Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans

Li Tan1, Allan M. Showalter2, Jack Egelund3, Arianna Hernandez-Sanchez4, Monika S. Doblin4 and Antony Bacic4*

1 Complex Carbohydrate Research Centre, The University of Georgia, Athens, GA, USA
2 Molecular and Cellular Biology Program, Department of Environmental and Plant Biology, Ohio University, Athens, OH, USA
3 Department of Plant Biology and Biotechnology, Faculty of Life Sciences, University of Copenhagen, Frederiksberg, Denmark
4 ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne, Melbourne, VIC, Australia

*Correspondence: abacic@unimelb.edu.au

Arabinogalactan-proteins (AGPs) are complex glycoconjugates that are commonly found at the cell surface and in secretions of plants. Their location and diversity of structures have made them attractive targets as modulators of plant development but definitive proof of their direct role(s) in biological processes remains elusive. Here we overview the current state of knowledge on AGPs, identify key challenges impeding progress in the field and propose approaches using modern bioinformatic, (bio)chemical, cell biological, molecular and genetic techniques that could be applied to redress these gaps in our knowledge.

Keywords: arabinogalactan-proteins, glycosylphosphatidylinositol anchor, cell surface signaling, AGP polysaccharide complexes, fasciclin-like AGP, galactosyltransferase

INTRODUCTION

Arabinogalactan-proteins (AGPs), ubiquitous cell surface proteoglycans in both terrestrial and aquatic plants (and algae), are proposed to play essential roles in a range of plant growth and development processes, including cell expansion, cell division, reproductive development, somatic embryogenesis, xylem differentiation, abiotic stress responses, and hormone signaling pathways. These roles emerge from largely indirect evidence and from the “recognition” potential arising from the incredible diversity of their gycan and protein backbone moieties as well as their location; attached to the outer leaflet of the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor and in some instances cross-linked into the wall. Despite intense research to unravel AGP function, their molecular mechanism(s) of action remain elusive. AGPs exhibit complexity at many levels: First, AGP protein backbone genes are part of large gene families, and this makes the study of AGP function through characterization of single AGP mutants a major challenge due to gene redundancy (Ma and Zhao, 2010; Showalter et al., 2010). Second, AGP protein backbones are highly glycosylated, hindering production of antibodies specifically directed to the protein moiety that would allow for identification and genetic techniques that could be applied to redress these gaps in our knowledge.

STRUCTURES

Arabinogalactan-proteins belong to the hydroxyproline-rich superfamily of glycoproteins (Schultz et al., 2002; Johnson et al., 2003) being composed largely of carbohydrate (90–98% w/w) with some protein typically rich in the amino acids, Hyp/Pro, Ala, Ser/Thr, that is usually covalently modified with a GPI anchor at the C-terminus (see Figure 1). Historically, AGPs were defined if they met three criteria: the presence of arabinogalactan chains, a Hyp-rich protein backbone, and the ability to bind to a class of synthetic phenylalanyl dyes, the β-glucosyl Yariv reagent (see Figure 1). The significant advances in our knowledge of their carbohydrate structures, protein backbone sequences, and variability in Yariv binding has considerably complicated how an AGP is defined. For instance, the diversity of protein backbones has led to a classification of the AGPs into different sub-classes based on the presence/absence of particular motifs/domains (Johnson et al., 2003). The carbohydrate moiety is typically in the form of type II arabinogalactans (AGs) although some AGPs also contain short arabinio-oligosaccharide chains (Figure 1; Fincher et al., 1983; Tan et al., 2004, 2010; Ellis et al., 2010). Type II AGs have also...
been reported either as free polysaccharides (Ponder and Richards, 1997) or as side chains of rhamnogalacturonan I (RG-I; Caffall and Mohnen, 2009). The existence of different forms of type II AGs raises a few questions. Are free type II AGs generated from AGPs by hydrolases in the wall or synthesized de novo? Are AG side chains of RG-I derived from either AGPs by transglycosylases or from covalently linked RG-I–AGP complexes? To understand how this diversity impacts biological function, we face the challenge of isolating “individual” AGPs and sequencing their glycans (and protein backbones).
Another aspect of AGP research is the intriguing possibility that they are one form of covalent cross linker for wall matrix polysaccharides. In the early 1970s, Kreeghta et al. (1973) hypothesized that Rha residues on AG side chains of AGPs might be attachment sites for RG-I. Since then, AGPs/AGs have been reported to form complexes with pectins (Yamada et al., 1987; Saulnier et al., 1988; Iwan and Morvan, 1991; Pellerin et al., 1995; Yamada, 2000; Duan et al., 2003, 2004) and xylans (Iwan and Morvan, 1995). However, residues involved in the covalent cross-link between AGPs and wall polysaccharides have not been defined. Several major challenges must be addressed to determine if AGP polysaccharide complexes (APCs) exist and to determine the structure and function of any such complexes.

CHALLENGE 1: ISOLATION AND PURIFICATION OF AGPs
The incredible heterogeneity of AGP structures has hampered purification of individual AGPs. As a consequence, most studies on AGPs have been on a family of molecules and often in the presence of contaminating polymers. There are a few examples of AGPs purified by a combination of traditional chromatographic methods (for example, anion exchange/lectin affinity/gel permeation using chaotropic reagents) and/or Yariv precipitation that are “pure” AGPs; for example, the AGPs from tobacco floral tissues (Gane et al., 1995) and larch AG exudates (Ponder and Richards, 1987). The application of molecular biology techniques to both isolate heterologously expressed AGP protein backbones or synthetic peptides as green fluorescent protein tagged (GFP)-fusion proteins by the Kieliszewski/Showalter, Matsuoka, and Somerville laboratories (Shpak et al., 1999; Zhao et al., 2002; Tan et al., 2004; Shimizu et al., 2005; Estève et al., 2006) was an ingenuous innovation that allowed the purification of AGPs with a single protein backbone and therefore the study of inherent glycan heterogeneity. However, the low DP of these glycans raises some questions on the fidelity of glycosylation in heterologous/high expression systems.

Thus, a combination of purification techniques is necessary to purify relatively homogeneous AGPs (and APC complexes extracted as described below). These techniques take advantage of the heterologous structural features of AGPs and wall polysaccharides including size, charge, hydrophobicity (Serpe and Nothnagel, 1996; Lampour et al., 2011), the ability to co-precipitate with Yariv reagents, the availability of anti-AG antibodies (Pattathil et al., 2010), and the use of tagged heterologously expressed protein backbones.

CHALLENGE 2: EXTRACTION AND PURIFICATION OF PUTATIVE APCS FROM WALLS
Because pectins and non-cellulosic polysaccharides are embedded within the highly cross-linked wall, the first obstacle to studying putative APCs is to extract intact macromolecules from the wall, especially from secondary walls. Traditional methods to release polysaccharides from the wall include either the use of wall-specific degrading enzymes (York et al., 1986) or the extraction of walls with increasingly harsh solvents (Fry, 1988). Since the enzymatic and strong base treatments could also potentially break covalent linkages between AGPs and wall polysaccharides, the released polymers may only contain partial structural information of potential APCs and may still contain contaminating wall polysaccharides.

To avoid these extraction complications it may be possible to source APCs from potentially rich sources such as suspension culture media, especially of xylogenic calli, polysaccharide-rich seed mucilages, and exudates, such as gums (Defeye and Wong, 1986) and root mucilages (Boddy et al., 1988) since these are released in a “solsibilized” form.

CHALLENGE 3: SEQUENCING OF AGs AND APCs
Our current knowledge of AG carbohydrate sequences is based on experiments using tools that include monosaccharide composition, linkage analysis, chemical or enzymatic degradation of glycans, mass spectrometry (MS), and NMR analysis. Partial acid hydrolysis (Defeye and Wong, 1986), acetylation, alkaline degradation, and Smith degradation (Churms et al., 1981; Bacic et al., 1987) have supported the basic structures summarized in Figure 1 and led to the suggestion that the AG glycans contain a backbone of β-(1,3)-galacto-oligosaccharides interrupted at regular intervals with a periodate-sensitive residue. However, few of the large AG chains have been de novo-sequenced due to the inherent biosynthetic heterogeneity and the current limitations of sequencing technologies.

The availability of linkage-specific enzymes has greatly assisted the sequencing of glycans although their lack of commercial accessibility has hampered progress. Thus a breakthrough in AGP analysis was the identification of an AGP-specific exo-β-(1,3)-galactanase that can bypass the β-(1,6)-galactosyl side chains (Tsumuraya et al., 1990; Kotake et al., 2005; Ichinoe et al., 2006). This enzyme, together with the recently characterized β-glucoaminidase (Haque et al., 2005), α-arabinofuranosidase (Hata et al., 1992), and endo-β-(1,6)-galactanase form a enzyme tool kit specific for AG side-chain analysis which enabled Tryfona et al. (2010) to characterize some long β-(1,6)-galacto-oligosaccharide AG side chains with the aid of MS/MS fragmentation (see Oxley et al., 2004). A recent study of Arabidopsis AGP31 (Hijazi et al., 2012), a chimeric AGP, illustrates the power of a multipronged approach to purification and characterization of AGPs.

Therefore, the best solution is to sequence small structural units of AGs, generated using a combination of chemical and enzymic techniques, and then to re-construct models of the intact AGs. Discovery of new chemicals and enzymes that can selectively cleave AGs would facilitate future progress in the sequencing of the AG glycan chains.

BIOINFORMATICS
Genomics and its related technologies have revolutionized the study of biology, facilitated the development of other ‘omics platforms, and created a need for bioinformatics to handle the acquisition, storage, and analysis of the vast amount of data generated from ‘omics and ‘omics-related projects. The AGP field has greatly benefited from genomics and bioinformatics. Given that AGP protein backbone sequences often have low sequence similarity, BLAST-type searches typically identify only a few closely related AGP family members and, therefore, are not a particularly effective means to comprehensively identify members of the AGP family. In contrast, bioinformatics approaches have provided a
Tan et al. The research challenges of arabinogalactan-proteins

The AGP protein backbones from the rice genome/proteome, including 13 classical AGPs, 15 AG-peptides, 21 fasciclin-like AGPs (FLAs), 17 plastocyanin AGPs, and 6 other chimeric AGPs. Ma and Zhao (2010) have conducted the only other comprehensive bioinformatics analysis for AGP protein backbones in rice. They identified 69 rice AGP protein backbones from the rice genome/proteome, including 13 classical AGPs, 15 AG-peptides, 3 non-classical AGPs, 3 early nodulin-like AGPs (eNod-like AGPs), 8 non-specific lipid transfer protein-like AGPs (nLTP-like AGPs), and 27 FLAs. A few other bioinformatic studies are reported for AGP protein backbones, but these studies were not focused exclusively on AGPs and/or concentrated only on one particular sub-class (e.g., GPI-anchored AGPs or FLAs). For example, Böerner et al. (2002, 2003) used bioinformatics to identify GPI-anchored proteins in A. thaliana from genomic and proteomic data. In addition, Iridah et al. (2008) applied bioinformatic analysis to their cell wall proteomic data in A. thaliana to identify several AGPs and Faulk et al. (2006) used bioinformatic analyses to identify 34 wheat and 24 rice FLAs. Bioinformatic tools have also been used to provide insight to the glycosyltransferases (GTs) involved in the assembly of AGP glycans (see Biosynthesis of Glycan Moieties) and in this way Bielesic and colleagues have proposed that the CAZy GT-family-31 consists of putative β-(1,3)-GalTs (Qiu et al., 2006; Ellis et al., 2010; Egelund et al., 2011).

The comprehensive bioinformatic studies on AGPs also took advantage of other related genomic technologies, including microarray data to reveal organ-specific expression patterns, abiotic- and biotic-regulated expression profiles, and genes which are co-expressed. Co-expression analysis has the potential to reveal networks of genes that are related to particular aspects of AGP biology, including their biosynthesis, interacting partners, and physiological functions. These kinds of downstream bioinformatic analyses are just in their infancy and many bioinformatics challenge lie ahead relating to AGPs, as outlined below.

CHALLENGE 4: IDENTIFYING AND CLASSIFYING AGPs FROM OTHER SEQUENCED PLANT GENOMES

Over 30 plant and algal genomes/proteomes are now known (see http://en.wikipedia.org/wiki/List_of_sequenced_eukaryotic_genomes#Algae). It would be useful to apply either the current or improved bioinformatics programs to these various datasets. Suggested enhancements to the programs would include making the bioinformatic analysis more automated and integrating the programs for predicting signal peptides, GPI anchor addition sites, gene expression, co-expression analysis, etc. into a single program. In addition, based on existing protein sequence and carbohydrate data on AGPs, a bioinformatics program predicting sites of prolyl hydroxylation and corresponding sites and type of glycosylation (i.e., AGs and arabinio-oligosaccharides) could be developed and used. This relies on our knowledge that the types of O-glycosylation on the AGP protein backbone can be predicted from the Hyp-contiguity hypothesis that defines Hyp (arabinio)galactosylation as occurring on the clustered, non-contiguous Hyp residues separated by Ala or Ser residues in a protein backbone whereas blocks of contiguous Hyp residues, such as occur in extensins, are arabinosylated with short oligosaccharides (Kieliszewski and Lamport, 1994; Shpak et al., 1999; Goodrum et al., 2000; Zhao et al., 2002). N-glycosylation, is predicted by the universally conserved consensus amino acid sequence Asn-X-Ser/Thr, where X can be any amino acid except Pro. Similarly, the specificity of prolyl hydroxylation by prolyl-4-hydroxylase, although not as well defined in plants as in mammalian systems (Gorres and Raines, 2010), can be used together with the Hyp-contiguity hypothesis to inform design of bioinformatics programs.

CHALLENGE 5: APPLING AND IMPROVING BIOINFORMATICS ANALYSES OF MICROARRAY DATA TO ELUCIDATE PATTERNS OF AGP (CO-)EXPRESSN

Unfortunately, not all of the sequenced plant genomes have extensive publicly available microarray data, unlike Arabidopsis and rice (e.g., see PLEXdb, http://www.plexdb.org/). Thus, in addition to generating new microarray data, it would be convenient to utilize and integrate expression analysis programs like Genevestigator and co-expression analyzer tools (see Table 1) to mine data and provide it in a more tailored manner. Analysis of such data can provide remarkable insight into the function (and functional redundancy) of AGP protein backbone genes as well as elucidate networks of AGPs and AGP-related genes involved in various metabolic pathways.

CHALLENGE 6: IMPROVING AND DEVELOPING NEW BIOINFORMATICS PROGRAMS TO ELUCIDATE MOLECULAR PHYLOGENIES OF AGP PROTEIN BACKBONE GENES

It would be interesting from an evolutionary standpoint to understand how AGPs are related within and between species, since such analysis may explain how the AGP gene family evolved and provide insight into AGP function. From a functional perspective, it would be useful to be able to identify AGP gene orthologs and paralogs. Software developers would use the gene families identified in Challenge 4 through packages summarized in Table 1 and the extensive web-based resources developed for studying gene orthology to focus on the AGP protein backbone genes.

CHALLENGE 7: DEVELOP BIOINFORMATICS TOOLS TO IDENTIFY AND CLASSIFY GENES/PROTEINS INVOLVED WITH AGP METABOLISM

Bioinformatic tools to identify genes involved with the biosynthesis and possible modification and degradation of AGPs would be of great benefit. In particular, bioinformatics analysis has the potential to identify GTs likely to be involved in the biosynthesis of AG chains. Currently, sequence similarities to mammalian GTs represent one approach to identifying these enzymes, for example, as recently described by Egelund et al. (2011) in which the authors adopted a bioinformatic approach to identify and systematically characterize putative GaTs from CAZy GT-family-31 responsible for synthesizing the β-(1,3)-Gal linkage. This study revealed that the Arabidopsis accessions grouped into four plant-specific clades (1, 7, 10, and 11; Table 2). Furthermore, the investigators attempted to predict the possible substrate specificity of these
Table 1 | Bioinformatic programs used to identify and characterize AGPs.

| Program                        | Program use                                                                 | Web address                                                                 |
|-------------------------------|-----------------------------------------------------------------------------|------------------------------------------------------------------------------|
| PAST percentage calculator    | Identification of AGP backbones                                             | http://www.adelaide.edu.au/directory/carolyn.schultz/under_files             |
| BIO OHO                       | Identification of AGP backbones and more                                     | http://tose.google.com/bio/protein-class/                                    |
| SignalP                       | Identification of signal peptides                                           | http://www.cbs.dtu.dk/services/SignalP/                                     |
| Plant big-PI predictor        | Identification of GPI anchor addition sites                                 | http://mendel.imp.ac.at/gp/plant_server.html                                  |
| Genevestigator                | Identification of gene expression                                           | https://www.genevestigator.ethz.ch/                                         |
| Arabidopsis Co-Response Database | Identification of co-expressed genes                                       | http://csbdb.mpimp-golm.mpg.de/csbdb/dbcor/ath.html                          |

CHALLENGE 1: DEVELOP BIOINFORMATICS TO IDENTIFY REGULATORY SEQUENCES IN AGP PROTEIN BACKBONE GENES

Bioinformatics has the potential to reveal gene regulatory sequences involved in regulated expression of AGP genes with respect to developmental expression (e.g., tissue- and temporal-specific expression) and a variety of stresses. Bioinformatic programs that have the ability to recognize either conserved nucleotide patterns alone or in combination with chromatin immunoprecipitation (ChIP) assays followed by DNA sequencing have the potential to reveal AGP gene regulatory sequences and the corresponding trans-acting factors. Knowledge of such regulatory sequences would reveal commonly regulated networks of AGP genes as well as other co-regulated genes. As such, this information may be complementary to co-expression data and would provide another avenue to elucidating AGP function(s).

BIOSYNTHESIS OF GLYCAN MOIEITIES

Many mammalian, fungal, and bacterial GTs have been identified, cloned, and biochemically characterized (Cantarel et al., 2009; Ellis et al., 2010). In contrast, only a few plant cell wall polysaccharide/proteoglycan-related GTs have been characterized biochemically (Doblin et al., 2010). From studies of Arabidopsis at the molecular and biochemical level (Strasser et al., 2007; Qu et al., 2008), and from assembly of mammalian proteoglycans, it is expected that AG glycan chains that decorate AGPs are synthesized by type II membrane-bound GTs located in the Golgi apparatus. This includes members of CAZY GT-family-31 with putative β(1,3)-GaT activity, that are suggested to be involved in synthesis of the β(1,3)-Gal backbone in AG glycan chains (Qu et al., 2008; Egeland et al., 2011).

Early studies showed that the Golgi apparatus plays an important role in synthesis of β(1,6)-Gal of the AG glycan chains of AGPs (Mascara and Fincher, 1982; Schibeci et al., 1984), whereas the initial enzyme in the AG biosynthetic pathway, adding the first Gal residue to a Hyp residue on the protein backbone (the Hyp-O-galactosyltransferase or HGT), is predominantly located in the ER (Oka et al., 2010). Outside of the development of in vitro assays to monitor GaT activity (Qu et al., 2008; Liang et al., 2010;
The research challenges of arabinogalactan-proteins (AGPs) have made significant advances in the past decades. However, the understanding of AGPs is still limited, and several key challenges remain.

**Challenge 1:** BIOSYNTHESIS OF PUTATIVE APCs

The challenge is to determine in which sub-cellular compartment putative APCs are assembled and by what mechanism. One possibility is that APCs are synthesized intracellularly in the ER/Golgi apparatus by multiple GTs (as proposed for AGPs and other non-cellulosic polysaccharides) by either en bloc transfer of pre-assembled oligosaccharides or stepwise sugar addition, followed by delivery into the wall. Another possibility is that APCs are assembled in the extracellular matrix, possibly by transglycosylases, a mechanism that has been well studied in xyloglucan remodeling (Seifert and Roberts, 2007).

**Challenge 2:** AN ALTERNATIVE APPROACH FOR THE IDENTIFICATION OF THE GLYCOSYLATION MACHINERY INVOLVED IN AG CHAIN SYNTHESIS

An alternative approach to the one described in Challenge 7, centers on the analysis of Gum Arabic, a tree exudate from the Acacia species, whose main fraction is an AG (DeFaye and Wong, 1986; Randall et al., 1989; Al-Assaf et al., 2005). AG chains comprise as much as 90–98% of the gum exudate (Osman et al., 1993), thus making Gum Arabic-producing cells from the Acacia trees an obvious choice as starting material to identify enzymes involved in AG biosynthesis.

**Challenge 3:** IDENTIFICATION OF THE GLYCOSYLATION MACHINERY INVOLVED IN AG CHAIN SYNTHESIS

The most obvious choice would be to develop an "in planta" system, possibly by participation in cellulose deposition within the wall (Rose et al., 2002) and is commonly utilized by yeast to modify their wall in response to abiotic/biotic stimuli (Kollar et al., 1997).

**Challenge 4:** HETEROLOGOUS EXPRESSION SYSTEMS

Expression of non-cellulosic/cellulosic plant GTs in functional assay systems remains a key challenge. The past lack of success of this approach has been ascribed to the mismatch between biochemical assays and native activity, failure of the expressed protein to accumulate to sufficient levels, incorrect folding or improper post-translational modifications (Petersen et al., 2009). The most obvious choice would be to develop an "in planta" system, however, the endogenous GT activities can make it difficult to distinguish the specific activity of the expressed protein (Petersen et al., 2009). Prokaryotes, of which some have limited capacity for post-translational processing, pose other problems. We therefore suggest developing multiple heterologous expression systems to maximize the likelihood that at least one will allow for successful expression where the biochemical activity is retained. Additionally, testing new expression systems that may prove "universal" (e.g., Aspergillus), which has served as one of the preferred expression systems in the biotechnology industry, as well as cell-free expression systems may prove useful for heterologous expression of plant GTs.

**Challenge 5:** A HIGH-THROUGHPUT ENZYME ACTIVITY SCREENING SYSTEM

The assignment of substrate specificity to GTs is often hindered by difficulties related to limited availability of relevant candidate acceptor molecules for biochemical assays. To overcome this challenge, the next step should be to employ carbohydrate array technology (Moller et al., 2007) with AGP/Gum Arabic-specific sugars and peptides, related acceptor substrates, i.e., natural acceptors from Gum Arabic and AGPs (e.g., β-(1,3)-galacto-oligosaccharides, generated by Smith degradation (see Challenge 3)), de-arabinosylated AGPs generated by mild acid, chemically synthesized β-(1,3)-galacto-oligosaccharides and isolated AGP protein backbones) together with other "AGP-enriched" fractions from wild type, AGP GT mutants, and Gum Arabic exudates.

Combining AGP-related arrays with established in vitro assays will facilitate a high-throughput screening system that can be used to test heterologously expressed candidate GTs in mixtures with either radio-labeled or fluorescently tagged NDP-sugar as the donor to identify AGP-specific carbohydrate acceptor molecules on the array. Development of such a comprehensive screening system would be a significant step in identifying the many GTs responsible for AG biosynthesis.

**Function**

Arabinogalactan-protein glycan-specific antibodies and β-Gal Variv reagent have recently been published (Li et al., 2010a,b; Ellis et al., 2010). The current use of these two indirect tools continues to provide information on AGP activity in new biological systems, e.g., European larch, Larix decidua (Rafinska and Bednarska, 2011), and little studied developmental processes, in this case, ovule development in gymnosperms, confirming the relevance and the conservation of function of these molecules within the plant kingdom. Unfortunately, the broad specificity of these techniques makes it impossible to assign function to a single AGP. This limitation has been partially overcome by genetic and molecular studies, including the characterization of AGP single or double mutants, RNAi and over-expressing lines, although these approaches also have complications.

The usefulness of reverse genetics approaches to investigate AGP backbone function is well demonstrated. The function of one cotton FLA, FLA5, in cotton fiber initiation and elongation (Li et al., 2010a) and four Arabidopsis members, FLA1, FLA3, FLA11, and FLA12 have recently been published (Li et al., 2010a,b; MacMillan et al., 2010; Johnson et al., 2011). Roles for FLA1 in lateral root and shoot development in tissue culture plant cell lines have been described (Johnson et al., 2011) and FLA3 in mesosome development, possibly by participation in cellulose deposition within...
The characterization of the promoters of AGP genes specifically expressed in pollen, have been demonstrated to be a guide to study gene function, we should be aware that in some cases mRNA levels have not been in agreement with protein levels (Yang et al., 2011).

CHALLENGE 14: PRODUCTION OF SPECIFIC AGP PROTEIN BACKBONE ANTIBODIES

The recent production of antibodies specifically recognizing the Lys-rich region of AtAGP17 and AtAGP19 protein backbones demonstrates both the veracity of this approach and also provides tools to study in more detail their tissue and cellular distribution and ultimately their function (Yang et al., 2011). Either these antibodies or alternatively antibodies to tagged versions of AGPs could be used in co-location and immunoprecipitation experiments to identify possible interacting partners.

CHALLENGE 15: DETERMINING THE FUNCTIONAL SIGNIFICANCE OF AG GLYCAN CHAIN HETEROGENEITY

One approach to address the functional importance of the glycan moiety of AGPs is to characterize AGP-specific GT mutants. Mutants implicated in AGP glycan moiety biosynthesis by transcript co-expression analysis could also be analyzed as single mutants and in combination with other GTs to potentially increase plant phenotypic severity. Limiting analyses to either single cell types (e.g., pollen/pollen tubes), or simple tissues with limited cell-types, would help in these analyses and provide a more restricted list of candidate GT genes. These genes could then be heterologously co-expressed and cellular fractions used in biochemical assays for functional assessment. While the initial aim of this work is to identify the GTs involved in AGP glycan synthesis, the underlying objective is to use these and other AGP mutants as functional assay systems to dissect the mechanism and pathway of AGP synthesis in greater detail.

Several questions arise from this ligand model of AGP function. Given the effects on cellulose in the fla11 fla12 double null mutant (MacMillan et al., 2010) and the abnormal cellulose deposition fla13 RNAs (Li et al., 2010b), may FLA11, FLA12, and FLA3, as well as other GPI-anchored or non-anchored AGPs, also be a part of this same network of components involved in wall sensing? Does this model explain the observation of AGPs as cell fate markers in tissues undergoing cell differentiation? Considering that the appearance of AGPs during specific developmental stages has been described using antibodies that recognize AGP-carbohydrate epitopes, is the heterogeneity of AGP glycosylation also involved in providing the necessary specificity to interact with different signal molecules and generate specific responses? What is the relevance of the presence and number of fascin domains of FLAs? Further investigation of the possible function of AGPs in wall sensing is of fundamental importance to uncover some of the components and mechanisms involved in the regulation of wall biosynthesis and ultimately plant cell growth. To address some of these challenges, we propose the use of the following experimental approaches, techniques, and resources:

- Production of specific AGP protein backbone antibodies
- Targeting functional redundancy of AGPs
interconnections of the macromolecular components. Plant Physiol. 91, 188–196.
Kieliszewski, M. J., and Lampert, D. T. A. (1994). Extensive repetitive motifs, functional sites, post-translational codes, and phylogeny. Plant J. 5, 157–172.
Kolke, B., Reinhold, B. B., Petrieva, E., Yeh, H. F. C., Isabell, C., Dignard, J., Kaptijn, J. C., Kh, F. M., and Caher, E. (1997). Architecture of the yeast cell wall – β(1 → 3)-glucan interacts monomannoprotein, β(1 → 3)-glucan, and chitin. J. Biol. Chem. 272, 17762–17775.
Kondo, T., Diao, S., Konishi, T., Kaneko, S., Igarashi, K., Samejima, M., Watanabe, Y., Kimura, K., and Tsuneya, Y. (2003). Molecular cloning of a glucanase from radiol which specifically hydrolyses β(1 → 3)- and β(1 → 6)-galactosyl residues of arabinogalactan protein. Plant Physiol. 135, 1563–1576.
Kwon, J. S., and Morvan, H. (1981). Extracellular branched xylos an acidic arabinogalactan from suspension cultured cells of white campion (Silene alba Mill.) E.H.L. Krause. Food Hydrocoll. 5, 155–166.
Kwon, J. S., and Seo, H. (1995). Characterization of extracellular β(1 → 6)-rhamnose backbone O-substituted by arabinogalactan type-II in a plant cell suspension. Carbohidr. Res. 272, 305–318.
Lampert, D. T., Tan, L., and Kieliszewski, M. J. (2011). Structural proteins of the primary cell-wall extraction, purification, and analysis. Methods Mol. Biology, ed. Z. Popper (New York: Springer-Nature New York Inc.), 209–218.
Li, Y., Liu, D., Li, T., Zhang, X., Wang, J., Song, T., Tan, L., and Lampert, D. T. F. (2010a). Suppression of GhAGP4 gene expression represses the initiation and elongation of cotton fiber. Plant Cell Rep. 29, 193–202.
Li, J., Yu, S., Meng, L.-L., and Zhao, J. (2010b). The fasciclin-like arabinogalactan protein gene, FLP, is involved in microtubule development of Arabidopsis. Plant J. 64, 482–497.
Liang, Y., Faik, A., Kieliszewski, M., Tan, L., Xu, W. L., and Shew, W. T. A. (2010). Identification and characterization of α-mannosyltransferase activities involved in arabinogalactan protein glycosylation. Plant Physiol. 154, 632–642.
Mia, H., and Zhao, J. (2010). Genome-wide identification, classification, and expression analysis of the plant cell-wall glycan-directed monoclonal antibodies. Plant Physiol. 153, 518–529.
Pellitteri, P., Vitali, S., Williams, P., and Brillouet, J. M. (1997). Characterization of five type II arabinogalactan-protein fractions from red wine of increasing arctic origin. Carb. Res. 277, 125–140.
Pierson, B., Engfeld, J., Damarter, J., Faber, K., Kruger Jensen, J., Yang, W., Bennett, E., Schlueter, H., andUknes, P. B. (2009). Asymmetric glycosylation expression in Pichia pastoris of plant cell wall type-II membrane anchored glycoconjugates. Glycobiology. 20, 1235–1246.
Premiér, G. R., and Ruhland, G. N. (1997). Arabinogalactan from Western hemlock. Part III alkaline degradation revisited, with novel conclusions on molecular structure. Carbohydr. Poly. 34, 253–261.
Qi, W., Fong, C., and Lampert, D. T. A. (1991). Gain of arabinogalactan is a translational hairy vac y: a novel model based on O-galactosylhydroxyproline as the polysaccharide attachment site. Plant Physiol. 96, 840–855.
Qu, Y., England, J., Galon, P. R., Houghton, E., Gleason, P. A., Schlueter, C. L., and Basic, A. (2008). Identification of a novel group of putative Arabidopsis thaliana f(1,3)-galactosyltransferases. Plant Mol. Biol. 66, 45–59.
Rafinska, K., and Bednarska, E. (2011). Localization pattern of homogalacturonan and arabinogalactan proteins in developing ovules of the gymnosperm plant Larix decidua Mill. Acta. Bot. Pol. 74, 75–87.
Randall, R. C., Phillips, G. O., and Williams, P. A. (1998). Identification and characterization of glycosylation of Araucaria angustifolia leaves. Food Hydrocoll. 12, 75–79.
Rose, J. K., Cram, J., Fry, S. C., and Nishizaki, K. (2005). The XTH family of enzymes involved in xylogenous endodermalization and endodermal synthesis: current perspectives and a new unifying nomenclature. Plant Cell Physiol. 46, 1421–1435.
Sundar, L., Brillouet, J.-M., and Jouve, J.-P. (1988). Structural studies of pectic substances from the pulp of grape berries. Carbohydr. Res. 182, 63–78.
Sukh, I., Ayak, J., and Fancher, G. B. (1994). Biosynthesis of arabinogalactan-protein in Lentinus edodes (Japanese reygy) endomycorrhizal cells. Rockym. J. 128, 653–656.
Sukh, I. J., Ruzomberok, M. J., John- son, J. K., Jones, B. J., Gaspar, Y. M., and Basic, A. (2002). Using genomic resources to guide research directions. The arabinogalactan protein gene family as a test case. Proc. Natl. Acad. Sci. U.S.A. 109, 1448–1453.
Seifert, G. J., and Roberts, K. (2007). The biology of arabinogalactan proteins. Annu Rev. Plant Biol. 58, 137–165.
Serce, M. D., and Nothnagel, E. A. (1998). Heterogeneity of arabinogalactan-proteins on the plasma membrane of rose cells. Plant Physiol. 112, 1261–1271.
Shi, H., Kim, Y., Gao, Y., Stevenson, B., and Zhu, J.-K. (2008). The Arabidop- sis AGP10 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. Plant Cell 19, 15–22.
Shinmin, M., Isgadk, T., Yamada, M., Yama, H., Hasegawa, J., Kato, T., Tsukagoshi, H., Nakamura, K., Fukuda, H., and Matsouka, K. (2005). Experimental determination of proline hydroxylation and hydroxyproline-arabinogalactan modification in soybean proteins. Plant J. 42, 877–888.
Showalter, A. M., Kapriel, B., Lichtenberg, J., Gu, D., and Wolk, L. B. (2010). A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. Plant Physiol. 153, 485–513.
Shukle, W., Leykam, J. F., and Kieliszewski, M. J. (1999). Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-O-glycosylation rules. Proc. Natl. Acad. Sci. U.S.A. 96, 14768–14774.
Strause, F., Bondoli, J. S., Viera, U., Schuster, J., Stover, K., Glass, J., Leonard, R., Stadtmann, J., Al- mann, F., Stendelhoff, H., and Mach, L. (2007). A unique f(1,3)- galactosyltransferase is indispensable for the biosynthesis of N-D-galactosylaminotransferase containing Lewis a structure in Arabidopsis thaliana. Plant Cell 19, 2278–2292.
Tan, L., Qu, F., Lampert, D. T. A., and Kieliszewski, M. J. (2004). Structure of a hydroxyproline (Hyp)- arabinogalactan polysaccharide from rice: Ala-Hyp expressed in transgenic Nicotiana tabacum. J. Biol. Chem. 279, 13136–13158.
Tan, L., Yamashita, T., Tsuneya, T., Xu, X., Qu, F., and Kieliszewski, M. J. (2010). Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short branched side chains. J. Biol. Chem. 285, 24577–24585.
Terry, D., Liang, H.-C., Kanata, T., Kaneko, S., Marec, J., Ichiho, H., Lovegrove, A., Tsuneya, Y., Schwo...
Tan et al. The research challenges of arabinogalactan-proteins

P. R., Stephens, E., and Dupree, P. (2010). Carbohydrate structural analysis of wheat flour arabinogalactan protein. Carbohydr. Res. 345, 2648–2656.

Tsumura, Y., Kochi, N., Hashimoto, Y., and Kori, Y. (1990). Purification of an exo-b-(1-3)-D-galactanase of Aspergillus oryzae (P. niger). Biochem. Biophys. Res. Commun. 170, 1110–1114.

Wu, Y., Williams, M., Bernard, S., Drysdale, A., Shaw, A., and Falah, A. (2010). Functional identification of two nonredundant Arabidopsis AtAGP17 and AtAGP19, two lysine-rich arabinogalactan proteins, in Arabidopsis. Plant Cell 22, 3665–3679.

Yamada, H. (2000). "Bioactive arabinogalactan-proteins and related products as antiangiogenic factors in Sino-Japanese herbal medicines," in Cell and Developmental Biology of Arabinogalactan-proteins, eds E. A. Nothnagel, A. Bacic, and A. E. Clarke (New York, NY: Kluwer Academic/Plenum), 211–229.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 March 2012; accepted: 10 June 2012; published online: 27 June 2012.

Citation: Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Dobler MS and Basic A (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front. Plant Sci. 3:140. doi: 10.3389/fpls.2012.00140

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

Copyright © 2012 Tan, Showalter, Egelund, Hernandez-Sanchez, Dobler and Basic. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

"fpls-03-00140" — 2012/6/25 — 20:19 — page 10 — #10
Author/s: 
Tan, L; Showalter, AM; Egelund, J; Hernandez-Sanchez, A; Doblin, MS; Bacic, A

Title: 
Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans

Date: 
2012-01-01

Citation: 
Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S. & Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. FRONTIERS IN PLANT SCIENCE, 3 (JUN), https://doi.org/10.3389/fpls.2012.00140.

Persistent Link: 
http://hdl.handle.net/11343/265112

File Description: 
Published version

License: 
CC BY-NC