Post-translational Modifications of Arabinogalactan-peptides of Arabidopsis thaliana

ENDOPLASMIC RETICULUM AND GLYCOSYLPHOSPHATIDYLinosITOL-ANCHOR SIGNAL CLEAVAGE SITES AND HYDROXYLATION OF PROLINE®

Received for publication, July 7, 2004, and in revised form, August 16, 2004
Published, JBC Papers in Press, August 18, 2004, DOI 10.1074/jbc.M407594200

Carolyn J. Schultz‡§, Kris L. Ferguson¶, Jelle Lahnstein‡, and Antony Bacic¶

From the §School of Agriculture and Wine, Waite Agricultural Research Institute, The University of Adelaide, PMB1, Glen Osmond, South Australia 5064, Australia and the ¶Australian Centre for Plant Functional Genomics and Plant Cell Biology Research Centre, School of Botany, The University of Melbourne, Parkville, Victoria, 3010 Australia

We have developed a method for separating the deglycosylated protein/peptide backbones of the small arabinogalactan (AG)-peptides from the larger classical arabinogalactan-proteins (AGPs). AGPs are an important class of plant proteoglycans implicated in plant growth and development. Separation of AG-peptides enabled us to identify eight of 12 AG-peptides from Arabidopsis thaliana predicted from genomic sequences. Of the remaining four, two have low abundance based on expressed sequence tag databases and the other two are only present in pollen (At3g20865) or flowers (At3g57690) and therefore would not be detected in our analysis. Characterization of AG-peptides was performed using matrix-assisted laser desorption ionization time-of-flight mass spectrometry and tandem mass spectrometry protein sequencing. These data provide (i) experimental evidence that AG-peptides are processed in vivo for the addition of a glycosylphosphatidylinositol (GPI) anchor, (ii) cleavage site information for both the endoplasmic reticulum secretion signal and the GPI-anchor signal for eight of the 12 AG-peptides, and (iii) experimental evidence that the Gly-Pro motif is hydroxylated in vivo. Furthermore, we show that AtAGP16 is GPI-anchored despite its unusually long hydrophobic C-terminal GPI-signal sequence. Prior to this work, the GPI-anchor cleavage site for only two plant proteins, NaAGP1 from Nicotiana alata and PgAGP1 from Pyrus communis, had been determined experimentally. Characterization of the post-translational modifications of AG-peptides contributes toward obtaining the complete primary structure of this class of biologically important plant proteoglycans and provides a greater understanding of post-translational modifications of plant proteins.

Arabinogalactan-proteins (AGPs) are proteoglycans of the plant extracellular matrix that are thought to have diverse functions in plant growth and development, including signaling, embryogenesis, and programmed cell death (1–3). They comprise a large family of Arabidopsis thaliana with almost 50 members, when the fasciclin-like AGPs (FLAs) are included (4). FLAs are chimeric AGPs that contain putative cell adhesion domains in addition to AGP domains (5, 6). Twelve Arabidopsis AGPs belong to a subgroup known as AG-peptides. Their encoded proteins have the hallmarks of the “classical” AGPs, namely an N-terminal secretion signal, a central domain rich in Pro/Hyp, Ser, Thr, and Ala, and a C-terminal region that is predicted to direct the addition of a glycosylphosphatidylinositol (GPI) anchor (7, 8). The distinguishing feature of the AG-peptides is that the central Pro-rich domain predicted to contain the entire mature protein backbone, is generally between 10 and 15 amino acid residues (9). One AG-peptide, AAGP24, has a longer protein backbone of 17 amino acids residues (4). It remains unclear whether the functionality for this class of proteoglycans resides in the glycan, the protein, the GPI-anchor, or some combination of these. One of the long-term goals of our research program is to define the complete primary structure of mature AGPs.

AGPs contain large O-linked arabinogalactan (AG) polysaccharides attached to the numerous hydroxyproline (Hyp) residues found throughout the protein backbone (1–3). In some instances, short arabinino-oligosaccharides chains (10–12) are attached to Hyp residues and single Gal residues are linked to Ser residues (6, 13). Modification of the Pro residues in the protein backbone is thought to occur co-translationally as the proteins are inserted into the endoplasmic reticulum (ER) (1). Recently a putative plant prolyl hydroxylase was shown to have hydroxylation activity in insect cells on both pro-collagen and plant-specific protein backbones (14). Because our understanding of the substrate specificity of individual prolyl hydroxylases is limited, sequencing the protein/peptide backbones of AGPs would make a substantial contribution. Other common post-translational modifications in plants and animals, which are not commonly associated with AGPs, include N-linked glycosylation, myristoylation, phosphorylation, and methylation (15).

It is now clear that the broad class of glycan attached to Hyp residues in the protein backbone depends on whether Hyp residues are contiguous or non-contiguous (12). The Hyp contiguity hypothesis (16) states that contiguous Hyp residues (e.g. Ser-Hyp-Hyp-Hyp), as typically found in extensions, are attachment sites for short arabinino-oligosaccharides and that

* This work was supported in part by Australian Research Council Grants A10020017 and DP0343454 (to A. B. and C. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

® The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables IV, V, and VI and Fig. 4.

‡ To whom correspondence should be addressed. Tel.: 61-8-8303-7109; E-mail: carolyn.schultz@adelaide.edu.au.

¶ Supported by Australian Research Council and Grants and Development Research and Development Corporation grant to the Australian Centre for Plant Functional Genomics.

The abbreviations used are: AG, arabinogalactan-protein; AG, arabinogalactan; ER, endoplasmic reticulum; FLAs, fasciclin-like AGPs; β-GlcY, β-glucosyl Yariv reagent; GPI, glycosylphosphatidylinositol; GT, glycosyltransferase; HF, hydrogen fluoride; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MS, mass spectrometry; RP-HPLC, reversed phase-high performance liquid chromatography; Rt, retention time; Hyp, hydroxyproline.
non-contiguous Hyp residues (e.g. Ser-Hyp-Ala-Hyp), as typically found in the AG-peptides, are the attachment sites for AG polysaccharides. Our understanding of the genes encoding the enzymes for the glycosylation of Hyp is limited although it is likely that many different glycosyltransferases (GTs) will be involved (6, 17). It is not clear whether the addition of AG polysaccharides to AGPs occurs one residue at a time, or whether a glycan core is added en bloc as is the case for N-linked glycans. The AG polysaccharide backbone is repetitive with \( \beta-(1 \rightarrow 3) \text{Galp} \) residues (degree of polymerization of approximately seven) separated by a periodate-sensitive linkage, reviewed in Ref. 1. The repetitive nature of the AG polysaccharide backbone provides some support for the \( \text{en bloc} \) addition of the sugars (18).

The finding that AGPs are GPI-anchored (19–22) provided new insights into how AGPs might function as signaling molecules (8). Most Arabidopsis AGPs are predicted to be, at least transiently, attached to the outer surface of the plasma membrane via a GPI anchor. A plasma membrane localization is consistent with the result of immuno localization experiments, reviewed in Ref. 23. Experimental evidence for the GPI anchoring of only a few plant AGPs has been obtained, PcAGP1 from *Pyrus communis* (19), NaAGP1 from *Nicotiana alata* (19), At-AGP10 from *Arabidopsis* (9), and LeAGP1 from *Lycopersicon esculentum* (12, 24, 25). To date, only two plant proteins have had their GPI-anchor cleavage site (\( \omega \)) determined experimentally; NaAGP1 from *N. alata* and PcAGP1 from *P. communis* (19). For many other AGPs the predicted in *vivo* location is based on the presence of a putative signal for GPI-anchor addition based on genomic sequences (4, 26).

With the completion of two plant genome sequencing projects, *Arabidopsis* (27) and *rice*, *Oryza sativa* (28, 29) there is a need to improve the certainty around GPI-anchor prediction. Although the basic structure of the GPI-signal sequence is conserved among a diverse group of organisms, prediction programs developed for animal and protozoan proteins (30) work poorly for the AGPs. Progress for the prediction of plant proteins has recently been made by two groups (26, 31). With a combination of simple *in silico* sequence analytic procedures, gene sequence refinement, and human expert judgment, Duree and colleagues (26) derived a list of 248 potentially GPI lipid-anchored proteins. With the improved genome annotation this approach predicts that there are \( \sim 248 \) GPI-anchored proteins in *Arabidopsis* belonging to at least 15 different gene families (32). The classical AGPs and AG-peptides comprise 12% of the total GPI-anchored proteins, and proteins containing AG-like domains (AG glycomodules) as part of longer proteins, e.g. certain FLAs and lipid transfer proteins, make up a further 28% of predicted GPI-anchored proteins (32).

The big-II Plant Predictor was developed by Eisenhaber and colleagues (31) and is available as an online predictor (mendel.imp.univie.ac.at/sat/gpi/plant_server.html). This predictor includes PcAGP1, NaAGP1, and selected AG-peptides among 219 plant proteins with experimentally verified or hypothetically anticipated GPI anchor modification. This has resulted in the prediction that 187 *Arabidopsis* proteins are GPI-anchored, and 165 of these are the same as predicted by Ref. 32. The discrepancies between the two approaches highlights that the knowledge of the sequence features of the GPI-anchor signal is incomplete and that the big-II Plant Predictor conservatively extrapolates the current knowledge on uncharacterized proteins.

Another new web-based GPI-anchor predictor, “detection of GPI” was developed by D. Buloz and J. Kronegg of the Swiss Institute of Bioinformatics (129.194.185.165/dgpi/DGPI_demo_en.html), however, it does not work well for plant proteins. Although the “detection of GPI” program suggests that both NaAGP1 and PcAGP1 are GPI-anchored, it does not accurately predict the cleavage site for these two AGPs. For example, the experimentally determined GPI cleavage site (\( \omega \)) for PcAGP1 (GenBank™ accession number U14009) is between Ser\( ^{120} \) and Gly\( ^{121} \) (19), whereas the detection of GPI program (Oct 24, 2003) predicts cleavage between Gly\( ^{106} \) and Ser\( ^{107} \).

The major limitation to accurate GPI-anchor prediction in plants is the lack of experimental verification of GPI anchoring, especially with regard to the determination of the cleavage site (\( \omega \)) residue. Here, we describe the separation of a family of deglycosylated AG-peptides from the bulk of the AGPs. The small size of the deglycosylated AG-peptides, 10 to 17 residues, allowed them to be analyzed without protease digestion. Using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF/MS) and tandem mass spectrometry (MS/MS) protein sequencing we determined the precise cleavage site for both N- and C-terminal signals. We also established that all Pro residues can be hydroxylated for Hyp for the eight AG-peptides analyzed. No other post-translational modifications were detected. These data contribute toward our goal of determining the complete primary structure of this functionally important group of plant proteoglycans.

**EXPERIMENTAL PROCEDURES**

*Plant Material—* Wild-type *A. thaliana* plants were used (Columbia-0 strain CS1092, Arabidopsis Biological Research Center, Columbus OH). For isolation of AGPs the plants were grown in liquid Gamborg’s B-5 medium (Inviron), pH 5.0. Approximately 50 seeds were surface sterilized by washing once in 70% ethanol for 2 min, followed by 15 min in 5% sodium hypochlorite, 0.1% Triton-X 100, then 4 washes in ultrapure water and finally placed into 50 ml of media in a 250-ml flask. The flasks were placed on a rotating shaker (\( \sim 120 \) rpm) in the laboratory under continuous fluorescent lighting at \(-22\) °C. The whole seedlings were harvested after 14 days and freeze dried. Five flasks gave \(-1 \) g dry weight of tissue after freeze drying.

*Purification of AG-peptides—* AGPs were extracted from freeze dried, liquid culture-grown seedlings as previously described (9). In brief, following homogenization in liquid nitrogen, detergent-soluble protein was extracted in a buffer (50 mM Tris-Cl, pH 8.0, 10 mM EDTA, 1% Triton X-100, 0.1% \( \beta \)-mercaptoethanol), freeze dried, and resuspended in a small volume of (\( w/v \)) NaCl. AGPs were selectively precipitated using \( \beta \)-glucosyl Yariv reagent (\( \beta \)-GlcY, Biosupplies, Parkville, Australia) and collected by centrifugation. After washing, the \( \beta \)-GlcY was dissociated from the AGPs using dimethyl sulfoxide and sodium dithionite and removed by desalting on a PD-10 column (Sephadex G-25, Amersham Biosciences). The \( \beta \)-GlcY-precipitated AGP sample was chemically deglycosylated without further chromatography as described above. To date, previous studies have shown that \( \beta \)-GlcY-precipitated AGPs contain no Coo-massie Blue-stained proteins on SDS-PAGE (33).

*Deglycosylation of AGPs by Anhydrous Hydrogen Fluoride—* Chemically deglycosylation of AGPs was performed using anhydrous hydrogen fluoride (HF) for 3 h at room temperature, according to the method of Mort and Lamport (34) as previously described (9). The HF was removed under vacuum, then the sample was desalted on a PD-10 column (Sephadex G-25, Amersham Biosciences). The desalted deglycosylated AGP fractions were freeze dried prior to resuspending in \( 500 \) \( \mu \)l of ultrapure water.

*HPLC—* Reversed-phase (RP)-high performance liquid chromatography (HPLC) was performed with a C18 column (Vydac, 218TP52, 5 \( \mu \)m 2.1 \( \times \) 250 mm) attached to a Hewlett-Packard 1090 Liquid Chromatograph with a diode array detector and analyzed using Chemstation software. AGPs were eluted and collected from the column equilibrated in solvent A (0.05% trifluoroacetic acid with a linear gradient of solvent B (0.04% trifluoroacetic acid in 70% acetonitrile): 0 to 30% solvent B in 30 min, then 30 to 100% in 30 min at a flow rate of 0.3 ml/min. Chromatography was monitored by absorption at 214 and 280 nm. A 250-\( \mu \)l aqueous aliquot of the deglycosylated AGP was mixed with 0.05% trifluoroacetic acid and loaded onto the C18 column maintained at ambient temperature (21 °C) to give sample A. The remaining deglycosylated AGP sample (approximately 250 \( \mu \)l) was mixed with an equal volume of 0.05% trifluoroacetic acid and loaded onto the same RP-HPLC column but the column temperature was increased to \( 28 \) °C in an attempt to improve the elution profile to give sample B. Sample B contained \(-80\%\) as much material as compared with sample A based on
Edman Degradation Protein Sequencing—N-terminal protein sequencing was performed by automated Edman degradation on a sequencer (Model LF 3400; Beckman Instruments) with on-line analysis on a Beckman System Gold HPLC.

Matrix-assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry—MALDI-TOF analysis was performed on a Voyager-DE STR MALDI mass spectrometer (Applied Biosystems), equipped with a 337-nm N2 laser. Parent ion masses were measured in reflector/delayed extraction mode, with accelerating voltage of 20 kV, grid voltage of 85%, and a 150-ns delay. Positive polarity was used. 50 scans were averaged per sample, and spectra were subject to 2 point external calibration using the most appropriate standards. Initially, MALDI-TOF was performed with external standards to avoid interference and suppression that can occur with internal standards (35). For sample A, the external standards used were bradykinin, angiotensin I, and adrenocorticotropic hormone peptide-(18–39) with Mₐ = 1060.5692, 1296.6853, and 2465.1989, respectively. For sample B, the internal standards used were bradykinin, angiotensin I, and adrenocorticotropic hormone-(1–17) with Mₐ = 1060.5692, 1296.6853, and 2093.0867, respectively. For both internal and external standards, two-point calibrations were used. The matrix used was a 2,5-dihydroxybenzoic acid-saturated solution in 50% acetonitrile with 0.1% trifluoroacetic acid.

Protein Sequencing by Tandem Mass Spectrometry—Most samples were analyzed on an Applied Biosystems QSTAR® XL system using a protana/new objective tip. MS/MS for most ions was performed with an ion source voltage of 1600 V, collision energy of 28 V, with Q1 in low resolution mode. For ions that were difficult to fragment (e.g., sodium and potassium adducts) the collision energy was increased to 40 V. Other analyses were analyzed on an LCQ mass spectrometer (Thermo-Finnigan, San Jose, CA). Sample (5 μl) was injected onto a C18 column (Vydac, 300 µm internal diameter × 10 cm). The solvents used were 0.5% acetic acid, 5% acetonitrile (solvent A) and 0.5% acetic acid, 60% acetonitrile (solvent B). Peptides were eluted with a 60-min gradient from Ref. 38.

Calculating the Predicted Masses of AG-peptides—To calculate the expected mass of each AG-peptide the expected post-translational modifications were included. The “standard” modifications were based on the (i) removal of the N-terminal signal sequence as predicted by SignalP (36), (ii) removal of the C-terminal signal for GPI-anchor addition (assuming the most likely cleavage site based on the consens data (37)), (iii) addition of an ethanolamine (43.0422 Da) at the N terminus, the mass for pyroglutamate was used because N-terminated GluGlu sequences produce the profiles shown in Fig. 1, A and B. The remaining material (sample B) was chromatographed at a later stage using slightly different conditions, in particular, elevated column temperatures, resulting in minor changes in the elution profiles (Fig. 1C).

Earlier work with the separation of AGPs showed that the peptide backbones of AG-peptides eluted before the classical AGPs and FLAs (5, 9). Thus, the fractions eluting between 16 and 24 min (fractions A1 to A12, Fig. 1, A and B) contain the AG-peptides and the fractions eluting after 27 min (Fig. 1A) contain the protein backbones from the classical AGPs and other proteins containing AG glycomodules such as FLAs (5, 9).

Edman Sequencing Confirms the Presence of AG-peptides—To confirm that the AG-peptides were in fractions A1 to A12, each fraction was sequenced by Edman sequencing. Some fractions contained multiple amino acid residues in each cycle of Edman sequencing suggesting that multiple peptides eluted in each fraction. Because the expected sequence of AG-peptides is known it was generally possible to assign most residues to a single AG-peptide (Table I). AtAGP13, AtAGP14, AtAGP15.2, AtAGP21, AtAGP22, and AtAGP24 were positively identified from the N-terminal Edman sequencing (Table I). Four of the identified AG-peptides, AtAGP13, AtAGP21, AtAGP22, and AtAGP24, have the expected N terminus based on SignalP prediction (36). The presence of AG-peptides in the β-GlcY-
Post-translational Modifications of Arabidopsis AG-peptides

**TABLE I**

| Fraction | Edman sequencea† | AG-peptide |
|----------|------------------|------------|
| A1       | 123456789 Cycle  | AGP15.2   |
| A2       | SEAOA            | AGP22     |
| A3       | SEAOAOS          | AGP15.2   |
| A4       | VEAOAOSOT        | AGP13     |
| A5       | AVDAOAOSO        | AGP14     |
| A6       | VEAOAOSOT        | AGP13     |
| A7       | HEGH            | AGP24     |
| A8       | ATVEOA            | AGP21     |
| A9       | ATVEAOA          | AGP21     |
| A10      | ATVEAOAOSOT      | AGP21     |
| A11      | ATVEAOAOS        | AGP21     |
| A12      | No signal        | Blocked   |

---

a† Amino acid sequence in single letter code. O is for Hyp and X is for undetermined amino acid. Several fractions had multiple residues per cycle. Because the expected sequence of each AG-peptide is known (Table II) it was possible to assign residues to a single AG-peptide and each peptide is shown on a separate line.

- Residues that do not match the indicated AG-peptide are underlined.
- The predicted N terminus for AtAGP15 is QSEAO and is designated AtAGP15.1 (see Table II).
- In cycle 2, there is an apparently modified residue eluting at retention time (Rt) = 10 min, approximately 3.5 min after E.
- The predicted N terminus for AGP14 is VDAO (see Table II).
- In cycle 4, there is an apparently modified residue eluting at Rt = 10 min, approximately 3.5 min after E.
- A in cycle 1 probably belongs to two peptides.
- In cycle 3, there is an apparently modified residue eluting at Rt = 8.8 min, approximately 3.9 min after D.

---

precipitated material provides indirect evidence that AG-peptides are glycosylated with AG polysaccharides because β-GlcY precipitation requires the presence of both carbohydrate and protein (39).

N-terminal sequence was obtained for all fractions except A12. Because fraction A12 has a significant absorbance at A214 (Fig. 1B) and was therefore expected to be abundant, it is likely that the peptide(s) in this fraction have blocked N termini. AtAGP12, AtAGP15.1, AtAGP16, AtAGP20, and AtAGP41 (Table II) are likely to be blocked to Edman sequencing because they are predicted to start with Gln (Q), a residue that is generally cyclized to pyroglutamate. AtAGP41 is a recently identified AG-peptide predicted from a full-length cDNA clone (AY088224). The gene encoding this AG-peptide was not correctly annotated in the December 2003 version of the Arabidopsis genome, but it has been corrected in the February 2004 version of GenBank™.

Surprisingly, the same AG-peptide often eluted in multiple fractions. For example, AGP21 has a unique N-terminal sequence ATVEA that was found in five fractions, A7 to A11 (Table I). This is likely because of cis-trans isomerization of Hyp that can lead to temperature-sensitive conformation changes that result in different retention times on RP-HPLC (40, 41).

Edman sequencing suggested that some of the AG-peptides contained modified amino acid residues. For example, two derivatized amino acid residues were observed in cycle 4 of fractions A9 to A11, whereas all the other cycles showed only one derivatized amino acid. The most abundant residue was the expected E (retention time (Rt) = 6.5 min) of AGP21, but a smaller peak at Rt = 10 min was also observed (data not shown). Because the other cycles contain only a single amino acid residue it is likely that there are two forms of AGP21, one of which contains a modified E in cycle 4. An unknown amino acid residue was also observed in cycle 2 of A3 to A7 as well as cycle 3 of A8.

MALDI-TOF Analysis of AG-peptides—To determine the cleavage sites for the N-terminal ER signal and the C-terminal GPI-anchor signal we used MALDI-TOF MS. External standards were used initially (sample A) to eliminate the possibility of signal suppression that can be observed with internal standards (35). All fractions showed at least one molecular ion, and many fractions showed two or more dominant ions (Fig. 2 and Supplemental Materials Table V). The MALDI-TOF data supported the Edman sequencing data that individual AGPs eluted in multiple fractions, e.g. AGP21 (expected monoisotopic mass of 1218.5853 Da) eluted in fractions A8 (Fig. 2A and Supplemental Materials Table V) and A9 (Fig. 2B and Supplemental Materials Table V). Ions with mass increases representing sodium (21.9819) and potassium (37.9559) adducts, respectively, were occasionally observed (e.g. A12, Fig. 2C), suggesting that salts were not completely removed in all samples (38).

Internal Standards Improves the Accuracy of MALDI—It is expected that the mass accuracy of MALDI-TOF will be less than 50 parts per million (ppm) (35), however, in some cases they were higher than this (e.g. AGP16 in A12 (Fig. 2C) has a Δmass of 120 ppm). One reason for the low accuracy in some fractions is the use of external standards for the calibrations (35). To improve the accuracy, the MALDI-TOF analysis was repeated using internal calibrations. The remaining aliquot of deglycosylated total AGPs was separated by RP-HPLC and a similar elution profile was obtained (Fig. 1C). The fractions from the second sample were numbered B1 to B12. There were minor differences in the elution profile because the column temperature was ~7 °C higher in the second run in an attempt to improve separation (see “Experimental Procedures”). Each fraction from sample B was analyzed by MALDI-TOF analysis (Supplemental Materials Table VI). Internal standards improved the accuracy with the largest Δmass being 32 ppm for AGP16 in B12 (Fig. 2D).

Tandem Mass Spectrometry Confirms That AG-peptides Elute in Several Fractions—Not all of the ions observed by MALDI could be assigned to a known AG-peptide, therefore we used MS/MS sequencing in an attempt to identify the unknown peptides and confirm the assignment of AG-peptides based on MALDI data. Using MS/MS we identified AGP12 (Fig. 3A), a methylated form of AGP13 (Fig. 3B), AGP14 (Fig. 3C), AGP15.1 (Fig. 3D), AGP16 (Fig. 3E), AGP21 in fractions B8b (Fig. 3F), and B10 (Fig. 3G) and a methylated form of AGP21 (Fig. 3H). Methylation of deglycosylated AGPs has been reported and is believed to be an artifact of HF deglycosylation in the presence of methanol (42). Sodium adducts of AGP15.1, AGP16, and AGP21 were also confirmed by MS/MS (data not shown). The presence of y ions consistent with the presence of ethanolamine at the C terminus of the AG-peptides confirmed that the AG-peptides are processed in vivo for the addition of a GPI-anchor.

One MS/MS spectrum, corresponding to a parent ion [M + H]+ of 1167 could not be fully interpreted because of poor fragmentation, but does not appear to represent a known AG-peptide. A partial sequence was obtained (V-O/L-O/L-H) and is too short and ambiguous to identify a single protein (data not shown).
In total, eight of the 12 AG-peptides, AtAGP 12, 13, 14, 15, 16, 21, 22, and 24, were identified by mass spectrometry and/or Edman degradation sequencing. An overview of the peptides identified by each method is shown in Table III. A more detailed summary is provided in Supplemental Materials Fig. 4, highlighting that most AG-peptides eluted in multiple peaks and that one gene can produce two different mature AG-pep-

### Table II

| Name  | AGI locus | Mature backbone sequence | Predicted mass | Number of ESTs |
|-------|-----------|--------------------------|----------------|---------------|
|       |           |                          |                | Mix | S | R | F | T |
| AGP12 | At3g13520 | qTEAOAOSOTS d           | 1159.5118      | 13  | 2 | 1 | 0 | 19 |
| AGP13 | At4g26320 | DAEOAOSOTS d           | 1046.5006      | 4   | 4 | 0 | 0 | 2  |
| AGP14 | At5g56540 | AVDAOAOSOTS s          | 1103.5220      | 5   | 1 | 0 | 0 | 9  |
| AGP15 | At5g11740 | qSEAOAOSOTS g          | 1145.4962      | 12  | 6 | 6 | 2 | 5  |
| AGP16 | At2g46330 | qSLAOAOTS d           | 1113.5428      | 11  | 1 | 2 | 5 | 21 |
| AGP20 | At3g61640 | qSLAOAOSOTS d         | 1129.5377      | 1   | 1 | 0 | 0 | 3  |
| AGP21 | At1g55330 | ATVEAOAOSOTS d         | 1218.5853      | 1   | 1 | 1 | 1 | 7  |
| AGP22 | At5g53250 | HSSSOAOAOTS d         | 1129.5125      | 4   | 0 | 1 | 0 | 6  |
| AGP23 | At3g57890 | AEEVGAOAOGOAS s       | 1243.5806      | 0   | 0 | 0 | 1 | 1  |
| AGP24 | At5g40730 | HEGHHHHAOAOAOGOAS ss  | 1813.8118      | 8   | 1 | 2 | 0 | 10 |
| AGP40 | At5g20865 |叫我AOAOSOGAS ss       | 1293.5632      | 0   | 0 | 0 | 0 | 1  |
| AGP41 | At5g24105 | qSAAOAOTS d           | 1071.4958      | 0   | 0 | 0 | 0 | 2  |

- The locus number given by the Arabidopsis Genome Initiative (27).
- The arrow (↓) denotes the most likely cleavage position for the C-terminal signal. The C-terminal ethanolamine, the only remnant of the GPI-anchor that remains after deglycosylation, is not shown, but is included in the mass calculations. q represents pyroglutamate, a modification common to N-terminal Q residues.
- The predicted mass, \([M + H]^{+}\), was determined taking into account the standard modifications of AGPs including hydroxylation of Pro to Hyp and an ethanolamine residue at the C terminus (see “Experimental Procedures”).
- Based on information at www.tigr.org, 24th February 2004. Mix, mixture of tissues; S, seedlings; R, root; F, flower; T, total.
- Predicted sequence based on SignalP version 1.1 (36).
- Observed sequence.

AtAGP40 was identified by Ref. 32. They also categorized AtAGP29 (At3g01730) as an AG-peptide, but it is not included here because the mature protein backbone is 44-amino acid residues in length and has two domains with no Pro residues (of 14 and 15 residues in length respectively). The predicted monoisotopic mass of AtAGP29 after removal of N- and C-terminal signals is 4273 Da.

**Discussion**

In total, eight of the 12 AG-peptides, AtAGP 12, 13, 14, 15, 16, 21, 22, and 24, were identified by mass spectrometry and/or Edman degradation sequencing. An overview of the peptides identified by each method is shown in Table III. A more detailed summary is provided in Supplemental Materials Fig. 4, highlighting that most AG-peptides eluted in multiple peaks and that one gene can produce two different mature AG-pep-
FIG. 3. Tandem mass spectrometry sequencing of AG-peptides. For the spectra obtained from the QSTAR the ions labeled have a mass accuracy of at least 0.06 Da, and from the LCQ the mass accuracy is 0.5 Da. A, MS/MS spectrum of a singly charged ion of \( m/z \) 1139.58 from sample A7 on the LCQ is identified as AGP12. B, MS/MS spectrum of a doubly charged ion of \( m/z \) 561.13 from sample A6 on the LCQ is identified as a methylated form of AGP13. C, MS/MS spectrum of a doubly charged ion of \( m/z \) 552.66 from sample A5 on the LCQ is identified as AGP14. D, MS/MS spectrum of a doubly charged ion of \( m/z \) 573.0 from sample B7 on the QSTAR-XL is identified as AGP15.1. The ion corresponding to \( b_{10} \) is not shown but was observed. E, MS/MS spectrum of a doubly charged ion of \( m/z \) 557.2 from sample B12 on the QSTAR-XL showing AGP16. F, MS/MS spectrum of a doubly charged ion of \( m/z \) 559.7 from sample B8b on the QSTAR-XL is identified as AGP21. G, MS/MS spectrum of a doubly charged ion of \( m/z \) 609.7 from sample B10 on the QSTAR-XL is identified as AGP21. This AG-peptide is also in sample B8b (F) confirming the MALDI results that the same modified AG-peptide can elute in multiple peaks from RP-HPLC. H, MS/MS spectrum of a doubly charged ion of \( m/z \) 616.7 from sample A10 on the QSTAR-XL is identified as a methylated form of AGP21. *e*, is a methylated E residue; \( e_t \), is ethanolamine.
Post-translational Modifications of Arabidopsis AG-peptides

| Table III | Summary of the identification of AG-peptides |
|-----------|--------------------------------------------|
| Name      | Edman  | MALDI  | MS/MS  |
|           |        | A      | B      | A | B |
| AGP12     | B       | +      | +      | - |  |
| AGP13     | +       | -      | -      | + |  |
| AGP14     | +       | +      | +      | - |  |
| AGP15.1   | B       | +      | +      | - |  |
| AGP15.2   | +       | -      | -      | + |  |
| AGP16     | B       | +      | +      | - |  |
| AGP20     | B       | -      | -      | + |  |
| AGP21     | +       | +      | +      | - |  |
| AGP22     | +       | -      | -      | + |  |
| AGP23     | -       | -      | -      | + |  |
| AGP24     | +       | +      | +      | - |  |
| AGP40     | -       | -      | -      | + |  |
| AGP41     | B       | -      | -      | - |  |

* Sample A was used for Edman sequencing, B, blocked to Edman degradation sequencing; +, observed; −, not observed.

GPI-anchor cleavage sites have been experimentally determined for plants (19). The data presented here show that the eight AG-peptides identified are GPI-anchored based on MS ions that match the mass of C-terminal processed AG-peptides that include an ethanolamine at the C terminus. The ethanolamine is the only remnant of GPI-anchor that remains after chemical (HF) deglycosylation.

Experimental proof for the GPI anchoring of AGP16 is noteworthy. When AtAGP16 was first discovered, it was unclear if it was GPI-anchored because it contained an unusually long C-terminal hydrophobic domain (9). A long GPI signal sequence has been shown to function in animal cell culture systems (45). Because this was only one example, it was difficult to know how general a phenomenon this was. We show here that AtAGP16 is processed in vivo for the addition of a GPI-anchor based on the observation of ions in MALDI spectra of m/z 1113 in fractions A12 (Fig. 2C) and B12 (Fig. 2D). The MALDI-TOF result was subsequently confirmed by MS/MS in both fractions (Fig. 3E and Supplemental Materials Fig. 4). The MALDI-TOF data shows the presence of sodium and potassium adducts for AtAGP16 at 1135.4069 (Fig. 2, C and D) and 1151.3899 (Fig. 2C), respectively, and the sodium adduct was confirmed by MS/MS (data not shown).

The data confirm that the permissible amino acid residues around the cleavage site residue (ω) are similar between plants and animals (37, 46). The cleavage site region, containing residues ω−1, ω, ω+1, and ω+2, are particularly important for recognition by the transamidase complex. The transamidase complex cleaves the GPI-anchor signal and adds the preformed GPI-anchor at the ω residue. The cleavage site regions of the AG-peptide propeptides are TS ↓ DA (AtAGP12, AtAGP13, AtAGP14, and AtAGP21), TS ↓ DG (AtAGP16 and AtAGP22), and AS ↓ SS (AtAGP24), where the arrow symbol (↓) represents the cleavage site. AtAGP15 is present with two different C-terminal cleavage sites TS ↓ GS (AGP15.1), GS ↓ SA (AGP15.2). The presence of ethanolamine at the C terminus of both forms of AtAGP15 confirms that they are the result of alternate cleavage sites rather than subsequent processing (Fig. 3D and Supplemental Materials Fig. 4). The cleavage site observed for AGP15.1 is listed as the most likely site for the big-II Plant Predictor (31), and the site used for AGP15.2 is listed as an alternate site. The two isoforms of AGP15 also differ at their N termini. One possibility is that the sequence near the N terminus of AG-peptides may influence the choice of cleavage site by the transamidase complex. This possibility is consistent with the information suggesting that residues in the mature protein up to ω−11 are important for GPI-anchor addition (31, 47).

The only other plant proteins that have had their cleavage sites determined experimentally have different cleavage site regions in their propeptides to the six confirmed regions in the Arabidopsis AG-peptides. NaAGP1 from N. alata has (P/O)IN ↓ AA (where O = Hyp) and PcAGP1 from P. communis has (P/O,S) ↓ GT (19). These sites are also consistent with the animal data and big-II Plant Predictor.

Prediction of GPI-anchored Proteins—The experimental determination of the cleavage site for GPI-anchor addition for the eight AG-peptides from Arabidopsis provides information for further refining GPI-anchor prediction tools. Among the 12 AGPs analyzed in this work, the big-II Predictor recognizes both the capacity for GPI lipid anchoring and the correct ω-site (as best or second best position) for 10 sequences. Even in the case of the overly long C-terminal signals, AtAGP16 and AtAGP20, the correct ω-site is found by big-II. It is only the built-in length constraint for permisive C termini that leads to rejection of AtAGP16 and AtAGP20 as possible transamidase

Summary of the identification of AG-peptides

| Name | Edman | MALDI | MS/MS |
|------|-------|-------|-------|
|      |       | A     | B     | A | B |
| AGP12 | B   | +     | +     | - |  |
| AGP13 | +   | -     | -     | + |  |
| AGP14 | +   | +     | +     | - |  |
| AGP15.1 | B   | +     | +     | - |  |
| AGP15.2 | +   | -     | -     | + |  |
| AGP16 | B   | +     | +     | - |  |
| AGP20 | B   | -     | -     | + |  |
| AGP21 | +   | +     | +     | - |  |
| AGP22 | +   | -     | -     | + |  |
| AGP23 | -   | -     | -     | + |  |
| AGP24 | +   | +     | +     | - |  |
| AGP40 | -   | -     | -     | + |  |
| AGP41 | B   | -     | -     | - |  |

* Sample A was used for Edman sequencing, B, blocked to Edman degradation sequencing; +, observed; −, not observed.

The program SignalP developed by von Heijne and colleagues (36) is widely used for predicting the targeting of proteins into the ER and the cleavage site for the N-terminal secretion signal. Six of the eight AG-peptides had a single N-terminal sequence that matched the cleavage site predicted by SignalP version 1.1 (36). For AtAGP14 the observed N terminus is, AVDA, suggesting that cleavage occurred at Ala27 rather than at Ala26. The N-terminal sequence of AtAGP14 was supported by both N-terminal Edman sequencing (Table I), MALDI-TOF analysis (Table V), and MS/MS (Fig. 3C). Because proteolysis is more likely to occur than de novo addition of an amino acid, we assume that AtAGP14 is the result of the use of a different cleavage site. The experimentally observed N terminus of AtAGP14 is predicted by a new version of SignalP (version 3.0) available at www.cbs.dtu.dk/services/SignalP/.

AtAGP15 was identified in two forms, one with the predicted N terminus of Gln25 (AtAGP15.1) and the other with an N terminus of Ser24 (AtAGP15.2) (Table II and Supplementary Materials Fig. 4). It is not known whether there are two different cleavage sites for the ER signal peptide, or whether AtAGP15.1 and AtAGP15.2 are both cleaved between Ala22 and Gln23 and in some cases Gln23 is removed by a different protease to give AtAGP15.2. If proteolytic trimming of the AG-peptides is occurring, then the peptide backbone that we have assumed is encoded by AtAGP13, VEOAOSOTS-εt, could in fact be because of proteolytic processing of AtAGP21.

Additional proteolysis of chimeric (non-classical) AGPs has been experimentally observed in PcAGP2 from pear (33) and for the 120-kDa protein from ornamental tobacco (44). For the chimeric AGPs it is not known whether proteolysis occurs in vivo or during extraction of the proteins.

Experimental Verification of GPI Anchoring of AG-peptides—Because it is technically more challenging to identify the C terminus of proteins than the N terminus, only two...
targets. For AtAGP12, the sites Ala\(^{33}\) → Pro\(^{34}\)-Ser\(^{35}\) and Ser\(^{38}\) → Asp\(^{39}\)-Ala\(^{40}\) are predicted by big-II Plant Predictor as the best and second best sites, respectively. Their relative ranking is determined by a volume constraint for the hydrophobic tail (term 20 big-II), a term that must be reconsidered during a revision of the big-II score function. The first site, Ala\(^{33}\) → Pro\(^{34}\)-Ser\(^{35}\), is unlikely to be used in vivo because the remaining peptide may be too short for further processing by the transamidase, because the cleavage of the ER signal gives an N terminus starting at Gln\(^{28}\). The current version of big-II Plant Predictor does not look for and remove the N-terminal ER targeting signal prior to GPI-anchor prediction. This was done intentionally because gene annotation from genomic sequences is not robust and therefore GPI-anchored proteins might be missed if they were incorrectly annotated (31).

As mentioned earlier AtAGP16 has an unusually long hydrophobic C terminus and is not predicted to be GPI-anchored by big-II Plant Predictor. Another AG-peptide, AtAGP20, contains a very similar C terminus to AtAGP16 and it is unclear whether AtAGP20 is GPI-anchored. We did not find conclusive proof that AtAGP20 is GPI-anchored, however, it is likely that it will be because it has the same cleavage site region as AGP16. It only differs by one amino acid in the mature protein and two in the putative signal sequence downstream of the cleavage site. The predicted GPI-signal sequence of AtAGP20 contains one additional Ser, eight residues from the C terminus, and has an Ser to Thr substitution three residues from the end of the C terminus. These differences will not significantly effect the hydrophobicity of the putative signal sequence and therefore are not likely to have a major effect. Recent data using an in vitro system showed that changing a charged residue (Glu) to a hydrophobic residue (Leu) was sufficient to change a GPI-anchor signal sequence to a classical transmembrane protein domain (48). This same report also provided strong evidence that the entire GPI-anchor signal is translocated across the ER membrane, prior to processing by the transamidase complex (48).

There are at least four distinct proteins in the transamidase complex for processing GPI-anchors and each of these is apparently encoded by a single copy gene in Arabidopsis (1, 49). Our results showing that AtAGP16 is GPI-anchored suggests that parameters such as the distribution and/or position of the hydrophobic region with respect to the cleavage site is more important than overall length of the signal for accurate prediction of GPI anchoring. The three-dimensional structure of the ω-1 to ω-11 region is also relevant with extended structures such as a polyproline II conformation, as found in AGPs (18, 50, 51), necessary for transamidase binding (31). Alternatively, the unusually long C-terminal stretch of AGP16 could be preprocessed prior to presentation to the transamidase complex. If big-II Plant Predictor is given a truncated form of AtAGP16 prepeptide (missing six residues at the C terminus), then GPI anchoring is predicted. Site-directed mutagenesis of AtAGP16 and the other AG-peptides could provide a useful system for studying the specificity of ER translocation and transamidase specificity in plants.

Hydroxylation of Pro Residues—In this paper we confirm that all of the Pro residues in the protein backbones can be, and usually are, modified to Hyp. The observation that AtAGP24 exists as a fully hydroxylated form (m/z 1813, Supplemental Materials Table V and Fig. 4) establishes for the first time that the Pro residues in the GP motif can be hydroxylated in vivo. GP motifs are not generally thought to be hydroxylated, as shown by the recent hydroxylation/glycosylation predictions for LeAGP1 from tomato (12). Underhydroxylated forms of AtAGP24 were observed in the MALDI-TOF data, however, the position of the unmodified Pro (rather than Hyp) is not known (Table V and VI). It is possible that the efficiency of hydroxylation of GP is reduced compared with AP. This may be one of the reasons why AtAGP24 was the only AG-peptide that was underhydroxylated.

Variable hydroxylation of Pro residues has been observed for some of the AGPs. In NaAGP1, the Pro residue closest to the GPI-anchor cleavage site is not always hydroxylated (19). The most likely reason for variable hydroxylation is the type of adjacent residue to the Pro residues. Experiments using fusion proteins in tobacco bright yellow (BY2) suspension cultured cells shows that TP and VP are hydroxylated less frequently than AP (52).

Complexity of AG-peptides in vivo—The size and complexity of the AG polysaccharides attached to each Hyp in the Arabidopsis AG-peptides is currently not known. In the current study our focus was on the peptide backbone and all the glycosylation was removed by anhydrous HF to facilitate purification of the peptides. Predictions about the size of AG polysaccharides in Arabidopsis can be made based on analyses of AG and AG-peptides in other plants. The length of AG polysaccharides reported on other AGPs is diverse ranging from 30 sugars/Hyp in gum arabic glycoprotein, up to 150 sugars/Hyp in radish leaf, reviewed in Ref. 39.

Native AG-peptides were first identified in wheat endosperm (53). Protein backbone sequence of the wheat endosperm AG-peptide identified the corresponding gene as a grain softness protein, GSP-1 (42). The purified backbone of GSP-1 is apparently processed to give, YAEVOSAAQATAD, an AG-peptide with a backbone of 1534 Da, containing three Hyp residues. The wheat AG-peptide is expected to have an AG polysaccharide of up to 32 residues in length attached to each Hyp based on the relative contribution of carbohydrate (91%) and protein (9%), and the observed ratio of Gal to Ara residues of 1.5.

More recently Tan et al. (18) showed that a recombinant AGP, (Ala-Hyp)\(^{31}\), expressed in tobacco contains AG polysaccharides ranging from 13 to 36 sugar residues/Hyp, with a median size of 16. The estimated molar ratios of the sugars of (Ala-Hyp)\(^{31}\) was Galp:Araf:GlcUA:Rhap at 10:2.9:2.0:0.5.

There is likely to be considerable diversity in the glycan structures in planta because of different GTs. The specific modifications of glycans can be dependent on either protein backbone sequence or the glycosylation status of neighboring amino acid residues providing an increasing complexity of glycosylation. For example, in animal mucins the α-N-acetylgalactosaminyl transferases have different specificities based on protein context and prior glycosylation status (54, 55).

Peptide differences have been shown to affect glycosylation in plants. Different AG polysaccharides are attached to AP containing fusion proteins compared with TP or VP fusion proteins in tobacco suspension-cultured cells (52). AP and TP containing synthetic AGPs have ~40 mol % Gal and 30 mol % Ara, whereas VP containing AGPs have more Ara with ~40 mol % of both Ara and Gal. The amounts of the other sugars, e.g. glucuronic acid (GlcUA), also differs in the three fusion proteins at 15, 23, and 13 mol %, respectively (52). This suggests that the Arabidopsis AG-peptides may have different polysaccharides even if they are expressed in the same cell types.

Additional heterogeneity in the carbohydrate moieties of AGPs is likely to exist because of tissue-specific differences in the expression of GTs. In radish, some AGPs are fucosylated. A tissue-specific fucosyltransferase has been found that adds fucosyl residues to an AGP-like synthetic acceptor (56). GT specificity has also been inferred from studies of carbohydrate-deficient mutants in Arabidopsis mur1 mutants lack-
ing an enzyme for the synthesis of GDP-i-fucose. The AGPs isolated from roots of mur1 mutants have a different carbohydrate composition and appear to have larger polysaccharide chains (57).

CONCLUSION

Understanding the functional importance of the diversity of AGPs is a major driving force of our research. It is still not known whether the functions of AGPs are because of the protein backbone, the carbohydrate moieties, the GPI-anchor, or some combination of these features. There are reports in the literature of AGPs binding and/or interacting with a diverse array of substrates including pectins (58, 59), receptor-like kinases (60), and flavonol glycosides (61).

Carbohydrate specificity is important for function of animal proteoglycans. For example, only heparan sulfates from fibroblasts bind to basic fibroblast growth factor thereby facilitating their uptake into cells by receptors (62), whereas others, such as the GPI-anchored Dally, bind to and regulate the extracellular movement of signaling peptides (63). The small size of the GPI-anchored AG-peptides makes them excellent targets for studying the importance of post-translational modifications for the function of AGPs in planta.

Acknowledgments—We are indebted to Drs. Frank and Birgit Eisenhaber for their interest in this research and their hospitality in Vienna and for useful comments on the manuscript. We also thank Dr. Anthony Purcell from the Department of Microbiology and Immunology, University of Melbourne, for allowing us to run one of our samples on their QSTAR Pulsar. We thank Dr. Graeme Currie from Applied Biosystems for advice with interpretation of some of the mass spectra.

REFERENCES

1. Gaspar, Y. M., Johnson, K. L., McKenna, J. A., and Bacic, A., and Schultz, C. J. (2001) Plant Mol. Biol. 47, 161–176
2. Majewska-Sawka, A., and Nothnagel, E. A. (2000) Plant Physiol. 122, 3–9
3. Showalter, A. M. (2001) Cell. Mol. Life Sci. 58, 1399–1417
4. Schultz, C. J., Rumsewicz, M. P., Johnson, K. L., Jones, B. J., Gaspar, Y. M., and Bacic, A. (2002) Plant Physiol. 129, 1448–1463
5. Johnson, K. L., Jones, B. J., Bacic, A., and Schultz, C. J. (2003) Plant Physiol. 133, 1911–1925
6. Johnson, K. L., Jones, B. J., Schultz, C. J., and Bacic, A. (2003) in The Plant Cell Wall (Rose, J., ed) pp. 111–154, Blackwell Publishing, Oxford, UK
7. Du, H., Clarke, A. E., and Bacic, A. (1996) Trends Plant Sci. 1, 426–431
8. Schulz, C. J., Gibson, P. O., Ouyang, Y., Ouyang, J. J., and Bacic, A. (1998) Trends Plant Sci. 3, 426–431
9. Schulz, C. J., Johnson, K. L., Currie, G., and Bacic, A. (2000) Plant Cell 12, 1–18
10. Qi, W., Fong, C., and Lamport, D. T. (1993) Plant Physiol. 96, 848–855
11. Goode, L. J., Patel, A., Leykam, J. F., and Kieliszewski, M. J. (2002) Phytochemistry 54, 99–106
12. Zhao, Z. D., Tan, L., Showalter, A. M., Lamport, D. T. A., and Kieliszewski, M. J. (2002) Plant J. 31, 431–444
13. Kieliszewski, M. J., and Shapak, E. (2001) Cell. Mol. Life Sci. 58, 1386–1398
14. Hietala, R., and Myllyharju, J. (2002) J. Biol. Chem. 277, 23965–23971
15. Wilkins, M. R., Gasteiger, E., Gooley, A. A., Herbert, B. R., Molloy, M. P., Binz, P. A., Ou, K. L., Sanchez, J. C., Bairoch, A., Williams, K. L., and Hochstrasser, D. F. (1999) J. Mol. Biol. 289, 645–657
16. Kieliszewski, M. J., and Lamport, D. T. (1994) Plant J. 5, 157–172
17. Wilson, I. B. (2002) Curr. Opin. Struct. Biol. 12, 569–577
18. Tan, L., Qiu, F., Lampert, D. T., and Kieliszewski, M. (2004) J. Biol. Chem. 279, 13156–13161
19. Ouyang, Y., Ouyang, J. J., and Lamport, D. T. (1994) Plant J. 5, 736–751
20. You, J., Bacic, A., and Ouyang, D. (1998) Proc. Natl. Acad. Sci. U.S.A 95, 7921–7926
21. Vteve, J., Yadav, M. P., and Nothnagel, E. A. (1999) J. Biol. Chem. 274, 14724–14733
22. Sherrer, D. J., Prime, T. A., and Dupree, P. (1999) Electrophoresis 20, 2027–2035
23. Ouyang, D., and Bacic, A. (1999) Proc. Natl. Acad. Sci. U.S.A 6, 14246–14251
24. Knox, J. P. (1995) FASEB J. 9, 1004–1012
25. Takos, A. M., Dry, I. B., and Soode, K. L. (2000) Plant J. 21, 43–52
26. Sun, W. X., Zhao, Z. D., Hare, M. C., Kieliszewski, M. J., and Showalter, A. M. (2004) Physiol. Plant. 129, 319–327
27. Borner, G. H., Sherrer, D. J., Stevens, T. J., Arkin, I. T., and Dupree, P. (2002) Plant Physiol. 129, 486–499
28. Arabidopsis Genomics Initiative (2000) Nature 408, 796–815
29. Yu, J., Ku, S. W., Jiang, W., Gong, K. S., Li, S. G., Liu, B., Deng, Y. J., Dai, L., Zhu, Y., Zhang, X. Q., Cao, M. L., Liu, J., Sun, J. D., Tang, J. B., Chen, Y. J., Huang, X. B., Lin, W., Ye, C., Tong, W., Cong, L. J., Geng, J. N., Han, Y. W. L., Li, W. Q., Huang, X. G., Li, W. J., Li, J., Liu, Z. X., Liu, J. P., Qi, Q. H., Liu, S. J., Li, T. T., Wang, X. G., Lu, H., Wu, T. T., Zhu, M. N., P. X., Han, H., Dong, W., Ren, X. Y., Feng, X. L., Cui, P. I., Li, X. H., Wang, X. Y., Zhai, W. X., Xu, X., Zhang, J. H., Xu, J. C., Zhang, R. L., Zheng, X. W., Dong, J. H., Zeng, W. Y., Tao, L. Ye, J., Tan, J., Ren, X. D., Chen, X. W., He, J., Liu, D. F., Tian, W., Cian, C. X., Xia, H. G., Bao, Q. Y., Li, G., Gao, H., Cao, T., Zhao, W. M., Li, P., Chen, W. X., Zhang, Y., Hu, J. F., Liu, S., Yang, J., Zhang, G. Y., Xiong, Q. J., Li, Z. J., Mao, L., Zhu, C. S., Zhu, Z. Chen, R. S., Hao, B. L., Zheng, W. M., Chen, S. Y., Gue, W. J., Liu, S. Q., Tao, M., Zhu, L. H., Yuan, L. F., and Yang, H. M. (2002) Science 296, 79–92
30. Eisenhaber, B., Bork, P., and Eisenhaber, F. (2001) Protein Eng. 14, 17–25
31. Eisenhaber, B., Wildman, P., Schultz, C. J., Birger, G. H. B., Ouyang, D., and Dupree, P. (2003) Plant Physiol. 133, 1691–1701
32. Borner, G. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003) Plant Physiol. 132, 568–577
33. Mau, S.-L., Chen, C.-G., Pu, Z.-Y., Moritz, R. L., Simpson, R. J., Bacic, A., and Clarke, A. E. (1999) Plant J. 18, 289–293
34. Mort, A., and Lamport, D. (2007) Anal. Biochem. 362, 289–309
35. Geomb, J., Mueller, M., Egelhefer, V., Theiss, D., Lehrach, H., and Nordhoff, E. (2002) Anal. Chem. 74, 3915–3923
36. Nielsen, H., Englundre, B., Brunak, S., and von Heijne, G. (1997) Protein Eng. 10, 1–6
37. Udenfriend, S., and Kodukula, K. (1995) Methods Enzymol. 250, 571–582
38. Simpson, K. J. (2003) Plant Cell and Developmental Biology of Arabinogalactan-
Post-translational Modifications of Arabinogalactan-peptides of Arabidopsis thaliana: Endoplasmic Reticulum and Glycosylphosphatidylinositol-Anchor Signal Cleavage Sites and Hydroxylation of Proline

Carolyn J. Schultz, Kris L. Ferguson, Jelle Lahnstein and Antony Bacic

J. Biol. Chem. 2004, 279:45503-45511.
doi: 10.1074/jbc.M407594200 originally published online August 18, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M407594200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/09/01/M407594200.DC1

This article cites 60 references, 24 of which can be accessed free at
http://www.jbc.org/content/279/44/45503.full.html#ref-list-1