The *O*-Hyp glycosylation code in tobacco and Arabidopsis and a proposed role of Hyp-glycans in secretion

Jianfeng Xu [a], Li Tan [a], Derek T.A. Lamport [b], Allan M. Showalter [c], Marcia J. Kieliszewski [a,*]

[a] *Department of Chemistry and Biochemistry, Ohio University, Athens, OH 45701, United States*
[b] *Department of Environmental and Plant Biology, Ohio University, Athens, OH 45701, United States*
[c] *School of Life Sciences, John Maynard Smith Building, University of Sussex, Falmer, Brighton BN1 9QG, UK*

Abstract

Most aspects of plant growth involve cell surface hydroxyproline (Hyp)-rich glycoproteins (HRGPs) whose properties depend on arabinogalactan polysaccharides and arabinosides that define the molecular surface. Potential glycosylation sites are defined by an *O*-Hyp glycosylation code: contiguous Hyp directs arabinosylation. Clustered non-contiguous Hyp directs arabinogalactosylation. Elucidation of this code involved a single species, tobacco (*Nicotiana tabacum*) BY-2 cells. However, recent work suggests species variation, perhaps tissue specific Hyp glycosylation. Thus, the extent to which the Hyp glycosylation code is 'global' needs testing. We compared the ability of distantly related Arabidopsis cell cultures to process putative HRGP glycosylation motifs encoded by synthetic genes. The genes included: repetitive *Ser-Pro*, *Ser-Pro₂*, *Ser-Pro₄* and an analog of the tomato arabinogalactan-protein, *LeAGP-1ΔGPI*. All were expressed as enhanced green fluorescent protein (EGFP) fusion glycoproteins, designated: AtSO-EGFP (O = Hyp), AtSO₂-EGFP, AtSO₄-EGFP and AtEGFP-LeAGP-1ΔGPI, respectively. The Arabidopsis glycosylation patterns were essentially similar to those observed in *Nicotiana*: non-contiguous Hyp residues in AtSO-EGFP were glycosylated exclusively with arabinogalactan polysaccharides while contiguous Hyp in AtSO₂-EGFP and AtSO₄-EGFP was exclusively arabinosylated. Mixed contiguous and non-contiguous Hyp residues in AtEGFP-LeAGP-1ΔGPI were also arabinosylated and arabinogalactosylated consistent with the code. However, slightly more arabinogalactosylated Hyp and less non-glycosylated Hyp in AtEGFP-LeAGP-1ΔGPI than tobacco NtEGFP-LeAGP-1ΔGPI suggested Arabidopsis prolyl hydroxylases have a slightly broader specificity. Arabidopsis Hyp-arabinogalactans differed from tobacco in decreased glucuronic acid content and lack of rhamnose. Yields of the EGFP fusion glycoproteins were dramatically higher than targeted EGFP lacking Hyp-glycomodules. This validates earlier suggestions that the glycosylation of proteins facilitates their secretion.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: Arabidopsis thaliana; Cruciferae; *O*-Glycosylation; Arabinogalactan-protein; Hydroxyproline-rich glycoprotein; (β-D-Galactosyl)₃-Yariv reagent

1. Introduction

Hydroxyproline-rich glycoproteins (HRGPs) expressed at the plant cell surface are of considerable current interest as they are involved in virtually all aspects of plant growth and development, from fertilization and cytokinesis (Hall and Cannon, 2002) to apoptosis and senescence (Nothnagel, 1997; Showalter, 2001; Motose et al., 2004; Lamport et al., 2006). In the broadest sense, HRGPs comprise numerous cell surface proteins that contain glycosylated hydroxyproline. This includes extensins, arabinogalactan-proteins (AGPs), proline-rich proteins, and numerous other protein chimeras that contain HRGP glycomodules. HRGPs are dominated by *O*-Hyp linked arabinosides (Lamport, 1967) and arabinogalactan polysaccharides (Lamport, 1977; Pope, 1977) which define most of the interactive molecular surface. The 'surface codes' are therefore of intrinsic interest.

Recent work with tobacco BY-2 cells supports the Hyp contiguity hypothesis (Fong et al., 1992; Kieliszewski et al., 1992; Kieliszewski and Lamport, 1994) where a simple

code based on peptide sequence directs *O*-Hyp glycosylation: contiguous Hyp blocks, typically the Ser-Hyp$_4$ motif of extensins, are preferred sites of arabinoside addition, while non-contiguous, clustered Hyp residues, typically Ser-Hyp and Ala-Hyp repeats of the AGPs, are preferred sites of arabinogalactan heteropolysaccharide addition. Earlier we tested this hypothesis by expressing synthetic genes encoding repetitive clustered non-contiguous Pro as Ser-Pro and Ser-Pro$_n$ repeats, or analogs of a naturally occurring tomato (*Solanum lycopersicon*) AGP, LeAGP-1. The products were hydroxylated and glycosylated as predicted (Shpak et al., 1999, 2001; Kieliszewski, 2001; Kieliszewski and Shpak, 2001; Zhao et al., 2002; Tan et al., 2003; Held et al., 2004).

Recently, Estévez et al., 2006 expressed the same synthetic gene constructs and an LeAGP-1 analog in Arabidopsis plants. However, the extent of glycosylation seemed to differ from tissue to tissue. For example, Pro residues in extensin Ser-Pro$_4$ motifs were hydroxylated and exclusively arabinosylated in BY-2 cells (Shpak et al., 2001). Interestingly when expressed in Arabidopsis leaves the same motif was only 50–60% hydroxylated with addition of arabinogalactan polysaccharide to some of the Hyp residues. Therefore, to ascertain possible differences in glycosylation codes between distantly related species like tobacco and Arabidopsis, we compared Hyp glycosylation in cultured cells of both species. This direct comparison between simple systems eliminated interpretation of HRGP glycosylation profiles arising from a mixture of cell and tissue types as occurs with analysis of whole plant organs as it eliminates the confounding effect of variable prolyl 4-hydroxylase (P4H) expression. Cell cultures tend to express high activities of P4H and they also provide abundant material for Hyp glycosylation profiling. Thus, here we compared the Hyp glycosylation profiles in Arabidopsis cell suspension cultures with those of similar tobacco cultures. Macromolecules expressed included both naturally occurring AGPs and the products of synthetic genes encoding LeAGP-1ΔGPI (Zhao et al., 2002), (Ser-Pro)$_{32}$ (designated SP), (Ser-Pro$_2$)$_{24}$ (designated SP$_2$), and (Ser-Pro$_4$)$_{18}$ (designated SP$_4$) all expressed as their EGFP fusions.

Finally, we note that addition of an AG glycomodule to EGFP significantly enhances its secretion and thus corroborates the original suggestion that one of the many roles of glycosylation facilitates secretion (Eylar, 1966; Varki, 1993; Borner et al., 2002).

2. Results

Transformation of the Arabidopsis cells with the genes *SP-EGFP*, *SP$_2$-EGFP*, *SP$_4$-EGFP* and *EGFP-LeAGP-1ΔGPI* (a gene encoding AtEGFP-LeAGP-1ΔGPI, which lacked the sequence directing addition of a glycosylphosphatidylinositol (GPI) anchor to LeAGP-1) resulted in several cell lines showing green fluorescence in the culture medium. Lines exhibiting the most fluorescence were cho-

sen for propagation in liquid culture and transgene product characterization. Isolation by hydrophobic interaction chromatography (HIC) and reversed-phase chromatography gave yields of AtSO-EGFP, AtSO$_2$-EGFP, AtSO$_4$-EGFP and AtEGFP-LeAGP-1ΔGPI (O = Hyp) that ranged from 25 to 40 mg/l medium. Cell lines expressing ER (endoplasmic reticulum) targeted, but non-glycosylated EGFP secreted <0.05 mg/l EGFP into the medium.

Amino acid composition analyses indicated that all Pro residues were hydroxylated in the isolated (SO)$_{32}$, (SO$_2$)$_{24}$ and (SO$_4$)$_{18}$ glycomodules of AtSO-EGFP, AtSO$_2$-EGFP and AtSO$_4$-EGFP (hereafter designated AtSO, AtSO$_2$ and AtSO$_4$) (data not shown), and except for Ser and Hyp, there were no other amino acids detected, confirming the identity and purity of these isolated glycomodules. The composition of AtLeAGP-1ΔGPI (after removal of EGFP) closely resembled that of EGFP-LeAGP-1ΔGPI expressed in tobacco (designated NtEGFP-LeAGP-1ΔGPI) after removal of EGFP (Table 1). N-terminal protein sequencing of AtLeAGP-1ΔGPI yielded TGQTOAAAOVGAKAGT-TOOA.... This confirmed the identity and purity of the isolated AtLeAGP-1ΔGPI further corroborated by SDS–PAGE that showed no contaminants (data not shown).

Hyp assays showed that cell walls isolated from leaves and stems of mature Arabidopsis plants contained 0.05% and 0.13% Hyp (dry weight) and the walls of Arabidopsis and BY-2 suspension cultured cells contained 0.3% and 0.5% Hyp, respectively.

AtSO-EGFP was rich in galactose (Gal) and arabinose (Ara) with lesser amounts of glucuronic acid (GlcUA) than

Table 1

Amino acid compositions of AtLeAGP-1ΔGPI, NtLeAGP-1ΔGPI, and the deglycosylated LeAGP-1 polypeptide from tomato compared to those predicted from the cDNA of *LeAGP-1ΔGPI*

Amino acid	Composition (mol%)[a]			
	AtLeAGP-1ΔGPI	NtLeAGP-1ΔGPIa	Tomato LeAGP-1ΔGPI[a]	LeAGP-1ΔGPI cDNA
Hyp	26.3 ± 0.1	25	29	–
Pro	1.9 ± 0.1	3	1	27.5
Asx	4.2 ± 0.2	4	2	3.5
Thr	9.6 ± 0.2	10	10	8.8
Ser	10.5 ± 0.3	10	12	11.7
Gly	3.7 ± 0.2	3	5	3.5
Glx	4.2 ± 0.2	5	3	3.5
Ala	23.3 ± 0.2	22	21	23.5
Val	8.8 ± 0.2	9	9	8.2
Cys	0	0	0	0.0
Met	1.1 ± 0.1	1	0	1.7
Ile	0	0	1	0.0
Leu	1.8 ± 0.2	1	1	1.2
Tyr	0	0	0	0.0
Phe	0	0	1	0.0
His	0	0	0	0.6
Lys	4.6 ± 0.1	4	5	6.4
Arg	0	1	0	0.0
Trp	nd[b]	nd	nd	0.0

Standard deviations represent the measurements from three cell lines.
[a] From Zhao et al. (2002).
[b] Not determined.

its tobacco counterpart, and it contained no rhamnose (Rha). In contrast, AtSO$_2$-EGFP contained only Ara while AtSO$_4$-EGFP contained mainly Ara and only 4% (mol) Gal, presumably as monogalactosyl Ser which is common in extensins (Shpak et al., 2001) (Table 2). The Hyp-glycoside profiles of AtSO-EGFP, AtSO$_2$-EGFP and AtSO$_4$-EGFP showed that each Hyp residue in AtSO-EGFP was arabinogalactosylated, whereas AtSO$_2$-EGFP and AtSO$_4$-EGFP contained only arabinosides and non-glycosylated Hyp (Table 3).

The monosaccharide composition (Table 2) and linkage analysis (data not shown) of AtEGFP-LeAGP-1ΔGPI resembled that of NtEGFP-LeAGP-1ΔGPI but was not identical; while both were rich in Ara (terminal and 5-linked) and Gal (terminal, 3-linked, 6-linked and 3,6-linked) the Arabidopsis fusion proteins AtEGFP-LeAGP-1ΔGPI and AtSO-EGFP contained no Rha and much less GlcUA (all terminal) similar to the wild-type AGPs from Arabidopsis (Fig. 1 and Table 2). The Hyp glycoside profile in Table 3 also showed that AtEGFP-LeAGP-1ΔGPI contained significantly more Hyp-arabinogalactan polysaccharides and less non-glycosylated Hyp than NtEGFP-LeAGP-1ΔGPI (t-test, P = 0.05).
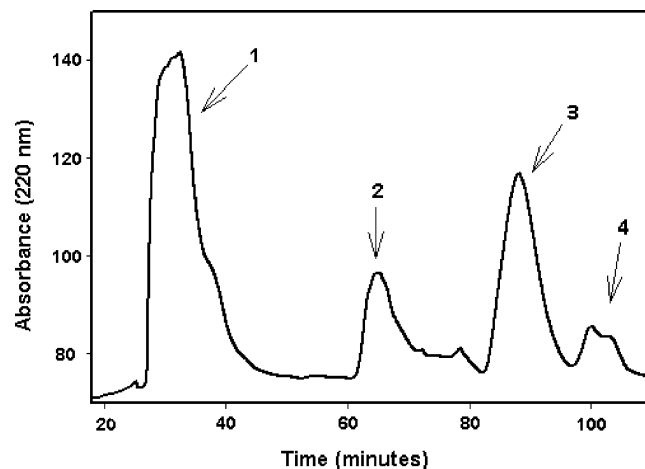

Fig. 1. AGPs from non-transformed wild-type Arabidopsis cells: final stage fractionation via reversed-phase HPLC on PRP-1. The four major UV-absorbing peaks, numbered 1 through 4 above, contained AGPs judging by reaction with (β-D-galactosyl)$_3$-Yariv reagent.

Surprisingly, neither AtEGFP-LeAGP-1ΔGPI nor AtLeAGP-1 (sans EGFP) reacted with (β-D-Galactosyl)$_3$-Yariv reagent at concentrations that reacted with

Table 2
Glycosyl compositions of Arabidopsis wild-type AGP (wtAtAGP pk1, pk2, pk3, and pk4 shown in Fig. 1) fractions and SO-EGFP, SO$_2$-EGFP, SO$_4$-EGFP, and EGFP-LeAGP-1ΔGPI expressed in Arabidopsis and tobacco cells

Sugar residue	Glycosyl composition (mol%)											
	AtSO-EGFP	NtSO-EGFP[a]	AtSO$_2$-EGFP	NtSO$_2$-EGFP[a]	AtSO$_4$-EGFP	NtSO$_4$-EGFP[a]	AtEGFP-LeAGP-1ΔGPI	NtEGFP-LeAGP-1ΔGPI[b]	wtAtAGP pk1	wtAtAGP pk2	wtAtAGP pk3	wtAtAGP pk4
Rha	0	8	0	0	0	0	0	11	Trace	2.8 ± 0.1	0	0
Ara	35.2 ± 0.2	28	100 ± 0.0	100	95.9 ± 0.1	95	44.1 ± 0.1	36	38.2 ± 0.2	36.8 ± 0.0	40.9 ± 0.2	32.3 ± 0.1
Gal	55.8 ± 0.2	45	0	0	4.1 ± 0.3	5	48.8 ± 0.2	39	53.1 ± 0.3	49.1 ± 0.1	48.2 ± 0.3	55.1 ± 0.3
GlcUA	9.0 ± 0.3	19	0	0	0	0	7.1 ± 0.1	14	8.7 ± 0.1	11.3 ± 0.2	10.9 ± 0.1	9.7 ± 0.1
Fuc	0	0	0	0	0	0	0	0	0	0	0	0
Glc	0	0	0	0	0	0	0	0	0	0	0	0
Xyl	0	0	0	0	0	0	0	0	0	0	0	0
Man	0	Trace	0	0	0	0	0	Trace	0	0	Trace	2.9 ± 0.2

Standard deviations represent the measurements from three cell lines.
[a] From Shpak et al. (2001).
[b] From Zhao et al. (2002).

Table 3
Hydroxyproline glycoside profiles of SO-EGFP, SO$_2$-EGFP, SO$_4$-EGFP, and EGFP-LeAGP-1ΔGPI isolated from Arabidopsis and tobacco cells

Hyp-glycoside[a]	Percent of total Hyp							
	AtSO-EGFP	NtSO-EGFP[b]	AtSO$_2$-EGFP	NtSO$_2$-EGFP[b]	AtSO$_4$-EGFP	NtSO$_4$-EGFP[b]	AtEGFP-LeAGP-1ΔGPI	NtEGFP-LeAGP-1ΔGPI[c]
Hyp-PS	100 ± 0.0	100	0	0	0	0	62.5 ± 0.4	56
Hyp-Ara$_4$	0	0	2.1 ± 0.3	4	27.0 ± 0.4	30	1.1 ± 0.2	2
Hyp-Ara$_3$	0	0	50.2 ± 0.6	48	52.7 ± 0.3	49	7.2 ± 0.3	7
Hyp-Ara$_2$	0	0	4.7 ± 0.3	4	8.3 ± 0.5	10	14.1 ± 0.5	17
Hyp-Ara$_1$	0	0	5.2 ± 0.4	3	6.8 ± 0.3	4	6.2 ± 0.3	4
NG-Hyp	0	0	37.8 ± 0.3	41	5.2 ± 0.4	7	8.9 ± 0.5	14

The deviations represent the standard deviation of the measurements from three cell lines.
[a] Hyp-PS, Hyp-polysaccharide; Hyp-Ara$_n$, Hyp-arabinoside$_{1-4}$; NG-Hyp, non-glycosylated Hyp.
[b] From Shpak et al., 2001.
[c] From Zhao et al., 2002.

Table 4
The (β-D-galactosyl)$_3$-Yariv reactivity of the AtSO, AtLeAGP-1ΔGPI, and Arabidopsis wild-type AGPs (wtAtAGP) compared to NtSO and NtLeAGP-1ΔGPI

Sample weight (μg)	Absorbance (420 nm)							
	AtSO	NtSO	AtLeAGP-1ΔGPI[a]	NtLeAGP-1ΔGPI	wtAtAGP pk1	wtAtAGP pk2	wtAtAGP pk3	wtAtAGP pk4
50	0.95	1.22	0.02	0.53	nd[b]	nd	nd	nd
100	1.65	2.18	0.05	1.02	2.14	1.24	2.02	1.52

[a] Measured more than five times from three cell lines at three different laboratories.
[b] Not determined.

NtEGFP-LeAGP-1ΔGPI, NtLeAGP-1ΔGPI, AtSO-EGFP, AtSO, and the isolated Arabidopsis wild-type AGPs (Fig. 1 and Table 4). Neither AtSO$_2$-EGFP nor AtSO$_4$-EGFP reacted with Yariv reagent.

Size fractionated base hydrolysates of AtEGFP-LeAGP-1ΔGPI, NtEGFP-LeAGP-1ΔGPI, Arabidopsis wild-type AGPs, and AtSO-EGFP showed that the Hyp-arabinogalactans of AtEGFP-LeAGP-1ΔGPI and AtSO-EGFP were smaller than those of NtEGFP-LeAGP-1ΔGPI and the Arabidopsis wild-type AGPs; they were also more polydisperse, having two distinct size populations (Fig. 2 and Table 5); the smaller glycans of both AtEGFP-LeAGP-1ΔGPI and AtSO-EGFP had much less glucuronic acid than the larger arabinogalactans (Table 5). The small molecular weight fractions (fractions 30–32) contained mainly arabinose and lesser amounts of galactose.

3. Discussion

The predictive value of the *O*-Hyp glycosylation code is well established based on expression of synthetic HRGP constructs in rapidly growing cell suspension cultures of solanaceous species. However, the same constructs in Arabidopsis plants gave complex glycosylation patterns difficult to interpret (Estévez et al., 2006). This raises questions about conservation of the Hyp glycosylation code both *in planta* and in distantly related species. Complex glycosylation patterns might conceivably arise from the tissue-specific expression and specificity *per se* of glycosyltransferases and prolyl hydroxylases especially considering reported variations in arabinogalactan structure (Himly et al., 2002; Leonard et al., 2005; Nothnagel, 1997). Therefore, we tested the predictive worth of the Hyp glycosylation code in relatively simple but comparable cell cultures of Nicotiana and Arabidopsis. Expression of the test substrates LeAGP-1 analog and other synthetic genes encoding non-contiguous Pro (*AtSP-EGFP*) and contiguous Pro (*AtSP$_2$-EGFP* and *AtSP$_4$-EGFP*) in cell suspension cultures has several advantages: rapid growth of essentially one tissue type favors non-limiting levels of endogenous prolyl hydroxylase and other enzymes involved in HRGP post-translational modifications and gives high yields of the purified fusion glycoproteins. Thus, subsequent biochemical analysis allowed direct comparison of Arabidopsis HRGPs with comparable products from cultured tobacco cells.

Transgenic products AtSO-EGFP, AtSO$_2$-EGFP, AtSO$_4$-EGFP and AtEGFP-LeAGP-1ΔGPI isolated from Arabidopsis cultures gave Hyp-glycoside profiles consistent with the Hyp-contiguity code. Thus, AtSO showed complete hydroxylation of Pro residues and the resulting clustered, non-contiguous Hyp residues were arabinogalactosylated exclusively. On the other hand contiguous Pro residues in AtSO$_2$ and AtSO$_4$ were also completely hydroxylated resulting in contiguous Hyp residues that were exclusively arabinosylated or remained non-glycosylated. Interestingly, however, Arabidopsis cultures yielded AtEGFP-LeAGP-1ΔGPI with somewhat more Hyp arabinogalactan polysaccharide and less non-glycosylated Hyp than its tobacco counterpart, NtEGFP-LeAGP-1ΔGPI (Table 3). This may reflect differences in prolyl hydroxylase specificity rather than glycosyltransferases as follows: the mature LeAGP-1ΔGPI analog expressed earlier in tobacco (Zhao et al., 2002) and expressed here in Arabidopsis encodes 171 amino acids. Eighteen of the 47 Hyp/Pro residues in LeAGP-1ΔGPI are contiguous. The remaining 29 residues predict 62% of the total Pro/Hyp residues are potential Hyp-arabinogalactosylation sites in good agreement with the 63% arabinogalactosylation determined experimentally by the Hyp glycoside profile of AtEGFP-LeAGP-1ΔGPI (Table 3). This is more than the 56% in NtEGFP-LeAGP-1ΔGPI (Zhao et al., 2002) but has a simple explanation: The lower figure in tobacco fits with the exclusion of Lys-Pro and Gly-Pro hydroxylation as observed in known HRGP sequences from the *Solanaceae*, *Gramineae*, and *Leguminosae* (Frueauf et al., 2000; Goodrum et al., 2000; Kieliszewski et al., 1990; Kieliszewski and Lamport, 1994; Smith et al., 1986). Broadening the prolyl hydroxylase specificity in Arabidopsis to include the dipeptide sequence Gly-Pro and perhaps Lys-Pro would account for increased arabinogalactosylation of AtEGFP-LeAGP-1ΔGPI compared with NtEGFP-LeAGP-1ΔGPI and is consistent with the identification of Gly-Hyp in arabinogalactan-peptides (Schultz et al., 2000, 2004). Thus, the Hyp glycosylation data for the Arabidopsis-expressed proteins characterized here fit a rule-based pattern that is secondarily dependent on prolyl hydroxylase specificity.

One should also note that rapidly growing cell cultures of Arabidopsis and tobacco yield relatively Hyp-rich cell walls (0.3% and 0.5% Hyp dry weight, respectively). This indicates
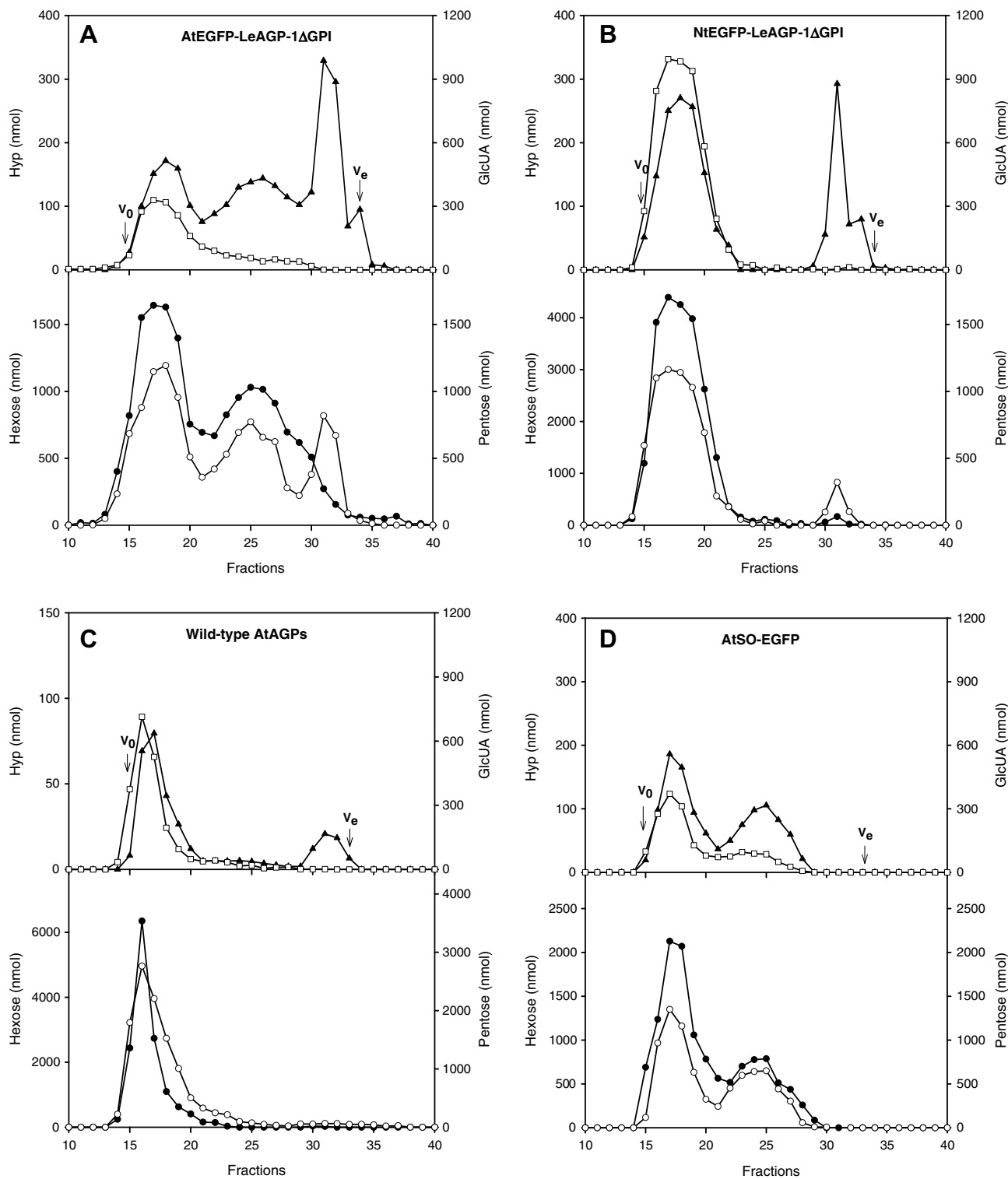
Fig. 2. Size fractionation of base hydrolyzates via Superdex Peptide HP 10/30. AtEGFP-LeAGP-1ΔGPI (A), NtEGFP-LeAGP-1ΔGPI (B), wild-type AtAGPs (C), and AtSO-EGFP (D). Hyp (▲) and hexose (●) left hand ordinates. GlcUA (□) and pentose (○) right hand ordinates. AtEGFP-LeAGP1ΔGPI had a broad glycan size distribution compared to NtEGFP-LeAGP-1ΔGPI and the wild-type AtAGPs. The Hyp-polysaccharides of AtEGFP-LeAGP-1ΔGPI (A) and AtSO-EGFP (D) eluted in two major peaks. Fractions 15–21 contained ∼18–27 glycosyl residues/Hyp and fractions 22–29 contained ∼12–15 glycosyl residues/Hyp (Table 5). In contrast the Hyp-polysaccharides from NtEGFP-LeAGP-1ΔGPI (B) were essentially monodisperse (fractions 15–21) with 23–39 glycosyl residues/Hyp (Table 5). The smaller molecular weight peaks in (A), (B) and (C) fractions ∼30–32 and ∼33–34 corresponded mainly to Hyp-arabinosides and non-glycosylated Hyp, respectively, judging by the pentose and Hyp content, however, the presence of some galactose suggests monogalactosyl serine and/or very small arabinogalactan oligosaccharides occur in these AGPs. The wild-type AtAGPs consisted mainly of Hyp-polysaccharide which eluted very close to the void volume of the peptide column (fraction 16); these glycans ranged from 65 to 142 glycosyl residues/Hyp.

Table 5
The compositions and sizes of glycans isolated from base hydrolysates of AtEGFP-LeAGP-1ΔGPI, NtEGFP-LeAGP-1ΔGPI, wild-type AtAGP (wtAtAGP), and AtSO-EGFP by Superdex peptide column

Component	Amount recovered and mol composition in each fraction (nmol)										
	16	17	18	19	24	25	26	27	30	31	32
AtEGFP-LeAGP-1ΔGPI											
Ara	879	1147	1193	956	693	773	657	625	380	819	671
Gal	1553	1643	1630	1398	955	1031	1016	912	509	272	155
Rha	0	0	0	0	0	0	0	0	0	0	0
GlcUA	275	329	318	257	62	55	40	45	4	0	0
Hyp	100	151	172	149	130	138	144	132	122	329	296
Sugar/Hyp	27	21	18	18	13	13	12	12	7	3	3
NtEGFP-LeAGP-1ΔGPI											
Ara	1100	1164	1141	1030	7	21	10	12	98	320	101
Gal	3911	4393	4250	3979	27	53	38	0	59	167	22
Rha	248	262	257	232	2	6	0	3	0	0	0
GlcUA	844	995	984	938	10	0	5	0	0	0	0
Hyp	147	250	270	256	0	0	0	0	56	292	72
Sugar/Hyp	39	26	23	23	–	–	–	–	3	2	2
wtAtAGP											
Ara	2763	2204	1523	1005	65	54	49	32	55	64	62
Gal	6349	2734	1097	629	21	23	20	18	37	26	0
Rha	0	0	0	0	0	0	0	0	0	0	0
GlcUA	714	526	195	95	16	11	0	6	0	0	0
Hyp	69	80	43	26	5	5	4	0	12	21	18
Sugar/Hyp	142	69	65	65	21	18	18	–	8	4	3
AtSO-EGFP											
Ara	966	1350	1160	631	640	649	442	301	0	0	0
Gal	1236	2128	2070	1058	776	788	511	437	0	0	0
Rha	0	0	0	0	0	0	0	0	0	0	0
GlcUA	275	370	311	127	88	85	49	26	0	0	0
Hyp	98	186	165	94	98	106	83	60	0	0	0
Sugar/Hyp	27	22	22	19	15	14	12	12	–	–	–

The fraction numbers shown below are from Fig. 2. Residue amounts are presented as nmol.

highly active prolyl hydroxylases with specificity for both extensin and AGP substrates and thus ensures complete substrate-specific hydroxylation. On the other hand, incomplete hydroxylation of contiguous proline residues would yield non-contiguous hydroxyproline and concomitant ambiguous glycosylation due to arabinogalactosylation rather than simple arabinosylation. Thus, limited expression of prolyl hydroxylases in Arabidopsis tissues may explain inconsistent processing of some HRGP gene constructions expressed in Arabidopsis plants as incompletely hydroxylated motifs could result in 'non-specific' addition of arabinogalactans to 'unintended' clustered non-contiguous Hyp residues. Indeed, Arabidopsis leaves and stems contain little firmly bound Hyp (0.13% and 0.05% dry weight, respectively) indicating that extensins, and presumably the enzymes for their post-translational modification, are not abundant or highly active in those tissues. In contrast to the under hydroxylation of ectopic extensin-like sequences however, EGFP-LeAGP-1ΔGPI was apparently fully hydroxylated and glycosylated when expressed in Arabidopsis tissues. This is not surprising as orthologs of LeAGP-1 are expressed in all tissues examined and simply indicates the presence of

the appropriate AGP enzymatic repertoire (Sun et al., 2005; Yang and Showalter, 2007).

The ability to purify specific AGPs as their EGFP fusions allows interesting comparisons of structure and properties, including their general ability to form insoluble complexes with the β-D-Glc and β-D-Gal Yariv reagents and the occasional exceptions (Smallwood et al., 1996). Purified AtLeAGP-1ΔGPI was remarkably unreactive with the β-D-Gal Yariv reagent, yet the same polypeptide expressed in tobacco was Yariv-reactive. Thus AtLeAGP-1ΔGPI gave less than 4% of the color yield given by NtLeAGP-1ΔGPI (Table 4). Currently, there is no simple explanation for Yariv-negative AGPs. Lack of Yariv reactivity here is not directly correlated with the lack of Rha in AtEGFP-LeAGP-1ΔGPI as AtSO-EGFP and Arabidopsis wild-type AGPs also lacked Rha (Tables 2 and 5) yet reacted strongly. Nor can the low reactivity of AtEGFP-LeAGP-1ΔGPI be ascribed simply to the smaller size of its Hyp-glycans (18–27 residues) as AtSO-EGFP also has small 18–27 residue Hyp-glycans but reacts with Yariv reagent similar to Arabidopsis wild-type AGPs which have large 65–142 residue glycans (Tables 4 and 5). Other possi-

ble explanations for differences in Yariv reactivity based on uronic acids (less in Arabidopsis than tobacco) or fucose content (Burget and Reiter, 1999; Gaspar et al., 2001) (only traces in minor AGP peaks of Fig. 1) are also inconclusive.

Finally, we note remarkably high yields of purified product (up to 40 mg/l) from Arabidopsis cultures transformed with EGFP-LeAGP-1ΔGPI compared with the low yield of non-glycosylated EGFP; this was barely detectable in the culture medium despite the signal peptide targeting it to "the" secretory pathway (Borisjuk et al., 1999; Shpak et al., 1999). This is consistent with similar observations in planta where significant secretion of EGFP only occurred when expressed as an HRGP fusion Estévez et al., 2006. These observations support the recent suggestion that proteins and cell wall "polysaccharides" (no doubt including AGPs) move through different secretory pathways (Leucci et al., 2007). It also explains such high yields of EGFP-HRGP fusion proteins in agreement with our earlier work where fusion glycoprotein levels in the growth medium reached as much as 110 mg/l (Held et al., 2004; Shpak et al., 1999, 2001; Tan et al., 2003; Xu et al., 2005; Zhao et al., 2002) and with our recent results expressing human interferon or other human proteins as fusion glycoproteins containing two, ten and twenty Ser-Hyp repeats (Xu et al., 2007; Kieliszewski and Xu, 2006). In all instances BY-2 cells expressing the ER targeted but non-glycosylated human protein control gave extremely low secreted yields, whereas the glycoprotein species were abundantly secreted (Xu et al., 2007; Kieliszewski and Xu, 2006). Thus, we agree with Borner et al. (2002) who, in light of the widespread occurrence of only a few AGP glycomodules in many secreted proteins, suggested that arabinogalactosylation helps target proteins to the cell surface. However, it is clear that this likely role also extends to Hyp arabinosylation, and perhaps to protein glycosylation in general as originally proposed more than 40 years ago by Eylar (1966). Examples of secreted proteins known to contain relatively few Hyp-glycomodules include glycerophosphodiesterase 1 (Borner et al., 2002), many fasciclins (Johnson et al., 2003; Tan and Kieliszewski, unpublished data), Zea expansin (Li et al., 2003), horse radish peroxidase (Shannon et al., 1966; Liu, 1971), systemin precursor (Pearce and Ryan, 2003), phytocyanin (Tan and Kieliszewski, unpublished data), stellacyanin (Mann et al., 1992), and umecyanin (van Driessche et al., 1995). Indeed, we found that addition of a single glycosylation site to human growth hormone expressed in BY-2 cells enhanced secretion nearly 100-fold (Xu and Kieliszewski, unpublished data).

To conclude, we note that yields of secreted transgenic proteins expressed in plant hosts are generally so low as to make production impractical (Gaume et al., 2003). However, we can now use the Hyp glycosylation code to design fusion glycoproteins with added HRGP glycan substituents. This ensures greatly enhanced secreted yields hence of practical significance to the biotechnology industry.

4. Experimental

4.1. General experimental procedures

Neutral sugars were analyzed as alditol acetate derivatives by GC (with EC detector) equipped with a Hewlett-Packard HP-5 column (crosslinked 5% PH ME Siloxane, 30 m × 0.32 mm × 0.25 μm) programmed from 130 to 177 °C at 1.2 °C/min (Bhatti et al., 1970). Myo-inositol was used as internal standard. Uronic acids were assayed by the colorimetric method based on reaction with *m*-hydroxydiphenyl using D-glucuronic acid as standard (Blumenkrantz and Asboe-Hansen, 1973). Hexose and pentose were assayed using the method described by Dische (1962). Sugar linkage and corroborative composition analyses were performed at the Complex Research Center, University of Georgia as described earlier (Merkle and Poppe, 1994).

Amino acid compositions of the transgenic glycoproteins were determined after chymotryptic cleavage and then removal of the EGFP tag by reversed-phase HPLC on a Beckman Gold System (Beckman Instruments Inc., CA). Samples were subjected to hydrochloric acid hydrolysis and subsequent PITC (phenylisothiocyanate) derivatization (Bergman et al., 1986). Protein sequence was determined at the Michigan State University Macromolecular Facility. Hydroxyproline (Hyp) was assayed colorimetrically using methods described earlier (Kivirikko and Liesmaa, 1959).

4.2. Plant materials and culture methods

Non-transformed *Arabidopsis thaliana* cells were maintained at room temperature on solid NT-1 medium composed of Murashige–Skoog (MS) basal salts (Murashige and Skoog, 1962), 180 mg/l KH_2PO_4, 100 mg/l myo-inositol, 1 mg/l thiamine HCl and 0.2 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D). The non-transformed stock cultures were also maintained in liquid culture at room temperature (80 ml NT-1 medium in 250 ml Erlenmeyer flasks), shaken at 90 rpm on an Innova gyrotary shaker (New Brunswick Scientific, Edison, NJ). Cells used for transformation experiments were grown in the dark, whereas transformed cells were exposed to general laboratory lighting. Transformed Arabidopsis cells were grown in Schenk–Hildebrandt (SH) medium (Schenk and Hildebrandt, 1972) containing 0.4 mg/l 2,4-D and 200 mg/l kanamycin (Sigma, St. Louis, MO). Subculture was carried out every week with a 2% (v/v) inoculum density.

4.3. Plant cell transformation and selection

The construction of the following four *pBI121*-based plant transformation plasmids were described previously (Shpak et al., 1999, 2001; Zhao et al., 2002). The plasmids pSS^{tob}-SP-EGFP (Shpak et al., 1999), pSS^{tob}-SP$_2$-EGFP and pSS^{tob}-SP$_4$-EGFP (Shpak et al., 2001) encoded a

tobacco extensin signal sequence (SS[tob]), and 32 Ser-Pro repeats, 24 Ser-Pro$_2$ repeats and 18 Ser-Pro$_4$ repeats, respectively, which were fused to enhanced green fluorescent protein (EGFP). The plasmid *pSS[tom]-EGFP-LeAGP-1ΔGPI* encoded the tomato LeAGP-1 signal sequence (SS[tom]) and EGFP fused to LeAGP-1 lacking its GPI anchor signal sequence (Zhao et al., 2002) (GenBank accession numbers: X99147, EU434000−EU434002). These four plasmids were introduced into *Agrobacterium tumefaciens* strain LBA4404 by the freeze–thaw method (An et al., 1988) and Arabidopsis cells were transformed as described earlier (Forreiter et al., 1997) with some modifications. Briefly, transformed bacteria were grown overnight in LB medium supplemented with 50 mg/l kanamycin and 40 mg/l streptomycin. Bacteria were collected by centrifugation at 1900*g* for 10 min and resuspended in the same amount of NT-1 medium. Arabidopsis cells (10 ml) collected 4 days after subculture were incubated with 5 ml of the transformed *Agrobacterium* suspension in a Petri dish at 25 °C in the dark with gentle agitation. After 2 days, the cells were washed with fresh NT-1 medium for four times in 50 ml tubes, then spread on plates containing solid NT-1 medium supplemented with 100 mg/l kanamycin and 250 mg/l timenten (SmithKline Beecham, PA). The plates were incubated at room temperature in the dark. After 2–3 weeks of selection for kanamycin resistance, calli were transferred to fresh plates.

4.4. Extraction and isolation

4.4.1. Cell wall isolation

Three grams of Arabidopsis leaves or stems harvested from mature plants, or suspension culture walls were frozen in liquid N_2 and ground to a powder. The ground tissue was washed ten times with 12 ml of 0.5 M K–Pi buffer (pH 7.0) and 10 times with 12 ml of H_2O, the pellet collected by centrifugation at 3000*g* for 5 min. Then 12 ml of $CHCl_3$–MeOH (1:1) were added to the walls, and the mixture was shaken vigorously overnight on a rotary shaker. The $CHCl_3$–MeOH (1:1) was removed by centrifugation and the cell walls were washed 10 times with 12 ml Me_2CO, and then dried overnight. Five hundred to seven hundred milligrams of walls was hydrolyzed with 1 ml 6 N HCl at 105 °C for 18 h. The hydrolysate was dried under N_2 gas at room temperature. The residues were tested for Hyp as described above.

4.4.2. Isolation of transgenic glycoproteins and removal of EGFP

The EGFP fusion glycoproteins were isolated from culture medium by hydrophobic-interaction chromatography (HIC) followed by reversed-phase chromatography on PRP-1 as described earlier (Tan et al., 2003; Zhao et al., 2002). The EGFP domain was removed from the fusion glycoproteins by chymotryptic cleavage (substrate to enzyme ratio, 100:1) (Zhao et al., 2002) and AtLeAGP-1ΔGPI, AtSO, AtSO$_2$ and the AtSO$_4$ glycomodules were subsequently isolated by reversed-phase HPLC chromatography (Zhao et al., 2002).

4.4.3. Isolation of native AGPs from non-transformed Arabidopsis suspension cultured cells

Non-transformed Arabidopsis cells (*Columbia*) were grown in NT-1 medium as suspension cultures at room temperature (400 ml NT-1 medium in 1000 ml Erlenmeyer flasks), shaken at 90 rpm on an Innova gyrotary shaker (New Brunswick Scientific, Edison, NJ). The culture medium was harvested after 8–10 days growth by filtration through a sintered funnel. The medium was concentrated by rotoevaporation, then dialyzed against water and freeze-dried. The dried medium was redissolved at 10 mg/ml in 20 mM Tris–HCl (pH 7.5) and fractionated on a DEAE column (16 mm i.d. × 700 mm, Amersham Pharmacia Biotech) equilibrated in 20 mM Tris–HCl. The column was eluted at a flow rate of 1.0 ml/min with a 400 ml linear gradient consisting of 0–100% 20 mM Tris–HCl containing 2 M NaCl. The column eluate was monitored for absorbance at 220 nm and the absorbing peaks collected, dialyzed against water and then freeze-dried. The Yariv-positive DEAE fractions were pooled, dialyzed and then loaded onto a semi-preparative Superose-12 column (16 mm i.d. × 500 mm, Amersham Pharmacia Biotech) equilibrated and isocratically eluted in 200 mM sodium phosphate buffer (pH 7) containing 0.05% sodium azide. The eluate was assayed for AGPs by reaction with the Yariv reagent, and the Yariv positive fraction was pooled then further fractionated on a Hamilton PRP-1 reverse phase column (7 mm × 305 mm, Hamilton Co., Reno) equilibrated in solvent A (0.1% TFA, aq.) and eluted with a 100 min gradient of 0–50% solvent B (0.1% TFA, 80% ACN).

4.5. Precipitation with (β-D-galactosyl)$_3$-Yariv reagent

We synthesized (β-D-galactosyl)$_3$-Yariv reagent (Yariv et al., 1962) and determined its reactivity with AtSO (EGFP removed), AtLeAGP-1ΔGPI (EGFP removed), NtLeAGP-1ΔGPI (EGFP removed) and Arabidopsis wild-type AGPs (wtAGP). Fifty and hundred micrograms samples were dissolved in 300 μl of H_2O then an equal volume of (β-D-galactosyl)$_3$-Yariv reagent (1 mg/ml in 2% NaCl aqueous solution) was added and the solutions were allowed to stand for 1 h at room temperature before pelleting the resulting precipitate in a micro-centrifuge. The pellets were washed with H_2O, then dissolved in 0.1 N NaOH and the absorbance measured at 420 nm (Jermyn and Yeow, 1975). As we did not initially believe AtLeAGP-1 was unreactive with the Yariv reagent, these experiments were repeated by five experimenters in three independent labs. The result from one set of experiments is presented in Table 4.

4.6. Hydroxyproline glycoside profiles

Samples (∼10 mg) were dissolved at 5 mg/ml in 0.44 N NaOH and hydrolyzed at 105 °C for 18 h, cooled on ice,

neutralized with ice-cold 1 M H_2SO_4, and then lyophilized. Hyp-glycoside profiles were determined as described earlier (Lamport and Miller, 1971).

4.7. Superdex peptide gel permeation chromatography of fusion glycoprotein base hydrolysates

Samples (~10 mg) were hydrolyzed in 1 ml 0.44 N NaOH at 105 °C for 18 h, followed by cooling on ice, then neutralization with 1 N HCl. The neutralized hydrolysates were applied to a Superdex Peptide HP 10/30 column (Amersham Pharmacia Biotech, Piscataway, NJ) equilibrated in 20% ACN aqueous solvent and eluted isocratically at a flow rate of 0.3 ml/min. Fractions (0.6 ml) were collected, lyophilized and assayed colorimetrically for Hyp, hexoses, pentoses and uronic acids. Some fractions were also assayed for Ara, Rha and Gal compositions by gas chromatography as above.

Acknowledgements

References

An, G., Ebert, P.R., Mitra, A., Ha, S.B., 1988. Binary vectors. In: Gelvin, S.B., Schilperoort, R.A. (Eds.), Plant Mol Biol Manual. Martinus Nijhoff, Dordrecht, pp. 1–19.

Bergman, T., Carlquist, M., Jornvall, H., 1986. Amino acid analysis by high performance liquid chromatography of phenylthiocarbamoyl derivatives. In: Wittmann-Liebold, W. (Ed.), Advanced Methods in Protein Microsequence Analysis. Springer-Verlag, Berlin, pp. 45–55.

Bhatti, T., Chambers, R.E., Clamp, J.R., 1970. The gas chromatographic properties of biologically important N-acetylglucosamine derivatives monosaccharides disaccharides trisaccharides tetrasaccharides and pentasaccharides. Biochim. Biophys. Acta 222, 339–347.

Blumenkrantz, N., Asboe-Hansen, G., 1973. New methods for quantitative determination of uronic acids. Anal. Biochem. 54, 484–489.

Borisjuk, N., Borisjuk, L.G., Logendra, S., Petersen, F., Gleba, Y., Raskin, I., 1999. Production of recombinant proteins in plant root exudates. Nat. Biotechnol. 17, 466–469.

Borner, G.H.H., Sherrier, D.J., Stevens, T.J., Arkins, I.T., Dupree, P., 2002. Prediction of glycosylphosphatidylinositol (GPI)-anchored proteins in Arabidopsis: a genomic analysis. Plant Physiol. 129, 486–499.

Burget, E.G., Reiter, W.D., 1999. The mur4 mutant of Arabidopsis is partially defective in the de novo synthesis of uridine diphospho L-arabinose. Plant Physiol. 121, 383–389.

Dische, Z., 1962. Color reactions of carbohydrates. In: Whistler, R.L., Wolfrom, M.L. (Eds.), Methods in Carbohydrate Chemistry. Academic Press, New York, pp. 475–514.

Estévez, J.M., Kieliszewski, M.J., Khitrov, N., Somerville, C., 2006. Characterization of synthetic hydroxyproline-rich proteoglycans with

Arabinogalactan protein and extensin motifs in Arabidopsis. Plant Physiol. 142, 458–470.

Eylar, E.H., 1966. On the biological role of glycoproteins. J. Theor. Biol. 10, 89–113.

Fong, C., Kieliszewski, M.J., de Zacks, R., Leykam, J.F., Lamport, D.T.A., 1992. A gymnosperm extensin contains the serine-tetra-hydroxyproline motif. Plant Physiol. 99, 548–552.

Forreiter, C., Kirschner, M., Nover, L., 1997. Stable transformation of an Arabidopsis cell suspension culture with firefly luciferase providing a cellular system for analysis of chaperone activity in vivo. Plant Cell 9, 2171–2181.

Frueauf, J.B., Dolata, M., Leykam, J.F., Lloyd, E.A., Gonzales, M., VandenBosch, K., Kieliszewksi, M.J., 2000. Peptides isolated from cell walls of Medicago truncatula nodules and uninfected root. Phytochemistry 55, 429–438.

Gaspar, Y.M., Johnson, K.L., McKenna, J.A., Bacic, A., Schultz, C.J., 2001. The complex structures of arabinogalactan-proteins and the journey towards understanding. Plant Mol. Biol. 47, 161–176.

Gaume, A., Komarnytsky, S., Borisjuk, N., Raskin, I., 2003. Rhizosecretion of recombinant proteins from plant hairy roots. Plant Cell Rep. 21, 1193–1199.

Goodrum, L.J., Patel, A., Leykam, J.F., Kieliszewski, M.J., 2000. Gum arabic glycoprotein contains glycomodules of both extensin and arabinogalactan-glycoproteins. Phytochemistry 54, 99–106.

Hall, Q., Cannon, M.C., 2002. The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. Plant Cell 14, 1161–1172.

Held, M.A., Tan, L., Kamyab, A., Hare, M., Shpak, E., Kieliszewksi, M.J., 2004. Di-isodityrosine is the intermolecular cross-link of isodityrosine-rich extensin analogs cross-linked in vitro. J. Biol. Chem. 279, 55474–55482.

Himly, M., Jahn-Schmid, B., Dedic, A., Kelemen, P., Wopfner, N., Altmann, F., van Ree, R., Briza, P., Richter, K., Ebner, C., Ferreira, F., 2002. Art v 1 the major allergen of Mugwort pollen is a modular glycoprotein with a defensin-like and a hydryproline-rich domain. FASEB J. 17, 106–108.

Jermyn, M.A., Yeow, Y.M., 1975. A class of lectins present in the tissues of seed plants. Aust. J. Plant Physiol. 2, 501–531.

Johnson, K.L., Jones, B.J., Bacic, A., Schultz, C.J., 2003. The fasciclin-like arabinogalactan proteins of Arabidopsis. A multigene family of putative cell adhesion molecules. Plant Physiol. 133, 1911–1925.

Kieliszewski, M.J., 2001. The latest hype on Hyp-O-glycosylation codes. Phytochemistry 57, 319–323.

Kieliszewski, M.J., Kamyab, A., Leykam, J.F., Lamport, D.T.A., 1992. A histidine-rich extensin from Zea mays is an arabinogalactan protein. Plant Physiol. 99, 538–547.

Kieliszewski, M.J., Lamport, D.T.A., 1994. Extensin: repetitive motifs functional sites posttranslational codes and phylogeny. Plant J. 5, 157–172.

Kieliszewski, M.J., Leykam, J.F., Lamport, D.T.A., 1990. Structure of the threonine-rich extensin from Zea mays. Plant Physiol. 92, 316–326.

Kieliszewski, M.J., Shpak, E., 2001. Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich proteins. Cell Mol. Life Sci. 58, 1386–1398.

Kieliszewski, M.J., Xu, J., 2006. Methods of producing peptides/proteins and peptides/proteins produced thereby. US Patent US-20060026719.

Kivirikko, K.I., Liesmaa, M., 1959. A colorimetric method for determination of hydroxyproline in tissue hydrolysates. Scand. J. Clin. Lab. 11, 128–131.

Lamport, D.T.A., 1967. Hydroxyproline-O-glycosidic linkage of the plant cell wall glycoprotein extensin. Nature 216, 1322–1324.

Lamport, D.T.A., 1977. Structure biosynthesis and significance of cell wall glycoproteins. In: Loewus, F.A., Runeckles, V.C. (Eds.), Recent Advances in Phytochemistry. Plenum Publishing Corp., New York, pp. 79–115.

Lamport, D.T.A., Kieliszewksi, M.J., Showalter, A.M., 2006. Salt-stress upregulates periplasmic arabinogalactan-proteins using salt-stress to analyse AGP function. New Phytol. 169, 479–492.

Lamport, D.T.A., Miller, D.H., 1971. Hydroxyproline arabinosides in the plant kingdom. Plant Physiol. 48, 454–456.

Li, L.C., Bedinger, P.A., Volk, C., Jones, D., Cosgrove, D.J., 2003. Purification and characterization of four beta-expansins (Zea m1 isoforms) from maize pollen. Plant Physiol. 132, 2073–2085.

Liu, E., 1971. The isozymic forms of peroxidase found in the horseradish plant (*Amoracia lapathifolia*). Ph.D. Dissertation, Michigan State University.

Leonard, R., Petersen, B.O., Himly, M., Kaar, W., Wopfner, N., Kolarich, D., van Ree, R., Ebner, C., Duus, J.O., Ferreira, F., Altmann, F., 2005. Two novel types of *O*-glycans on the Mugwort pollen allergen At V 1 and their role in antibody binding. J. Biol. Chem. 280, 7932–7940.

Leucci, M.R., Di Sansebastiano, G-P., Gigante, M., Dalessandro, G., Piro, G., 2007. Secretion marker proteins and cell-wall polysaccharides move through different secretory pathways. Planta 225, 1001–1017.

Mann, K., Schafer, W., Thoenes, U., Messerschmidt, A., Mahrabian, Z., Nalbandyan, R., 1992. The amino acid sequence of a type I copper protein with an unusual serine- and hydroxyproline-rich C-terminal domain isolated from cucumber peelings. FEBS Lett. 314, 220–223.

Merkle, R.K., Poppe, I., 1994. Carbohydrate composition analysis of glycoconjugates by gas–liquid chromatography/mass spectrometry. Method Enzymol. 230, 1–15.

Motose, H., Sugiyama, M., Fukuda, H., 2004. A proteoglycan mediates inductive interaction during plant vascular development. Nature 429, 873–878.

Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bioassay with tobacco tissue culture. Physiol. Plant 15, 473–497.

Nothnagel, E.A., 1997. Proteoglycans and related components in plant cells. Int. Rev. Cytol. 174, 195–291.

Pearce, G., Ryan, C.A., 2003. Systemic signaling in tomato plants for defense against herbivores: isolation and characterization of three novel defense-signaling glycopeptide hormones coded in a single precursor gene. J. Biol. Chem. 278, 30044–30050.

Pope, D.G., 1977. Relationships between hydroxyproline-containing proteins secreted into the cell wall and medium by suspension-cultured Acer pseudoplatanus cells. Plant Physiol. 59, 894–900.

Schenk, R.U., Hildebrandt, A.C., 1972. Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. Can. J. Bot. 50, 199–204.

Schultz, C.J., Johnson, K.L., Currie, G., Bacic, A., 2000. The classical arabinogalactan protein gene family of Arabidopsis. Plant Cell 12, 1751–1767.

Schultz, C.J., Ferguson, K.L., Lahnstein, J., Bacic, A., 2004. Post-translational modifications of arabinogalactan-peptides of *Arabidopsis thaliana*: ER and GPI-anchor signal cleavage sites and hydroxylation of proline. J. Biol. Chem. 279, 1–48.

Shannon, L.M., Kay, E., Lew, J., 1966. Peroxidase isozymes from horseradish roots. I. Isolation and physical properties. J. Biol. Chem. 241, 2166–2172.

Showalter, A.M., 2001. Arabinogalactan-proteins: structure expression and function. Cell Mol. Life Sci. 58, 1399–1417.

Shpak, E., Barbar, E., Leykam, J.F., Kieliszewski, M.J., 2001. Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum*. J. Biol. Chem. 276, 11272–11278.

Shpak, E., Leykam, J., Kieliszewksi, M.J., 1999. Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-*O*-glycosylation codes. Proc. Natl. Acad. Sci. USA 96, 14736–14741.

Smallwood, M., Yates, E.A., Willats, W.G.T., Martin, H., Knox, J.P., 1996. Immunochemical comparison of membrane-associated and secreted arabinogalactan-proteins in rice and carrot. Planta 198, 452–459.

Smith, J.J., Muldoon, E.P., Willard, J.J., Lamport, D.T.A., 1986. Tomato extensin precursors P1 and P2 are highly periodic structures. Phytochemistry 25, 1021–1030.

Sun, W., Xu, J., Yang, J., Kieliszewski, M.J., Showalter, A.M., 2005. The lysine-rich arabinogalactan-protein subfamily in Arabidopsis: gene expression glycoprotein purification and biochemical characterization. Plant Cell Physiol. 46, 975–984.

Tan, L., Leykam, J.F., Kieliszewski, M.J., 2003. Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins. Plant Physiol. 132, 1362–1369.

van Driessche, G., Dennison, C., Sykes, A.G., van Beeumen, J., 1995. Heterogeneity of the covalent structure of the blue copper protein umecyanin from horseradish roots. Protein Sci. 4, 209–227.

Varki, A., 1993. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 3, 97–130.

Xu, J., Shpak, E., Gu, T., Moo-Young, M., Kieliszewski, M.J., 2005. Production of recombinant plant gum with tobacco cell culture in bioreactor and gum characterization. Biotechnol. Bioeng. 90, 578–588.

Xu, J., Tan, L., Goodrum, K.J., Kieliszewski, M.J., 2007. High-yields and extended serum half-life of human interferon α2b expressed in tobacco cells as arabinogalactan-protein fusions. Biotechnol. Bioeng. 97, 997–1008.

Yang, J., Showalter, A.M., 2007. Expression and localization of AtAGP18 a lysine-rich arabinogalactan-protein in Arabidopsis. Planta 226, 169–179.

Yariv, J., Rapport, M.M., Graf, L., 1962. The interaction of glycosides and saccharides with antibody to the corresponding phenylazo glycosides. Biochem. J. 85, 383–388.

Zhao, Z.D., Tan, L., Showalter, A.M., Lamport, D.T.A., Kieliszewski, M.J., 2002. LeAGP-1 arabinogalactan-protein purified from transgenic tobacco corroborates the Hyp contiguity hypothesis. Plant J. 31, 431–444.