soluble and have been isolated from extracellular fluids (11, 13), which may indicate that these AGPs are transiently present on the plasma membrane and are released from the membrane by processing of the C-terminal domain. In view of the proposed roles of AGPs as recognition molecules, such a mechanism could be important for their biological function. We therefore have investigated the potential for C-terminal processing of the AGP*Na*1 and AGP*Pc*1 proteins. We show that the C-terminal transmembrane domain of both of these classical AGPs is absent from the mature protein. Furthermore, ethanolamine was found at the C terminus of the mature AGP protein backbone, which, together with the presence of inositol, glucosamine, and mannose in the native AGP, is evidence of glycosylphosphatidylinositol (GPI) anchoring. C-terminal processing is likely to occur in the endoplasmic reticulum as occurs for GPI-anchored proteins of protozoa, yeast, and mammals (see ref. 15).

## MATERIALS AND METHODS

**Plant Material.** *Nicotiana alata* Link et Otto (self-incompatibility genotype $S_6S_6$) plants were maintained under standard glasshouse conditions and styles (stigma plus style) and were collected and stored at −70°C before use as described (16). *Pyrus communis* (pear) cell suspension was initiated from fruit and maintained as described (13, 17).

**Extraction of AGPs from *Nicotiana alata* Styles.** Buffer-soluble AGPs were extracted from *N. alata* styles as described by Gane *et al.* (16).

**Extraction of AGPs from Pear Cell Suspension Culture.** High molecular mass material was precipitated from cell suspension culture medium with ethanol (four volumes) (13) and then dissolved in 1% (wt/vol) NaCl and centrifuged (10,000 × *g*, 30 min, 4°C). The supernatant was collected, and the AGPs were purified as described below.

**Purification of AGPs.** Most of the contaminating proteins and polysaccharides present in the crude AGP extracts were precipitated by the addition of $(NH_4)_2SO_4$ to 100% saturation at 0°C. After stirring for 30 min at 0°C, the supernatant was collected by centrifugation (10,000 × *g*, 20 min, 4°C), dialyzed extensively against distilled water (molecular mass cut-off 8,000 Da), and freeze-dried. AGPs were purified further by precipitation with β-glucosyl Yariv reagent as described by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



Abbreviations: AGP, arabinogalactan-protein; AQC, 6-aminoquinolyl-*N*-hydroxysuccinimidylcarbamate; ESI-MS, electrospray-ionization MS; GPI, glycosylphosphatidylinositol; Hyp, hydroxyproline; RP-HPLC, reversed-phase HPLC.

†To whom reprint requests should be addressed. e-mail: a.bacic@botany.unimelb.edu.au.

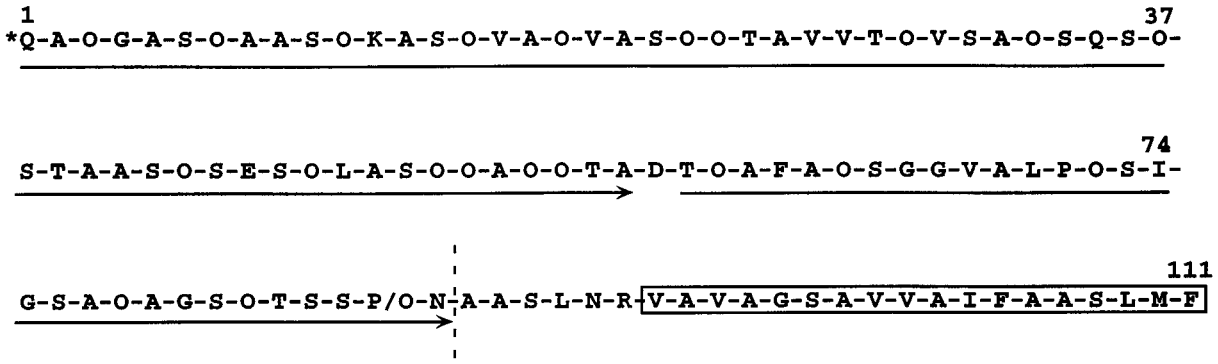


FIG. 1. Amino acid sequence of the AGP*Na*1 protein backbone as deduced from its cDNA sequence (11); Pro (P) residues shown by ESI-MS (Table 1 and Fig. 4) to be modified by posttranslational processing to Hyp (O) residues are indicated. Variability in hydroxylation at $Pro_{86}$, giving rise to the two major forms of peptide H2 (Fig. 2*B* and Fig. 4*A*), is indicated. The N-terminal secretion signal is not shown. *Q, pyroglutamate. The putative transmembrane domain is boxed. The products of selective acid hydrolysis are indicated by arrows, and the site of GPI attachment is indicated by the vertical dotted line.

Gane *et al.* (16). A small portion of *N. alata* and *P. communis* AGPs were fractionated by reversed-phase (RP)-HPLC for monosaccharide analysis (see below).

**Deglycosylation of AGPs by Anhydrous Hydrogen Fluoride.** AGPs (≈10 mg) were deglycosylated by anhydrous hydrogen fluoride based on the method of Mort and Lamport (18) as described by Du *et al.* (11) and then fractionated by RP-HPLC. In the case of the pear AGPs, the fraction containing the AGP*Pc*1 protein was fractionated further by gel filtration and RP-HPLC (peak 1B in figure 1*E* from ref. 13).

**HPLC.** RP-HPLC was carried out as described by Oxley and Bacic (19). Samples were applied to a RP-300 column (2.1 × 100 mm; Applied Biosystems) previously equilibrated in 0.1% trifluoroacetic acid. Proteins/peptides were eluted from the column with a linear gradient of acetonitrile (0–80% in 30 min) in 0.1% trifluoroacetic acid at a flow rate of 0.4 ml/min. Where greater separation was required, a shallower gradient (0–24% acetonitrile in 60 min) was used. Gel filtration chromatography was performed by using a Superdex 75 column (10 × 300 mm) (Pharmacia) eluted with 0.1% trifluoroacetic acid at a flow rate of 0.4 ml/min. In all cases, chromatography was monitored by absorption at 215 and 280 nm.

**N-Terminal Deblocking by Pyroglutamate Aminopeptidase.** Deglycosylated AGPs were deblocked by treatment with pyroglutamate aminopeptidase as described by Du *et al.* (11).

**Protein/Peptide Sequencing.** Automated Edman degradation of peptides was performed by using a Beckman sequencer (model LF 3400) with on line analysis on a Beckman system Gold HPLC.

**Ethanolamine Analysis.** Deglycosylated AGPs were hydrolyzed in 6 M HCl *in vacuo* at 110°C for 18 h and then dried over NaOH. The products were derivatized with 6-aminoquinolyl-*N*-hydroxysuccinimidylcarbamate (AQC) (20) and separated on a Beckman Sepherogel C-18 column (2.1 × 150 mm) at 0.2 ml/min. Buffer A was 40 mM NaAc (pH 6.0) containing 0.1% triethylamine, and buffer B was 80% acetonitrile. AQC derivatives were eluted with a gradient of 10–30% B in 30 min and were detected by using a fluorescence detector (Jasco, Tokyo, Japan) with excitation at 250 nm and emission at 395 nm.

**Monosaccharide Analysis.** Native AGP (≈50 μg) was *N*-acetylated by treatment with acetic anhydride (5%) in saturated $NaHCO_3$ at room temperature for 1 h and then desalted by RP-HPLC. The products then were solvolyzed with 2 M methanolic HCl at 80°C for 16 h, and the reagents were removed by a stream of nitrogen. The products were converted to trimethylsilyl derivatives by reaction with Tri-Sil (Pierce) at 70°C for 10 min, dried under nitrogen, dissolved in dichloromethane, and analyzed by GC-MS on a CP-SIL5 column (Chrompack, The Netherlands).

**Selective Acid Hydrolysis of AGP Protein Backbone.** Selective acid hydrolysis was carried out as described by Inglis (21). Deglycosylated AGPs (4 μg) were hydrolyzed with 13 mM HCl (Pierce, sequencing grade) at 110°C for 2 h *in vacuo*. The products were separated by RP-HPLC.

**Proteolytic Cleavage of Deglycosylated AGPs.** Deglycosylated AGPs (11 μg) were dissolved in 20% 2-propanol/0.1 M $NH_4HCO_3$ (100 μL) and then boiled for 2 min. Trypsin (Sigma, sequencing grade, 0.1 μg) was added to the cooled solution, and the reaction was incubated at 22°C for 16 h. The products were separated by RP-HPLC.

**Electrospray Ionization-MS (ESI-MS).** Fractions from HPLC were analyzed on a Finnigan LCQ ESI-MS. Typical conditions were: heated capillary 200°C and needle voltage 4.5 kV. Samples were introduced by infusion in 50% methanol or 0.1% acetic acid. Spectra were acquired by scanning from 200 to 2000 atomic mass units. MS/MS was performed with the parent ion selected by using a 3-atomic mass unit window and a relative collision energy of 33%. High resolution mass spectra were acquired by using the zoom-scan facility. All molecular masses quoted are average values unless otherwise stated.

## RESULTS

**Purification of the AGP*Na*1 Protein Backbone.** Buffer-extracted material from styles of *Nicotiana alata* was used for the isolation of AGPs. Most of the contaminating proteins and polysaccharides were removed by ammonium sulfate precipitation, and then AGPs were precipitated selectively by the β-glucosyl Yariv reagent. The AGPs were deglycosylated with anhydrous hydrogen fluoride, and the resulting protein backbones were fractionated by RP-HPLC (Fig. 2*A*). Fractions containing the AGP*Na*1 protein were chosen on the basis of previous studies (11).

No N-terminal sequence could be obtained from the AGP*Na*1 protein due to cyclization of the N-terminal Gln residue to pyroglutamate (11). After treatment with pyroglutamate aminopeptidase, to remove N-terminal pyroglutamate, the sequence Ala-Hyp-Gly-Ala was obtained, which matches that of the mature protein (lacking the N-terminal secretion signal) predicted from the cDNA sequence (Fig. 1; ref. 11).

ESI-MS of the AGP*Na*1 protein backbone (Fig. 3) shows that deglycosylation was complete because no clusters of ions differing from each other by 162/z (i.e., a hexose) were observed. However, the pseudomolecular ions were broad because of the variability in the numbers of hydroxylated Pro residues, leading to multiple (unresolved) ions differing by 16 atomic mass units. The data indicate a protein of molecular mass 8,289 ± 10 Da, which is ≈2,000 lower than the value

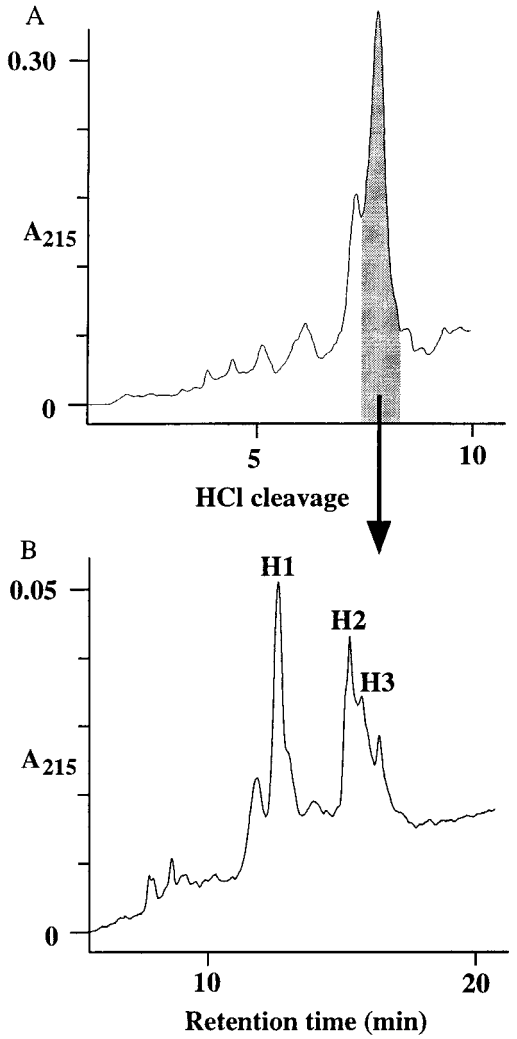


FIG. 2. RP-HPLC separation of the hydrogen fluoride deglycosylated AGP protein backbones (*A*) and the products of selective acid hydrolysis of the protein backbone of AGP*Na*1 (*B*) from *N. alata* styles. The shaded area in *A*, containing the AGP*Na*1 protein backbone, was collected and subjected to HCl hydrolysis, and after separation, fractions H1, H2, and H3, were collected for additional analyses (Table 1; Figs. 3 and 4).

calculated from the cDNA [10,187 (minimum; 0 Hyp/23 Pro) − 10,555 (maximum; 23 Hyp/0 Pro)], indicating that the

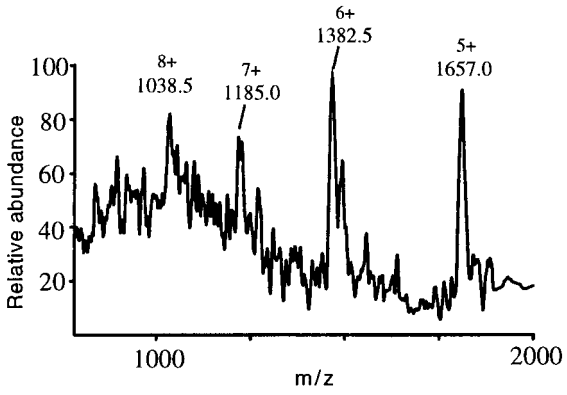


FIG. 3. ESI-MS of the purified AGP*Na*1 protein backbone. Numbers above peaks indicate the charge state of the corresponding ion. The molecular mass was determined for each charge state, and the average was taken. Average molecular mass = 8,289 ± 10 Da.

mature protein had been processed proteolytically. Amino acid analysis of the AGP*Na*1 protein (11) indicated that 85% of the Pro residues were hydroxylated to Hyp. By using these figures, we found that the AGP*Na*1 protein backbone should be C-terminally processed after $Ala_{88}$ (Fig. 1) for the calculated molecular mass (8,276 Da; 21 Hyp/2 Pro) to approach that determined experimentally (molecular mass 8,289 Da). This would require removal of the putative transmembrane domain in the mature protein predicted from the cDNA sequence (Fig. 1). To determine the exact site of C-terminal processing, it was necessary to obtain smaller C-terminal peptides that could be analyzed with greater accuracy by ESI-MS.

**Selective Acid Hydrolysis of the AGP*Na*1 Protein Backbone.** The AGP*Na*1 protein was hydrolyzed with HCl under conditions that selectively cleave at Asp residues and liberate free Asp (21). Because the AGP*Na*1 protein backbone contains only one Asp residue ($Asp_{58}$), two peptide products were expected, i.e., a N-terminal peptide $*Gln_1$-$Ala_{57}$ (where *Gln denotes pyroglutamate) and a C-terminal peptide $Thr_{59}$-$Ala_{88}$ (where $Ala_{88}$ is the putative C terminus of the proteolytically processed protein) (see Fig. 1).

The products of the HCl hydrolysis were separated by RP-HPLC (Fig. 2*B*) and subjected to Edman sequencing and ESI-MS analysis (Table 1; Fig. 4). Fraction **H1** contained a peptide, molecular mass 5,434.0 Da, which matches the calculated value (5, 434.7 Da; 16 Hyp/0 Pro) for the $*Gln_1$-$Ala_{57}$ peptide. The N terminus was blocked to Edman sequencing, which is consistent with this assignment.

ESI-MS of fraction **H2** showed the presence of several species due to variability in the degree of Pro hydroxylation on a single peptide. The molecular mass of the two major species (2,716.2 and 2,732.2 Da, Table 1), are close to, but not identical with, the values calculated (2,711.9 Da; 3 Hyp/4 Pro and 2,727.9 Da; 4 Hyp/3 Pro) for the peptide $Thr_{59}$-$Ala_{88}$. N-terminal sequencing of this fraction (Table 1) confirmed this assignment. To determine the reason for the mass discrepancy between the measured and calculated molecular mass, this peptide was investigated further by MS/MS (Fig. 4*B*) and MS/MS/MS (Fig. 4*C*).

In the MS/MS spectrum of the doubly charged molecular ion at $m/z$ 1359.1 from the AGP*Na*1 C-terminal peptide **H2**, all of the b-series ions, arising from fragmentation in which the charge is retained on the N-terminal side of the cleavage site (22), match the values calculated from the cDNA sequence (Figs. 1 and 4*B*). However, the y″ series of ions, arising by fragmentation in which the charge is retained on the C-terminal side of the cleavage site (22), display the same mass discrepancy as the molecular ion (Fig. 4*B*), indicating there is a modification on the C-terminal end of the peptide. Because of the mass range limitation of the LCQ-MS (2,000 atomic mass units), no fragment ion could be observed beyond $Ala_{79}$, and the precise location of the modification could not be determined. Therefore, a MS/MS/MS experiment was performed in which the $y''_{18}$ fragment ion from the first MS/MS spectrum (Fig. 4*B*) was fragmented further (Fig. 4*C*). Again, the y″ series of ions displayed the same mass discrepancy as before, whereas the b-series ions matched those calculated from the cDNA sequence. In this experiment, the b-series ions were observed to the penultimate residue ($Asn_{87}$), indicating that the C-terminal residue was the site of the anomaly.

It was also apparent from these spectra that this peptide actually contained 5 Hyp/2 Pro and not 3 Hyp/4 Pro, as was used for the molecular mass calculation (Table 1). Thus, the molecular mass of the peptide $Thr_{59}$-$Ala_{88}$ now can be calculated to be 2743.9 Da, which is 27.7 more than the measured molecular mass of 2716.2 Da. Thus, the C-terminal residue cannot be $Ala_{88}$, or indeed any other amino acid, but instead must be a residue with a (anhydro-) molecular mass of 43.3 Da. This result is consistent with the unknown compound at the C terminus of the AGP*Na*1 protein backbone being ethanol-

Table 1. N-terminal sequencing and ESI-MS of the products of selective acid hydrolysis of the AGP*Na*1 protein backbone (Fig. 2)

| Fraction | N-terminal sequence* | Molecular mass, Da | | Peptide |
|---|---|---|---|---|
| | | Observed | Calculated† (Hyp/Pro) | |
| H1 | Blocked | 5,434.0 | 5,434.7 (16/0) | ‡$Gln_1$–$Ala_{57}$ |
| H2 | TOAFAOSGGVAL | 2,716.2 | 2,711.9 (3/4) | $Thr_{59}$–$Ala_{88}$ |
| | | 2,732.2 | 2,727.9 (4/3) | $Thr_{59}$–$Ala_{88}$ |
| H3 | Blocked | 8,261.2 | 8,259.7 (20/3) | ‡$Gln_1$–$Ala_{88}$ |

*Determined by N-terminal protein sequencing. O, Hyp.
†Calculated with number of Hyp/Pro residues chosen to give a molecular mass most similar to the experimentally determined value.
‡Gln, pyroglutamate.

amine. To positively identify the ethanolamine, the AGP*Na*1 protein backbone was subjected to total acid hydrolysis. The products were derivatized with AQC and separated by RP-HPLC under conditions optimized for the separation of AQC ethanolamine from AQC amino acids. A peak was observed that coeluted with an ethanolamine standard but was not present in the negative control (lysozyme). The AQC ethanolamine peak from the AGP*Na*1 protein backbone had the same molecular mass (231 Da) as the derivatized standard by ESI-MS (not shown). This result shows conclusively that ethanolamine is linked to the C terminus of AGP*Na*1 protein backbone after C-terminal proteolytic processing of the protein.

The presence of ethanolamine at the C terminus implies that the protein contains a GPI membrane anchor (23). To provide additional evidence of the presence of a GPI-anchor, monosaccharide analyses were performed on the native AGP to look for the diagnostic sugars inositol, glucosamine, and mannose, which are present in all GPI-anchors. Because the glucosamine is not *N*-acetylated, the glucosamine-inositol linkage is resistant to solvolysis. Therefore, the native AGP*Na*1 (partially purified by RP-HPLC) was *N*-acetylated before methanolysis. GC-MS of the trimethylsilyl derivatives of the products showed the presence of inositol, glucosamine (as *N*-acetylglucosamine), and mannose (<0.1% each) as well as arabinose, galactose, rhamnose, and glucuronic acid.

**Analysis of the AGP*Pc*1 Protein Backbone.** In a series of similar experiments with the deglycosylated protein backbone of AGP*Pc*1 (purified as described in *Materials and Methods*) but by using trypsinolysis rather than selective acid hydrolysis to generate peptides for ESI-MS analysis, we isolated and characterized a C-terminal peptide fragment ($Ser_{74}$-$Ser_{97}$) with ethanolamine attached to $Ser_{97}$, the C-terminal amino acid (data not shown). Furthermore, inositol, glucosamine, and mannose were detected in a sugar analysis of *N*-acetylated native AGP*Pc*1. Thus, the protein backbone of AGP*Pc*1 is processed in the same manner as described for AGP*Na*1.

## DISCUSSION

We present evidence that two classical AGPs, AGP*Na*1 and AGP*Pc*1, undergo C-terminal processing involving GPI anchors. The presence of ethanolamine attached to the C-terminal amino acids ($Asn_{87}$ and $Ser_{97}$ for AGP*Na*1 and AGP*Pc*1, respectively) of the proteins, together with the presence of inositol, glucosamine, and mannose, is compelling evidence that they carry GPI membrane anchors. Supportive evidence is that the AGP*Na*1 and AGP*Pc*1 protein backbones have other features that are present in animal, yeast, and protozoan parasite membrane proteins that are destined to become GPI-anchored. These include an N-terminal secretion signal sequence, C-terminal hydrophobic domain, and specific features around the GPI attachment site required for processing (Fig. 5).

More than 125 GPI-anchored glycoproteins have been identified, and 20 of these have had the GPI attachment site (designated $\omega$) determined experimentally (24). Only Ser, Asn, Ala, Gly, Asp, and Cys are found at the $\omega$ site, and only Ala, Gly, Thr, or Ser is found at the critical $\omega$+2 site (15, 24). The type of amino acid does not appear to be critical at the $\omega$+1 site although it is generally similar to those at the $\omega$ site. The AGP*Na*1 protein ($\omega$ = $Asn_{87}$, $\omega$+1 = $Ala_{88}$, $\omega$+2 = $Ala_{89}$) (see Fig. 1) and AGP*Pc*1 protein ($\omega$ = $Ser_{97}$, $\omega$ +1 = $Gly_{98}$, $\omega$ +2 = $Thr_{99}$) (13) conform to this pattern.

The $\omega$+2 residue of known GPI-anchored glycoproteins usually is followed by a spacer of five to seven amino acids rich in charged and/or Pro residues, which is analogous to the hinge region observed for amino-terminal, signal–sequence cleavage (15). This hinge region is followed by a stretch of 8–20 hydrophobic amino acids (15). The AGP*Na*1 and AGP*Pc*1 proteins contain four and five amino acids before the hydrophobic domain, respectively, but in both only the last residue, Arg, is charged. A hydrophobic domain of 18 amino acids following the potential hinge region also is observed in both the AGP*Na*1 (see Fig. 1) and AGP*Pc*1 (13) proteins predicted from the cDNA clones.

GPI anchors consist of a core glycan (Man$\alpha$1–2Man$\alpha$1–6Man$\alpha$1–4GlcN$\alpha$1-) linked via ethanolamine phosphate to the C terminus of the protein, with a reducing-end glucosamine linked via inositol to a glycerolipid moiety (or occasionally a ceramide) that is embedded in the plasma membrane (23) (see Fig. 5). Because the AGP*Na*1 and AGP*Pc*1 proteins were purified from the buffer-soluble fraction, it is likely that the lipid had been removed. No fatty acids were detected in the monosaccharide analysis of either AGP, suggesting that the lipid moiety may have been cleaved by the action of a phospholipase, possibly as a result of some signaling event. Regulated release of GPI-anchored proteins has been suggested for a number of mammalian cell surface glycoproteins (25). Irrespective of the presence or absence of the lipid moiety, hydrogen fluoride deglycosylation would have removed the GPI anchor, leaving only ethanolamine attached to the C terminus, as was observed.

An examination of the protein backbones predicted from cDNA clones of other putative classical AGPs shows that GPI anchors may be a common feature of this class of AGPs. For example, the putative AGP cDNA clone *Sta 39–3* from *Brassica napus* pollen (26) has a potential GPI attachment site, $\omega$ = Gly, $\omega$+1 = Ser, $\omega$+2 = Ala, four residues before a 17-residue hydrophobic C-terminal domain. Five putative AGP cDNA clones, pCK-H6 from *Gossypium hirsutum* (cotton) (27), pPs ENOD5 from *Pisum sativum* (pea) (28), *Pt*X3H6 and *Pt*X14A9 from *Pinus taeda* (loblolly pine) (29), and *Le*AGP-1 from *Lycopersicon esculentum* (tomato) (30, 31), have $\omega$ = Ser, $\omega$+1 = Gly, $\omega$+2 = Ala before a hydrophobic C-terminal domain (16–20 residues). All of the clones have a basic residue 6–16 residues C-terminal to the predicted $\omega$ site. All of these genes are identified as encoding AGP protein backbones on the basis of homology with the cDNAs, which are known to code backbones of individual AGPs. The characteristics of these putative processing sites are further evidence that they indeed encode AGP protein backbones. It is

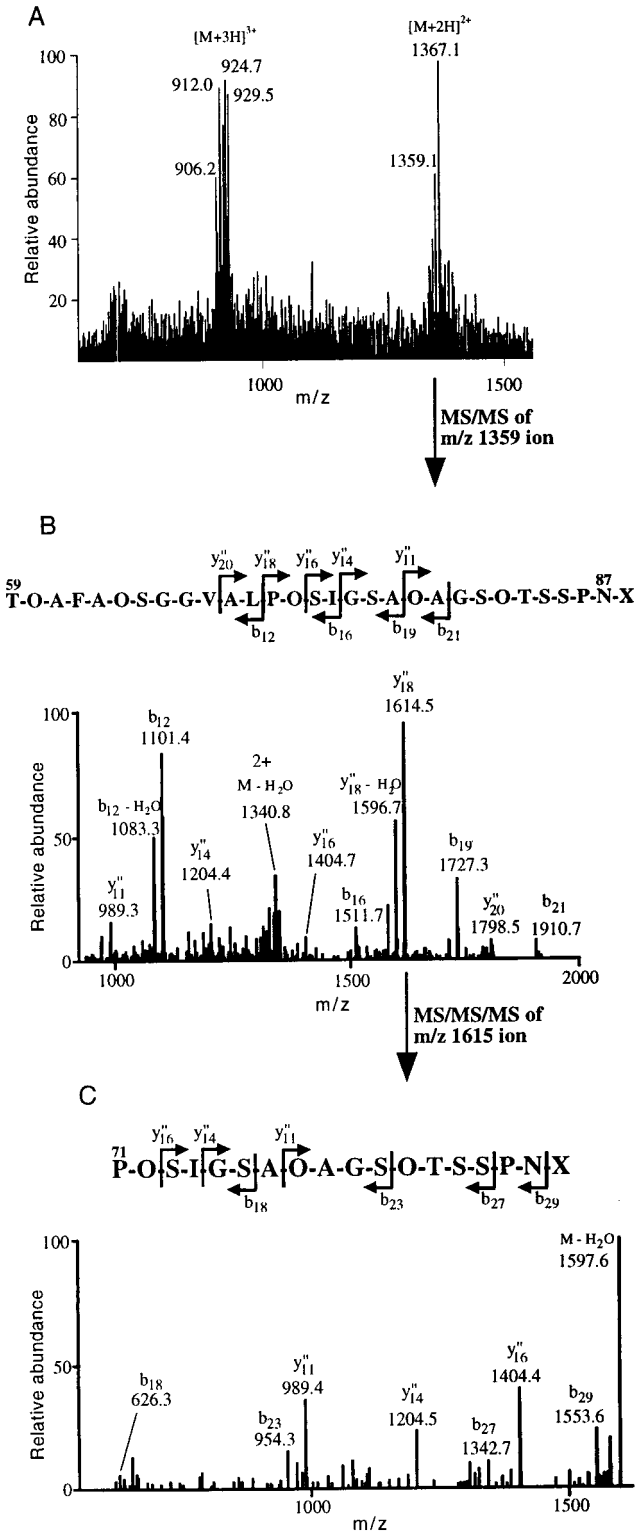


FIG. 4. (*A*) ESI-MS of peptide H2 (Fig. 2*B*) from selective hydrolysis of the AGP*Na*1 protein backbone. (*B*) ESI-MS/MS of the fragment ions produced by collision-induced dissociation of the *m/z* 1359.1-Da ion in *A*. (*C*) ESI-MS/MS/MS of the fragment ions produced by collision-induced dissociation of the $y''_{18}$ fragment ion at *m/z* 1614.5 Da in *B*. Fragment ions (b- and y″-series) assigned are labeled according to Biemann (22) and are indicated on the peptide sequence. X, unknown modification.

particularly interesting to note that the *Le*AGP-1 gene contains an intron that separates the central Hyp-rich glycosylation domain from the C-terminal hydrophobic domain imme-

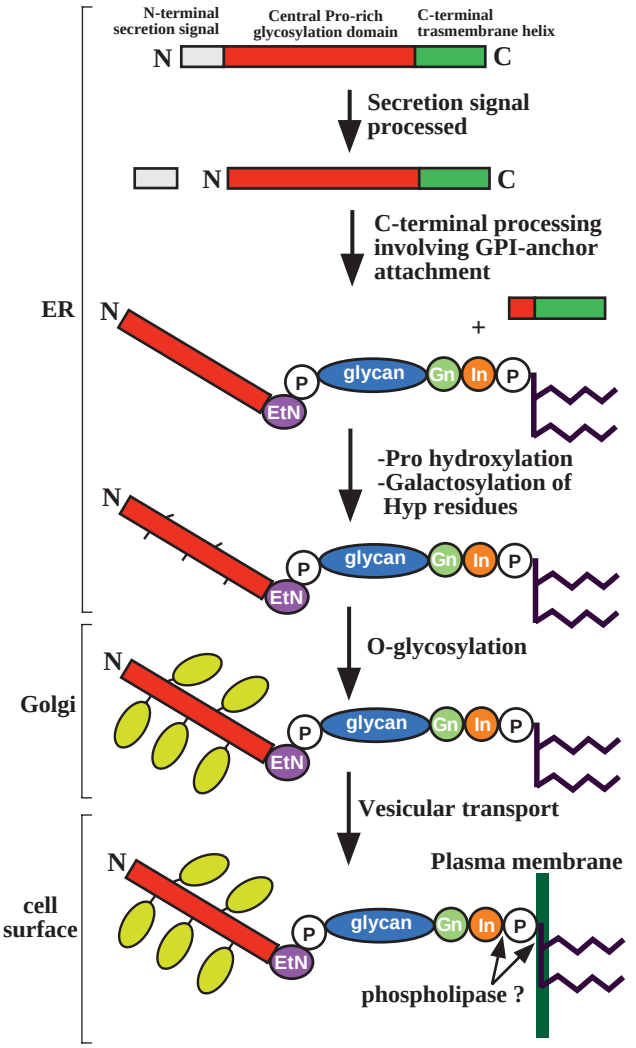


FIG. 5. Schematic representation of the synthesis and processing of classical AGPs. Synthesis of the protein backbone occurs by ribosomes on the endoplasmic reticulum (ER) by vectorial transport into the lumen of the ER. The N-terminal secretion signal is cleaved cotranslationally, and on completion of protein synthesis, the C-terminal transmembrane domain is processed with addition of a pre-formed GPI anchor by a transamidase. The structure of the GPI anchor on AGPs is unknown but is presumed to contain the minimal structure EtN-$PO_4$-Man$\alpha$1–2Man$\alpha$1–6Man$\alpha$1–4GlcN$\alpha$1–6In-$PO_4$ linked to either a glycerolipid or a ceramide. Within the ER, the Pro residues are hydroxylated by prolyl hydroxylase and *O*-glycosylated by a galactosyl transferase. The degree of *O*-glycosylation occurring in the ER is unknown. The AGP then is transported to the Golgi apparatus where *O*-glycosylation to form the type II arabino-3,6-galactan chains is completed and then transported to the cell surface in vesicles. At the cell surface, the AGP is proposed to be released to the extracellular space through the action of a phospholipase (arrows). EtN, ethanolamine; Gn, glucosamine; In, inositol; P, phosphate.

diately preceding the predicted site ($\omega$) for GPI attachment (31).

GPI anchors may be a common feature in classical AGPs, but the nonclassical AGPs that have been studied to date do not have the required features for GPI anchor attachment, although there is evidence for proteolytic processing at both the N and C termini in two nonclassical AGPs (AGP*Pc*2, AGP*Na*2; ref. 14).

There are several recent reports of GPI-anchored proteins in plants (32, 33) and green algae (34) based on mobility shifts on SDS/PAGE and phase separation properties in Triton X-114 after phosphatidylinositol-specific phospholipase C digestion. Of interest, the best characterized GPI-anchored

protein, alkaline phosphatase from the aquatic plant *Spirodela* (35), is resistant to phosphatidylinositol-specific phospholipase C.

We present evidence for the presence of GPI anchors on AGPs, a finding that may help to gain further insights into the function of this widely distributed class of plant proteoglycans. In mammalian, yeast, and protozoan systems, the function of GPI anchors is unclear. They are not limited to a class of proteins but are found linked to many proteins, including membrane regulatory proteins, receptors, adhesion proteins, enzymes, and numerous other cell surface markers of unknown function (23, 36). Of particular interest is evidence that GPI-anchored proteins are specifically transported to the apical surface of polarized cells, suggesting that GPI may act as an apical transport signal. This may explain the very specific temporal and spatial distribution of some cell surface AGPs and may suggest a mechanism by which they could be involved in signaling events that lead to the determination of cell fate.

This work was supported by funds from a Special Research Centre grant from the Australian Research Council. We thank Professor Adrienne E. Clarke for her support and ongoing input into these studies and for her guidance in preparation of this manuscript. We also thank Dr. Ed Newbigin for comments provided during the preparation of the manuscript. We thank Ms. Susan Mau and Mr. Bruce McGinness for the collection of *N. alata* styles and Dr. David McManus of the Co-operative Research Centre for Industrial Plant Biopolymers for the pear suspension cell culture filtrate.

1. Showalter, A. M. (1993) *Plant Cell* **5,** 9–23.
2. Sommer-Knudsen, J., Bacic, A. & Clarke, A. E. (1998) *Phytochemistry* **47,** 483–497.
3. Clarke, A. E., Anderson, R. C. & Stone, B. A. (1979) *Phytochemistry* **18,** 521–540.
4. Bacic, A., Du, H., Stone, B. A. & Clarke, A. E. (1996) in *Essays in Biochemistry*, ed. Apps, D. K. (Portland, London), Vol. 31, pp. 91–101.
5. Du, H., Clarke, A. E. & Bacic, A. (1996) *Plant J.* **9,** 313–323.
6. Nothnagel, E. A. (1997) *Int. Rev. Cytol.* **174,** 195–291.
7. Samson, M. R., Klis, F. M., Corrien, A. M., Stegwee, S. & Stegwee, D. (1983) *Planta* **159,** 322–328.
8. Norman, P. M., Kjellbom, P., Bradley, D. J., Hahn, M. C. & Lamb, C. J. (1990) *Planta* **181,** 365–373.
9. Komalavilas, P., Zhu, J.-K. & Nothnagel, E. A. (1991) *J. Biol. Chem.* **266,** 15956–15965.
10. Knox, J. P. (1995) *FASEB J.* **9,** 1004–1012.
11. Du, H., Simpson, R. J., Moritz, R. L., Clarke, A. E. & Bacic, A. (1994) *Plant Cell* **6,** 1643–1653.
12. Du, H., Simpson, R. J., Clarke, A. E. & Bacic, A. (1996) *Trends Cell Biol.* **6,** 413–416.
13. Chen, C.-G., Pu, Z.-Y., Moritz, R. L., Simpson, R. J., Bacic, A., Clarke, A. E. & Mau, S.-L. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 10305–10309.
14. Mau, S.-L., Chen, C.-G., Pu, Z.-Y., Moritz, R. L., Simpson, R. J., Bacic, A. & Clarke, A. E. (1995) *Plant J.* **8,** 269–281.
15. Udenfriend, S. & Kodukula, K. (1995) *Annu. Rev. Biochem.* **64,** 561–591.
16. Gane, A. M., Craik, D., Munro, S. L. A., Howlett, G. J., Clarke, A. E. & Bacic, A. (1995) *Carbohydr. Res.* **277,** 67–85.
17. Pech, J.-C., Latché, A. & Fallot, J. (1979) *Physiol. Plant* **46,** 260–264.
18. Mort, A. J. & Lamport, D. T. A. (1977) *Anal. Biochem.* **82,** 289–309.
19. Oxley, D. & Bacic, A. (1995) *Glycobiology* **5,** 517–523.
20. Cohen, S. A & Michaud, D. P. (1993) *Anal. Biochem.* **211,** 279–287.
21. Inglis, A. S. (1983) *Methods Enzymol.* **91,** 324–332.
22. Biemann, K. (1990) *Methods Enzymol.* **193,** 455–479.
23. McConville M. J. & Ferguson M. A. J. (1993) *Biochem. J.* **294,** 305–324.
24. Udenfriend, S. & Kodukula, K. (1995) *Methods Enzymol.* **250,** 571–582.
25. Lisanti, M. P., Rodrigeuz-Bolan, E. & Saltiel, A. R. (1990) *J. Membr. Biol.* **117,** 1–10.
26. Gerster, J., Allard, S. & Robert, L. S. (1996) *Plant Physiol.* **110,** 1231–1237.
27. John, M. E. & Keller, G. (1995) *Plant Physiol.* **108,** 669–676.
28. Scheres, B., van Engelen, F., van der Knaap, E., van de Wiel, C., van Kammen, A. & Bisseling, T. (1990) *Plant Cell* **2,** 687–700.
29. Loopstra, C. A. & Sederoff, R. R. (1995) *Plant Mol. Biol.* **27,** 277–291.
30. Pogson, B. J. & Davies, C. (1995) *Plant Mol. Biol.* **28,** 347–352.
31. Li, S.-X. & Showalter, A. M. (1996) *Plant Mol. Biol.* **32,** 641–652.
32. Takos, A. M., Dry, I. B. & Soole, K. L. (1997) *FEBS Lett.* **405,** 1–4.
33. Kunze, M., Riedel, J., Lange, U., Hurwitz, R. & Tischner, R. (1997) *Plant Physiol. Biochem.* **35,** 507–512.
34. Stöhr, C., Schuler, F. & Tischner, R. (1995) *Planta* **196,** 284–287.
35. Morita, N., Nakazato, H., Okuyama, H., Kim, Y. & Thompson, G. A. (1996) *Biochim. Biophys. Acta* **1290,** 53–62.
36. Hirose, S., Knez, J. J. & Medof, E. M. (1995) *Methods Enzymol.* **250,** 582–614.