IDENTIFICATION AND CHARACTERIZATION OF THE ARABIDOPSIS GAUT1:GAUT7 PECTIN HOMOGALACTURONAN:GALACTURONOSYLTRANSFERASE COMPLEX

by

MELANI AGUSTINA ATMODJO

(Under the Direction of Debra Mohnen)

ABSTRACT

Pectin is a family of structurally complex polysaccharides defined by the presence of α-1,4-linked galacturonic acid (GalA) residues. It is present in all plant cell walls and contributes substantially to normal plant growth and development. Prior research has elucidated the intricate structure of pectin, yet a detailed understanding of pectin biosynthesis remains elusive. The membrane-bound Arabidopsis GALACTURONOSYLTRANSFERASE 1 (GAUT1) catalyzes the biosynthesis of homogalacturonan (HG), a linear homopolymer of α-1,4-linked GalA that makes up ~65% of total pectin. GAUT1 is an α-1,4-galacturonosyltransferase (GalAT) that transfers GalA from UDP-GalA onto HG acceptors to synthesize HG oligomers and polymers. This study presents the identification of an Arabidopsis HG:GalAT complex composed of GAUT1 and its homolog GAUT7, both of which are Golgi-localized, type II transmembrane proteins. Co-immunoprecipitation using anti-GAUT1 and anti-GAUT7 polyclonal antibodies and bimolecular fluorescent complementation demonstrated protein-protein interaction between GAUT1 and GAUT7. Transcript analyses indicate GAUT1 and GAUT7 co-expression in most tissues and developmental stages of Arabidopsis. Upon non-reducing SDS-PAGE, both GAUT1 and GAUT7 resolve at ~185 kDa, a size larger than their monomeric forms (~60 kDa and ~75

kDa, respectively) observed by reducing SDS-PAGE, thus indicating disulfide bond involvement in GAUT1:GAUT7 complex formation. Proteomics analyses identified GAUT1 and GAUT7 as the only components of the non-reducing SDS-PAGE-stable GAUT1:GAUT7 complex. The observed size of monomeric Arabidopsis GAUT1 (~60 kDa) is smaller than its predicted mass (77.4 kDa). N-terminal sequencing of GAUT1, mass spectrometry protein sequence coverage, and primary structure-dependent GAUT1 reactivity against a series of anti-GAUT1 antibodies, all suggest post-translational cleavage of Arabidopsis GAUT1 between Met_{167} and Arg_{168}, resulting in a mature GAUT1 with no transmembrane domain. Transient expression of several GFP fusion constructs of GAUT1 in *N. benthamiana* leaves, in the presence or absence of GAUT7, showed a GAUT7-dependent Golgi accumulation of GAUT1-GFP. It is concluded that Golgi retention of GAUT1 is due to the formation of a complex with GAUT7. Models of the GAUT1:GAUT7 HG:GalAT complex, with GAUT1 as the catalytic subunit and GAUT7 as the membrane anchor, are proposed. Twelve additional proteins were also identified by proteomic analyses and proposed as GAUT1:GAUT7 complex putative interacting proteins.

INDEX WORDS: plant cell wall, pectin, homogalacturonan, biosynthesis, glycosyltransferase, galacturonosyltransferase, Golgi apparatus, *Arabidopsis thaliana*, GAUT, protein complex, membrane anchor

IDENTIFICATION AND CHARACTERIZATION OF THE ARABIDOPSIS GAUT1:GAUT7 PECTIN HOMOGALACTURONAN:GALACTURONOSYLTRANSFERASE COMPLEX

by

MELANI AGUSTINA ATMODJO

B.S., Bogor Agricultural University, Indonesia, 1997

M.S., University of Adelaide, Australia, 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

IDENTIFICATION AND CHARACTERIZATION OF THE ARABIDOPSIS GAUT1:GAUT7 PECTIN HOMOGALACTURONAN:GALACTURONOSYLTRANSFERASE COMPLEX

by

MELANI AGUSTINA ATMODJO

Major Professor:	Debra Mohnen
Committee:	Maor Bar-Peled Michael Pierce Lance Wells

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2010

DEDICATION

This doctoral work is dedicated to my husband Frank V. Gagliano, my family back home in Bandung, Indonesia, and my new family here in Athens and Crawford, Georgia, U.S. Their continuous support, love, and prayers are what kept me going and made all this possible.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Debra Mohnen, for giving me the opportunity to learn and work on this project. I appreciate her guidance, friendship, and encouragement throughout my time in her laboratory, as well as the knowledge, experience, and inspiration that she shared me. I would also like to thank my committee members Dr. Maor Bar-Peled, Dr. Michael Pierce, and Dr. Lance Wells, for their support, advice, and constructive criticism that helped me get to where I am now.

It was a great honor for me to have the chance to collaborate with these excellent researchers: Dr. Yumiko Sakuragi (University of Copenhagen, Denmark), Dr. Henrik Scheller (Joint BioEnergy Institute), Xiang Zhu (CCRC), Dr James Atwood III (Bioinquire), and Dr. Ron Orlando (CCRC). I am indebted as well to many people who generously provided me with assistance and materials to do my research: Drs. Carl Bergmann and Gerardo Gutierrez-Sanchez (CCRC–for exo- and endopolygalacturonase enzymes), Stefan Eberhard (CCRC–for Arabidopsis suspension culture, RG-I and RG-II polysaccharides), Sarah Inwood (CCRC–for technical assistance and enzyme), Drs. Jae-Min Lim and Lance Wells (CCRC–for LC-MS/MS identification of GAUT1 and GAUT7), Dr. Mary Tierney (The University of Vermont-for pBI101 vector containing *AtPRP3* promoter), Dr. Malcolm O'Neill (CCRC–for RG-II monomer and dimer), Dr. Henk Schols (Wageningen University–for RG-I hydrolase enzyme), Dr. Simon Turner (University of Manchester, UK-for KOR1 antibody), and Sangita Mohanty (CCRC–for technical assistance).

My gratitude goes to all past and present members of the Mohnen lab, whom I have been privileged to know, work with, and learn from over the years, and to all the undergraduate students who contributed to my research, especially Amy Burrell and Ioana Petrascu, for their excellent work. I would also like to acknowledge people at CCRC for their administrative (especially Sheilah Dixon-Huckabee, Karen Howard, and Carol Connelly) and technical assistance, scientific advice, and friendships.

TABLE OF CONTENTS

Page

INTRODUCTION

The carbohydrate-rich cell wall is an extracellular matrix that encircles and connects all cells within a plant. A properly constructed wall is essential for normal plant growth and development, as has been exemplified by numerous mutant lines defective in their wall structures. It is therefore fundamentally important to have a clear and comprehensive understanding of how the wall is synthesized, if we are to fully utilize all that plants and plant cell walls have to offer us.

Pectin is a family of galacturonic acid-rich polysaccharides that makes up a major portion of the wall carbohydrates, in addition to cellulose and hemicelluloses. Pectin is probably the most structurally complex polysaccharide in nature, and requires ~70 different transferase activities for synthesis, including glycosyl-, methyl-, and acetyl-transferases. The different types of pectic polysaccharides are homogalacturonan (HG), xylogalacturonan (XGA), apiogalacturonan (AGA), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II). HG is a linear homopolymer of α-1,4-linked galacturonic acid and is the most abundant type of pectin. Chapter 1 provides an overview of HG structure and function, and summarizes the current knowledge in HG biosynthesis based on over 20 years of extensive studies focusing on α-1,4-galacturonosyltransferase (GalAT) activity. Significantly, previous research by Dr. Jason Sterling in our laboratory resulted in the identification of *Arabidopsis thaliana* GALACTURONOSYLTRANSFERASE 1 (GAUT1), a functionally proven HG:GalAT and a member of the GAUT1-related gene family.

The biosynthesis of pectin, and of plant cell walls in general, is thought to be facilitated, at least in part, by protein complexes containing the biosynthetic transferase activities. In consideration of this view, and since this dissertation describes the first pectin biosynthetic enzyme complex, a review of the current literature on the involvement of protein complexes in the biosyntheses of selected polysaccharides from both plants and animals is presented in Chapter 2. This review presents information against which the results of the thesis research can be compared and contrasted.

The goal of the research in this dissertation was to advance our understanding of how pectic polysaccharides, and homogalacturonan in particular, are synthesized. The work described in Chapter 3 focuses on Arabidopsis GAUT1 and its homolog GAUT7, specifically on the protein-protein interaction between two proteins. Results from co-immunoprecipitation and bimolecular fluorescent complementation experiments demonstrate that GAUT1 and GAUT7 interact in a protein complex. In agreement with this finding, GAUT1 and GAUT7 are shown to be co-expressed in most tissues and developmental stages of the Arabidopsis plant, as judged by transcript co-expression analyses employing online microarray databases and by generation and analyses of transgenic Arabidopsis containing promoter:*GUS* gene reporters. The GAUT1:GAUT7 complex exhibits GalAT activity with an acceptor substrate preference for HG oligomers (oligogalacturonan) in comparison to other pectic polysaccharide acceptors (i.e. RG-I and RG-II), thus establishing the enzymatic activity of the GAUT1:GAUT7 complex as HG:GalAT. An Arabidopsis SP-sepharose-purified solubilized membrane preparation (SP fraction) containing GAUT1 and GAUT7 was subjected to non-reducing SDS-PAGE. Western blotting of the separated proteins using anti-GAUT1 and anti-GAUT7 antibodies identified a higher molecular weight protein band at ~185 kDa that is common to both GAUT1 and GAUT7

and is of a larger size than the monomeric forms of GAUT1 and GAUT7 observed by reducing SDS-PAGE. These results indicate involvement of disulfide bond(s) in the formation of the GAUT1:GAUT7 complex. Proteomics analyses of the ~185 kDa protein band identified GAUT1 and GAUT7 as the only components of the non-reducing SDS-PAGE-stable GAUT1:GAUT7 complex.

Upon reducing SDS-PAGE, Arabidopsis GAUT1 resolves as an ~60 kDa protein that is smaller than its predicted size based on DNA sequence. Based on this observation, and other data including GAUT1 N-terminal sequence (see Chapter 3), we propose that Arabidopsis GAUT1 is post-translationally processed into a mature protein by a proteolytic cleavage between Met_{167} and Arg_{168}. Transient expression in *Nicotiana benthamiana* leaves, of several GFP fusion constructs of GAUT1 (full length versus C-terminally truncated) in the presence or absence of GAUT7, demonstrated that the cleaved GAUT1 is retained in the Golgi, the site of HG synthesis, through interaction with GAUT7 in the GAUT1:GAUT7 complex. Based on all available data, we present models of the GAUT1:GAUT7 complex, in which GAUT1 acts as the catalytic HG:GalAT subunit, and GAUT7 serves to anchor GAUT1 to the Golgi membrane. Proteomics analyses of proteins pulled down by anti-GAUT1 and anti-GAUT7 antibodies identified twelve additional proteins that putatively interact with the GAUT1:GAUT7 complex.

CHAPTER 1

LITERATURE REVIEW:

BIOSYNTHESIS OF PECTIN HOMOGALACTURONAN

INTRODUCTION

A carbohydrate-rich extracellular matrix known as the plant cell wall surrounds all plant cells. The cell wall is a dynamic and plastic entity that protects plant cells from the outside environment, defines the cell size and shape, mediates the interactions between cells, and provides mechanical strength and rigidity to allow plants to grow upright. The plant cell wall is categorized into the primary and secondary wall based on wall structure and cell developmental stage. The primary wall is laid down during cell growth and expansion, and is composed of ~90% carbohydrate and ~10% proteins. Upon cessation of growth, some cells lay down secondary cell wall, which is thicker and may be lignified.

Approximately 60 years of plant cell wall research has led to the elucidation of the structures of the wall polysaccharides and the identity of many wall proteins. However, an understanding of how the wall is synthesized, which is essential to manipulate the plant cell wall for agronomic and industrial purposes, is lagging behind. This is especially true for pectin, which is likely the most structurally complex polysaccharide in nature and a major component of plant cell walls along with cellulose and hemicellulose. Pectic polysaccharides comprise up to ~35% of the primary wall of dicots, non-graminaceous monocots, and gymnosperms; up to ~10% of the primary wall of grasses and other commelinids; and up to ~5% of walls in woody

tissues. Pectin is enriched in the middle lamella and cell corners, but is also found throughout the wall (Mohnen, 2008). While it has been estimated that ~70 different enzymatic activities are involved in the biosynthesis of pectin, so far only a handful of these activities have been demonstrated and their encoding genes identified (Mohnen, 2008; Caffall and Mohnen, 2009). This literature review summarizes the current understanding of the biosynthesis of homogalacturonan, the most abundant type of pectin.

OVERVIEW OF PECTIN STRUCTURE AND FUNCTIONS

Pectin constitutes a family of structurally complex polysaccharides characterized by the presence of D-galactopyranosyluronic acid (GalA) residues in α-1,4 linkages. Pectic polysaccharides are structurally classified into homogalacturonan (HG), xylogalacturonan (XGA), apiogalacturonan (AGA), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (Caffall and Mohnen, 2009). HG is a linear homopolymer of α-1,4-linked GalA residues, which are partially methyl-esterified at the C-6 carboxyl and may also be partially *O*-acetylated at the O-2 and/or O-3 (Figure 1.1A) (Ridley et al., 2001). XGA and AGA are substituted galacturonans consisting of an HG backbone with β-D-xylosyl substitutions at the C-3 and β-D-apiosyl substitutions at the C-2 or C-3 of the backbone GalA residues, respectively (Ridley et al., 2001). RG-II is also a branched substituted galacturonan comprised of a short HG chain (7-9 GalA residues) decorated at the C-2 and C-3 by four structurally complex side chains designated A through D (Figure 1.1C) (O'Neill et al., 2004). The complex structure of RG-II, which consists of 12 different types of glycosyl residues connected by more than 20 different linkages, is highly conserved among vascular plants including lycophytes and pteridophytes (O'Neill et al., 2004). In contrast to the other four classes of pectic polysaccharides, RG-I has a

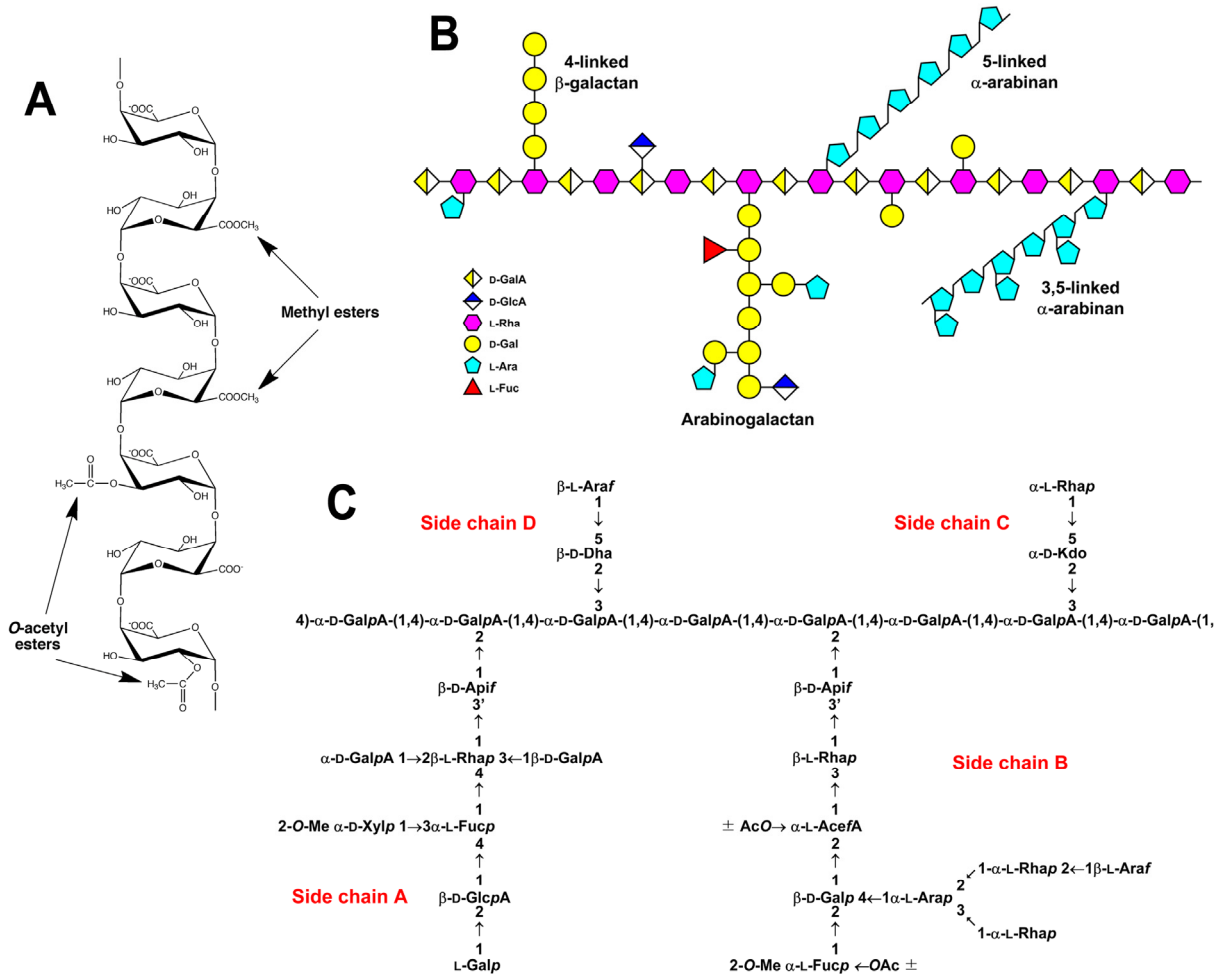


Figure 1.1. Representative structures of the pectic polysaccharides:

(A) homogalacturonan (HG) (adapted from Ridley et al., 2001);

(B) rhamnogalacturonan I (RG-I) (adapted from http://www.ccrc.uga.edu/~mao/rg1/rg1.htm);

(C) rhamnogalacturonan II (RG-II) (adapted from http://www.ccrc.uga.edu/~mao/rg2/intro.htm).

backbone of repeating disaccharides [→4)-α-D-GalA-(1→2)-α-L-Rha-(1→]$_n$, of which 20-80% of the backbone rhamnosyl residues are substituted at the C-4 position with side chains of arabinan, galactan, and/or arabinogalactan (Figure 1.1B) (Mohnen, 1999; Willats et al., 2001). HG is the most abundant type of pectin, accounting for 55-70% of total pectin, while RG-I and RG-II are typically found to be approximately 20-35% and 10% of pectin, respectively (Mohnen, 2008). The other two classes, XGA and AGA, are less prevalent. XGA is detected mostly in storage or reproductive tissues in numerous plant species, although it has also been found in the leaves and stems of *A. thaliana* (Zandleven et al., 2007), while AGA is present in cell walls of aquatic plants such as duckweeds and marine seagrasses (Caffall and Mohnen, 2009).

In the wall, pectic polysaccharides form, via non-covalent as well as covalent interactions, a matrix network in which the load-bearing cellulose/hemicellulose network is embedded. Non-esterified carboxyl groups of HG can engage in ionic interactions with calcium ions (Ca^{2+}), forming junction zones between non-esterified HG chains. This structure is referred to as an "egg-box" structure and may result in the formation of a calcium pectate gel (Thakur et al., 1997). RG-II is known to form a dimer via borate-diester cross-linking between the apiosyl residues of side chain A of two RG-II molecules (Ishii et al., 1999; O'Neill et al., 2004). Domains of HG, XGA, RG-I, and RG-II are believed to be covalently linked to each other through their backbones as illustrated by a model of pectin structure in Figure 1.2, because these polysaccharides can be extracted from the wall by digestion with endopolygalacturonase (EPG) (Caffall and Mohnen, 2009). The presence of covalent linkages between HG and RG-II has been demonstrated (Ishii and Matsunaga, 2001), as has a covalent connection between HG/XGA and RG-I (Nakamura et al., 2002; Coenen et al., 2007). In addition to pectin-pectin interactions, cross-linking has also been reported between pectin and other wall components, such as the

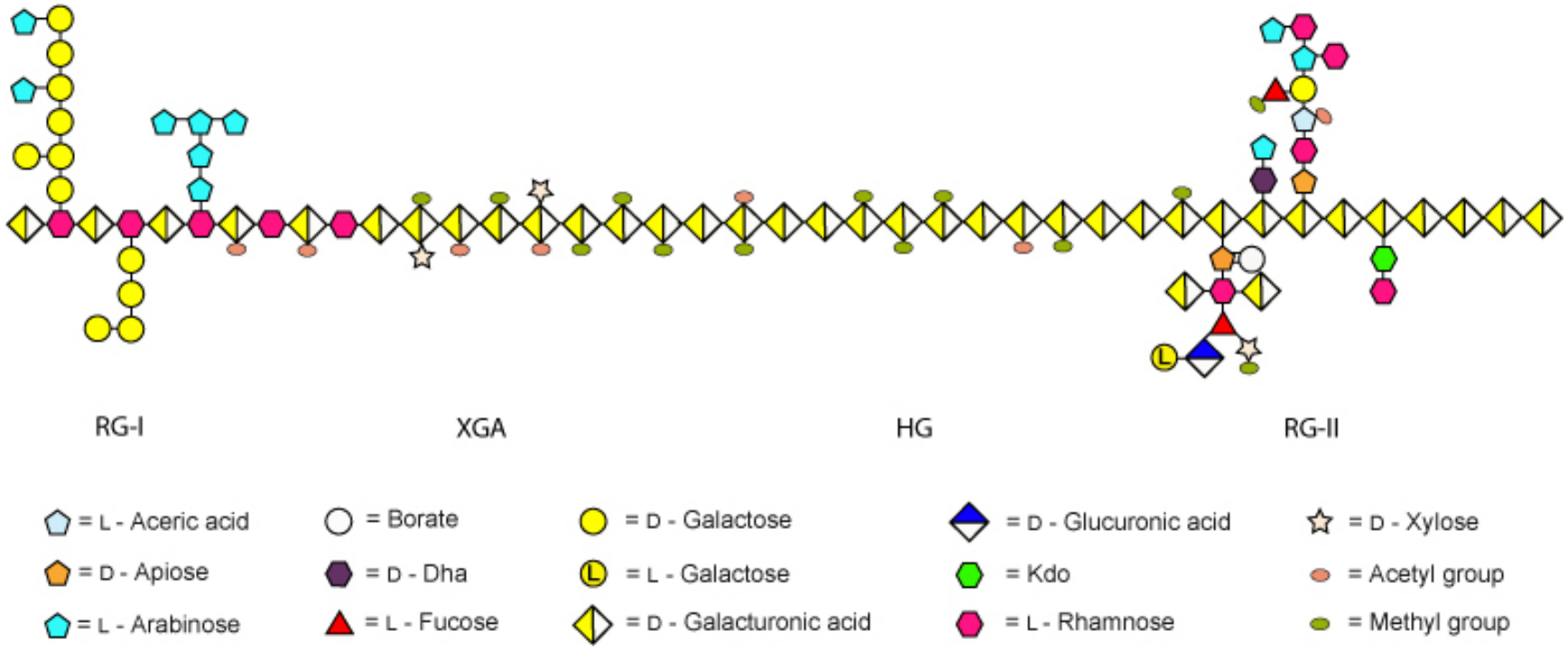


Figure 1.2. A structural model of pectin illustrating the connectivity between HG, XGA, RG-I and RG-II through their backbones. The model is not drawn to scale quantitatively, therefore the HG and RG-I domains should be increased by ~12.5-fold and ~2.5-fold, respectively, to fit the estimated actual amounts of these polymers in the wall. (Adapted from Mohnen, 2008).

hemicelluloses xylan (Nakamura et al., 2002) and xyloglucan (Popper and Fry, 2005, 2008), phenolic compounds, and wall proteins (Caffall and Mohnen, 2009). Nevertheless, the exact ultrastructural organization of the pectic polysaccharides and other components in the plant cell wall is not yet fully understood (Caffall and Mohnen, 2009).

Pectin has been implicated as having numerous physical and biochemical functions in plant growth, development, and defense against pathogens. The pectin matrix is the major determinant of wall porosity and thus the sieving properties of the wall (Bacic et al., 1988; Baron-Epel et al., 1988; McCann and Roberts, 1991). Wall porosity is attributed, at least partially, to Ca^{2+}-bridging of un-esterified HG (Willats et al., 2001) and borate-diester cross-linking of RG-II (O'Neill et al., 2004), and these characteristics are thought to be important for regulating wall permeability and cell-cell signaling (Bacic et al., 1988). RG-II dimerization via a borate-diester linkage is also essential for normal plant growth, as shown by the dwarf phenotype of Arabidopsis *mur1* mutants defective in the formation of this linkage (O'Neill et al., 2001). Enrichment of pectin at the middle lamella and at cellular junctions is believed to be responsible for cell-cell adhesion, the regulation of which is critical for such processes as pollen tube development (Mollet et al., 2000), leaf and flower abscission, and fruit ripening (Knox, 1992; Jarvis et al., 2003). Lower wall pectin content and reduced cell-cell adhesion have been observed in tobacco and Arabidopsis mutants defective in the putative pectin biosynthetic genes *NpGUT1* and *AtQUA1/GAUT8* (Bouton et al., 2002; Iwai et al., 2002; Leboeuf et al., 2005). Pectin is also a reservoir of biologically active oligosaccharides important for plant development and defense. HG in particular, upon enzymatic cleavage in the wall, is the source of oligogalacturonides that have been demonstrated to function as signaling molecules. The physiological responses elicited by oligogalacturonides include induction of ion fluxes, oxidative

burst, phosphorylation (Messiaen et al., 1993; Mathieu et al., 1998; Navazio et al., 2002), production of phytoalexins (Hahn et al., 1981), transcriptional activation of defense genes (Messiaen and Van Cutsem, 1993), fruit ripening (Melotto et al., 1994), as well as regulation of morphogenesis and elicitation of extracellular alkalinization and H_2O_2 accumulation in *in vitro* cell andtissue cultures (Spiro et al., 1998). The wall-associated kinases (WAKs), a class of plasma membrane-embedded serine/threonine receptor kinases, have been implicated as receptors of OGA signaling (Decreux and Messiaen, 2005; Kohorn et al., 2009; Brutus et al., 2010). Members of this kinase family, WAK1 and WAK2, have been shown to bind polygalacturonic acid and oligogalacturonides (DP ≥ 9) in a calcium-induced conformation through the extracellular receptor domains, and to initiate signaling cascades through the intracellular kinase domains.

In addition to its important functions within plants, pectin also has numerous practical and industrial applications. The gelling properties of pectin have been widely exploited for food and other industrial purposes, for example, pectin is used as a gelling agent, thickener, emulsifier, fat/sugar replacer (Thakur et al., 1997), and has pharmaceutical potential such as in drug and fragrance delivery (Liu et al., 2005; Itoh et al., 2007). With regards to biofuel production, as a component of plant cell walls pectin also affects biomass recalcitrance, and it has recently been shown that reducing the proportion of non-methylesterified HG in the wall improves the enzymatic saccharification of plant biomass for bioethanol production (Lionetti et al., 2010).

Multiple benefits of pectic polysaccharides on human health have been described. Pectin has been demonstrated to lower the levels of blood glucose as well as serum lipid and cholesterol, and to inhibit tumor/cancer growth and metastasis (Kumar et al., 1993; Pienta et al.,

1995; Thakur et al., 1997; Sudheesh and Vijayalakshmi, 1999; Nangia-Makker et al., 2002; Jackson et al., 2007; Theuwissen and Mensink, 2008), illustrating the potential use of pectin in managing diseases such as diabetes, cardiovascular, and cancer. Immunomodulating activity and prebiotic properties of pectin have also been reported (Manderson et al., 2005; Paulsen and Barsett, 2005).

BIOSYNTHESIS OF HOMOGALACTURONAN

To better understand the structure and function of pectin in the plant cell wall and to improve the use of pectin in food, pharmaceutical, and other industries, it is essential to know how pectin is synthesized. It has been proposed that at least 67 distinct enzyme activities, comprising glycosyltransferases, methyltransferases, and acetyltransferases, are required to synthesize the complex structure of pectin (Mohnen, 2008). Nevertheless, so far only a handful of these enzymatic activities have been studied, and even less have had the encoding genes conclusively identified (Mohnen, 2008; Caffall and Mohnen, 2009).

The best-studied pectin biosynthetic process to date is the biosynthesis of HG. Seemingly the simplest in structure, HG also serves as the backbone of XGA, AGA, and RG-II, suggesting that all these substituted galacturonan classes of pectin may be built on HG, and thus illustrating the fundamental importance of understanding the biosynthesis of this most abundant pectic polysaccharide. The α-1,4-linked GalA chain of the HG backbone is synthesized by α-1,4-galacturonosyltransferases (GalATs). GalATs transfer GalA from nucleotide sugar precursor uridine diphosphate-GalA (UDP-GalA), which in the cell is made from UDP-glucuronic acid (UDP-GlcA) by UDP-GlcA 4-epimerases (Gu and Bar-Peled, 2004). It is not yet known how the GalA chain is initiated, how the polymerization processes, and how the process is terminated.

HG chains with degrees of polymerization (DP) of up to ~100 GalA residues have been identified in pectin fractions (Yapo et al., 2007). Modification of HG by C-6 carboxyl methylesterification and O-2/O-3 *O*-acetylation occurs through the actions of pectin methyltransferase and acetyltransferase activities, respectively, which use the substrates *S*-adenosyl-methionine (SAM) and acetyl-coenzyme A (acetyl-CoA), respectively, as donors (Caffall and Mohnen, 2009). Furthermore, pectin methylesterase activity in the wall can remove the methyl ester groups from HG, thereby altering the degree and pattern of methylesterification of HG (Willats et al., 2001; Wolf et al., 2009).

HG and the other pectic polysaccharides are synthesized in the Golgi apparatus and subsequently transported via secretory vesicles to the extracellular space for incorporation into the walls. Following incubation of pea seedlings with D-[^{14}C]Glc, radioactive sugars characteristic of pectic polysaccharides, i.e. galactose, arabinose, and galacturonic acid, were detected in isolated Golgi bodies (Harris and Northcote, 1971). α-1,4-Galacturonosyltransferase activity co-localized with latent UDPase activity, a Golgi-specific enzyme marker, in the fractions obtained following sucrose gradient centrifugation of pea (*Pisum sativum*) homogenates, demonstrating that HG synthesis occurs in the Golgi apparatus (Sterling et al., 2001).

Immunocytochemistry studies employing antibodies against several pectic epitopes provide a more detailed picture of pectin biosynthesis. In sycamore maple culture cells, polygalacturonic acid/rhamnogalacturonan I (PGA/RG-I) epitopes were found in the *cis*- and medial Golgi cisternae (Zhang and Staehelin, 1992). However, upon deesterification using Na_2CO_3, heavier labeling of the same epitopes was detected in all Golgi compartments, especially in medial cisternae and onwards. Methylesterified HG epitopes were found

predominantly in the medial cisternae and in lesser amounts in the *trans* cisternae and *trans*-Golgi network (TGN), but not in the *cis* cisternae. On the contrary, RG-I side chain epitopes were detected only in the *trans* cisternae and TGN (Zhang and Staehelin, 1992). These observations indicated that HG is synthesized as early as the *cis*-Golgi cisternae; that HG synthesis continues into the medial Golgi; that methylesterification of the backbone takes place in the same compartments; and that side chain additions occur in the *trans* cisternae (Zhang and Staehelin, 1992; Staehelin and Moore, 1995). This biosynthetic compartmentalization in the plant Golgi stacks may be not absolute, but rather cell-type dependent, since in clover root tips the PGA/RG-I epitopes were observed in the *cis* and medial Golgi cisternae in the cortical cells while the same epitopes were seen in the *trans* cisternae and TGN of the mucilage-secreting root cap cells (Moore et al., 1991; Lynch and Staehelin, 1992; Zhang and Staehelin, 1992; Staehelin and Moore, 1995). Further immunocytochemistry studies on root tips and pollen tubes using antibodies against low- versus highly-methylesterified HG epitopes, demonstrated that upon synthesis, HG is transported via secretory vesicles and deposited in the wall in highly-methylesterified form (Sherrier and VandenBosch, 1994; Staehelin and Moore, 1995; Dolan et al., 1997; Lennon and Lord, 2000; Dardelle et al., 2010). HG in the wall may subsequently be de-esterified by pectin methylesterase activity.

SYNTHESIS OF HOMOGALACTURONAN BACKBONE

Synthesis of the HG backbone by homogalacturonan:α-1,4-D-galacturonosyltransferase (HG:GalAT)

α-1,4-Galacturonosyltransferase (GalAT, EC 2.4.1.43) is the enzyme that synthesizes the HG backbone by catalyzing the transfer of GalA from uridine diphosphate-GalA (UDP-GalA)

Table 1.1. GalAT activity identified in various plant species.

([a]) Enzyme sources tested were microsomal membrane preparations (microsome), detergent-permeabilized microsomes (permeabilized), detergent-permeabilized isolated Golgi (permeabilized Golgi), or detergent-solubilized microsomes (solubilized).

([b]) Exogenous acceptors are differentiated into polygalacturonic acid (PGA) or oligogalacturonic acid (OGA); the latter was fluorescently labeled with 2-aminopyridine (PA) or 2-aminobenzamide (2AB), or immobilized on Sepharose support (imm.). ND – not determined.

Plant species	Enzyme source[a]	pH optimum	Temperature optimum	Apparent Km for UDP-GalA (μM)	Vmax (pmol/min/mg)	Type of acceptor[b]	Reference
Mung bean (*Vigna radiata*)	microsomes	6.3 – 7.0	30°C	1.7	4700	endogenous	(Villemez et al., 1965; Villemez et al., 1966)
	solubilized	ND	ND	ND	ND	endogenous exogenous-pectin, PGA, OGA	(Crombie et al., 2003)
Tomato (*Lycopersicon esculentum*)	microsomes	ND	ND	ND	ND	endogenous	(Lin et al., 1966)
Sycamore (*Acer pseudoplatanus*)	microsomes	ND	ND	770	ND	endogenous	(Bolwell et al., 1985)
Tobacco (*Nicotiana tabacum*)	microsomes	7.8	25°-30°C	8.9	150	endogenous	(Doong et al., 1995)
	solubilized	6.3 - 7.8	ND	37	290	exogenous-OGA	(Doong and Mohnen, 1998; Scheller et al., 1999)
Pea (*Pisum sativum*)	microsomes	ND	ND	ND	ND	endogenous	(Abdel-Massih et al., 2007)
	permeabilized Golgi	ND	ND	ND	ND	exogenous-OGA	(Sterling et al., 2001)
	solubilized	7.0	30°C	170	0.7	exogenous-OGA-PA	(Ohashi et al., 2007)

Azuki bean (*Vigna angularis*)	permeabilized	6.8 - 7.8	25°-35°C	140	2700	exogenous-PGA	(Takeuchi and Tsumuraya, 2001)
Pumpkin (*Cucurbita moschata*)	permeabilized	6.75 - 7.25	25°-35°C	700	7000	exogenous-OGA-2AB	(Ishii, 2002)
Petunia (*Petunia axillaris*)	solubilized	7	30°C	170	480	exogenous-OGA-PA	(Akita et al., 2002)
Arabidopsis thaliana	microsomes	ND	ND	ND	ND	endogenous	(Orfila et al., 2005)
	solubilized	ND	ND	ND	ND	exogenous-OGA-imm.	(Guillaumie et al., 2003)
	solubilized	ND	ND	ND	ND	exogenous-OGA	(Sterling, 2004; Orfila et al., 2005; Sterling et al., 2006)

onto endogenous and exogenous HG acceptors (Doong and Mohnen, 1998). Membrane-bound GalAT activity has been described in several plant species as summarized in Table 1.1. Enzyme sources prepared using different methods (microsomal membranes, permeabilized-membranes, and solubilized membranes) were used in these studies, however, all these GalATs appear to have similar pH and temperature optima, and strictly require Mn^{2+} for catalysis. The variability in kinetic parameters among these enzymes may reflect the diverse plant materials and species from which the enzymes were isolated as well as the different methods used to assay the GalAT activity. With regards to the direction of chain elongation, it has been shown that GalAT *in vitro* polymerizes HG by addition of new GalA residues onto the non-reducing end of the growing polymer (Scheller et al., 1999; Akita et al., 2002; Ishii et al., 2002). As mentioned in the previous section, GalAT activity was detected in the Golgi-enriched fraction from pea membranes and, upon proteinase K digestion of the Golgi vesicles in the presence or absence of detergent, was shown to have its catalytic site within the Golgi lumen (Sterling et al., 2001). Characterization of the product generated by tobacco membrane-bound GalAT revealed ~50% esterification of the product, of which at least 40% was found to be methyl esters (Doong et al., 1995), demonstrating the cooperation of different enzyme activities in the biosynthesis of HG.

While GalATs in microsomal membrane preparations were able, in the presence of UDP-GalA, to utilize endogenous acceptors to give reaction products of high molecular weight (e.g. ~105 kDa in tobacco (Doong et al., 1995) and >500 kDa in pea (Sterling et al., 2001)), solubilized GalATs were shown to be dependent on the presence of exogenously-added acceptor, preferably OGAs with degrees of polymerization (DP) ≥10 (Doong and Mohnen, 1998; Akita et al., 2002; Ishii, 2002; Ohashi et al., 2007). An exception was the solubilized GalAT from mung bean hypocotyls, which was demonstrated to have a low level of activity in the absence of

exogenous acceptor (Crombie et al., 2003). The degree of GalA incorporation onto the elongating OGA acceptors by solubilized GalATs appears to be dependent on the molar ratio between the donor substrate UDP-GalA and the acceptor substrate OGAs. For example, when acceptor OGAs were present in molar excess in comparison to UDP-GalA, or when a relatively low molar excess of UDP-GalA was used in comparison to OGAs (UDP-GalA:OGA molar ratios of 1:54 and 10:1, respectively), solubilized GalAT from tobacco was able to add only one GalA residue to OGA acceptors (Doong and Mohnen, 1998). A similar situation was also noted with pea GalAT, specifically between intact Golgi and detergent-permeabilized Golgi membranes (Sterling et al., 2001). In contrast, when UDP-GalA was used in relatively high molar excess (>20-fold) compared to OGAs, solubilized enzymes from petunia and Arabidopsis were able to successively transfer >10 GalA residues onto the OGAs (Akita et al., 2002; Sterling et al., 2006). The fact that the solubilized GalATs require, and are greatly activated by, addition of exogenous OGA acceptors with preference of DP ≥10, indicates that the solubilized enzymes have lost access to endogenous acceptor and are not capable of initiating *de novo* polymerization of HG. Thus, the solubilized GalATs are most likely functioning in the elongation step of the HG biosynthesis.

The above results indicate that permeabilized and solubilized GalATs lack the ability to generate the high molecular weight reaction products produced by the membrane-bound enzymes (Doong et al., 1995; Doong and Mohnen, 1998; Sterling et al., 2001; Akita et al., 2002; Ishii, 2002). One possible reason may be that the exogenous OGA acceptors are not the ideal substrate for the GalAT (Doong and Mohnen, 1998; Sterling et al., 2001). Another possible reason for this enzymatic difference may be that GalAT activity *in vivo* occurs via an enzyme complex which may still be intact in microsomes (Doong and Mohnen, 1998; Sterling et al.,

2001). Permeabilization or solubilization of the microsomal membranes may disrupt the complex, resulting in a lack of access to endogenous acceptor and/or a loss of potential processivity (Doong and Mohnen, 1998; Sterling et al., 2001).

It has recently been demonstrated that GalAT activity may indeed exist in high molecular weight, presumably multi-protein, complexes (Ohashi et al., 2007). Gel filtration chromatography of CHAPS-solubilized membrane proteins from azuki bean reproducibly resulted in three peaks associated with GalAT activity (referred to as polygalacturonic acid synthase by the authors) which, upon subjection to sucrose density gradient centrifugation, were determined to correspond to molecular masses of 21,000 kDa, 5,000 kDa, and 590 kDa, respectively. The 590 kDa peak appears to correspond to a component of the two larger peaks, since gel filtration chromatography following pre-incubation of the peak fractions with 500 mM NaCl caused the two larger peaks to resolve at the same position as the 590 kDa peak, while the 590 kDa peak itself did not shift (Ohashi et al., 2007). These results suggest that the structural integrity of the putative larger complexes may involve ionic interactions (Ohashi et al., 2007). The authors, unfortunately, did not determine the protein components of the described complexes.

ARABIDOPSIS GALACTURONOSYLTRANSFERASE 1 (GAUT1) AND THE GAUT GENE FAMILY

The first functional identification of a gene encoding a HG:GalAT resulted from a combination of proteomic and biochemical studies (Sterling et al., 2006). Two putative glycosyltransferases, namely galacturonosyltransferase 1 and 7 (GAUT1 and GAUT7, respectively), were identified from trypsin digests of a partially purified, GalAT-enriched,

Arabidopsis solubilized enzyme preparation followed by tandem mass spectrometry of the resulting tryptic peptides. Based on its amino acid sequence, GAUT1 (At3g61130) is predicted to be a protein of 673 amino acids (a.a.), with a predicted molecular mass of 77.4 kDa and a predicted pI of 9.95. It is also predicted to have type II transmembrane topology and has been shown to localize to the Golgi apparatus (Dunkley et al., 2004; Dunkley et al., 2006), which is in agreement with the location of pectin biosynthesis as suggested by immunocytochemistry (Moore et al., 1991; Zhang and Staehelin, 1992) and as shown by subcellular fractionation of GalAT activity in pea (Sterling et al., 2001). Transient expression of an N-terminally-truncated construct of GAUT1 in mammalian HEK293 cells resulted in secretion of the recombinant enzyme and recovery of GalAT activity in the growth media (Sterling et al., 2006). Furthermore, anti-GAUT1 polyclonal antibodies specifically immunoabsorbed GalAT activity from a SP-Sepharose-purified, GalAT-enriched fraction of Arabidopsis solubilized membranes. Depletion of GAUT1 coincided with a decrease in the GalAT activity in the supernatants, while both the protein and enzyme activity were recovered in the pellets. Sensitivity of the products made by the immunoabsorbed enzyme to a HG-specific exopolygalacturonase, confirmed that the anti-GAUT1-immunoprecipitated activity was indeed α-1,4-GalAT (Sterling et al., 2006). These data provided conclusive biochemical evidence that GAUT1 is a HG:GalAT (Sterling et al., 2006).

BLAST analysis of GAUT1 against the Arabidopsis genome revealed 14 other Arabidopsis proteins (GAUTs 2 – 15) that share 36-68% a.a. sequence identity and 56-84% sequence similarity with GAUT1 (Sterling et al., 2006). These proteins, with GAUT1, constitute the Arabidopsis GAUT gene family. Additionally, 10 other Arabidopsis proteins with less similarity to GAUT1 (23-29% a.a. sequence identity and 42-53% similarity) were also identified.

These latter proteins form the Arabidopsis GAUT-like (GATL) gene family that is distinct from the GAUT gene family (Sterling et al., 2006) and thus will not be covered further in this review, except when comparison to the GAUT gene family is warranted. The GAUT (and also GATL) gene family belong to glycosyltransferase family (GT) 8, a family of retaining glycosyltransferases according to the CAZy database (http://www.cazy.org/), and characterized by the presence of the Pfam Glyco_transf_8 (PF01501) domain (Sterling, 2004; Sterling et al., 2006; Cantarel et al., 2009). Orthologs of Arabidopsis GAUT1 and other GAUT family members exist in all other plant species studied, including in the moss (*Physcomitrella patens*) and the spike moss (*Selaginella moellendorffii*) but not in most green algal species (Yin et al., 2010). Phylogenetic analysis of the GAUT homologs from *Arabidopsis thaliana*, rice (*Oryza sativa*), and poplar (*Populus trichocarpa*) identified seven clades in the GAUT gene family (Figure 1.3) (Caffall, 2008; Caffall et al., 2009), which were confirmed by further phylogenetic studies involving more GAUT homologs from other plant species whose genomes were recently sequenced (Yin et al., 2010; Yin et al., In press). Within the GT8 family, the plant cell wall synthesis-related GAUT and GATL clades are distant from the other plant GT8 clades, i.e. galactinol synthase (GolS) and plant glycogenin-like starch initiation protein (PGSIP), as shown in a phylogenetic analysis of 378 plant GT8 proteins from 15 genomes (Yin et al., 2010). This suggests that the GAUT and GATL clades may have evolutionarily diverged relatively early from the rest of plant GT8 members, and may have developed a distinct catalytic mechanism compared to the other clades.

All the GAUT proteins share the GAUT family motif (H-x(2)-[ILV]-x-[ST]-D-N-[IV]-[IL]-A-[ASTV]-S-V-V-[AIV]-x-S-x-[AIV]-x(2)-[AS]-x(2)-[PS]-x(3)-V-[FL]-H-[ILV]-[ILV]-T-[DN]-x(2)-[NST]-x(2)-[AGP]-[IM]-x(3)-F) and contain the DxD motif, which has been

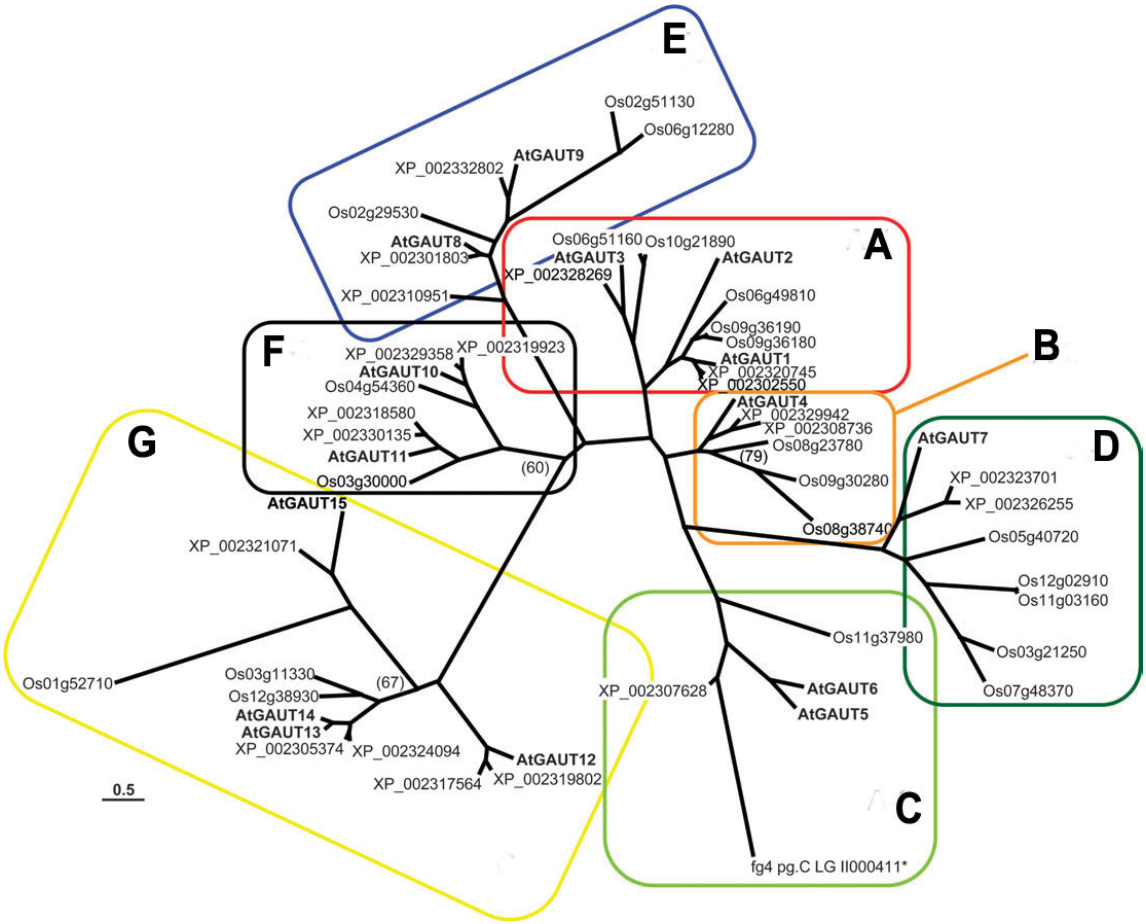


Figure 1.3. A phylogenetic tree of the GAUT gene family identifying seven clades: clade A (GAUTs 1-3), clade B (GAUT4), clade C (GAUTs 5 and 6), clade D (GAUT7), clade E (GAUTs 8 and 9), clade F (GAUTs 10 and 11), and clade G (GAUTs 12-15). (Adapted from Caffall et al., 2009).

identified in many well-characterized glycosyltransferases and is thought to be involved in the coordination of divalent cation required for catalysis (Sterling, 2004; Sterling et al., 2006). Similar to GAUT1, the other Arabidopsis GAUT proteins are predicted to be 61-78 kDa in size, and most are likely to have type II transmembrane protein topology with the putative transmembrane and catalytic domains located at the N- and C-terminal regions, respectively. One exception is GAUT2, which is predicted to have neither a transmembrane domain nor a signal peptide. The Arabidopsis GATL proteins, in contrast, are predicted to be 39-44 kDa in size and to have only N-terminal signal peptides, suggesting that they are soluble, secreted proteins (Sterling et al., 2006). However, two poplar orthologs of Arabidopsis GATL1/PARVUS have recently been shown to reside in the endoplasmic reticulum (ER) and Golgi apparatus, indicating that a mechanism exists to retain these supposedly soluble proteins within the secretory organelles (Kong et al., 2009; Lee et al., 2009).

Functions of other GAUT family members

GAUT1 has been functionally established as a HG:GalAT. The other GAUT proteins have been proposed to also function as GalATs that catalyze the transfer of GalA from UDP-GalA onto various acceptors to synthesize the different types of pectic polysaccharides (Sterling et al., 2006). *GAUT* genes appear to be expressed at some level in all major tissues of Arabidopsis as indicated by a transcript expression study combining data from publicly available microarray databases and from RT-PCR, with the exception of *GAUT2*, which was suggested to be a non-functional truncated homolog of *GAUT1* (Caffall et al., 2009). Glycosyl residue composition analyses of 26 homozygous T-DNA mutant lines that represent 13 *GAUT* genes revealed eight *gaut* mutants, namely *gauts 6, 8, 9, 10, 11, 12, 13*, and *14*, having significantly

different wall content (in mol%) of GalA, Xyl, Rha, Gal, or Ara, in comparison to wild-type (Caffall et al., 2009). The data suggested that mutations in the corresponding *GAUT* genes influence wall pectin and/or xylan polysaccharides in the mutants (Caffall et al., 2009). Intriguingly, despite the changes in wall compositions, most of the mutants did not show any major growth phenotypes different from wild-type, except for *gaut8* and *gaut12* which showed severely dwarfed growth phenotypes and sterility. The lack of a severe phenotype for most of the *GAUT* mutants may be due to the functional redundancy among members of the gene family (Caffall et al., 2009). Unfortunately, homozygous mutant lines for *GAUT1* and *GAUT4* were not available/obtainable when this research work was undertaken, suggesting that these genes encode proteins that are crucial for plant growth and development, and thus, that mutations in these genes may have caused lethality. More detailed studies of several of the Arabidopsis *GAUT* genes, mostly through mutant analyses, have also been reported and are summarized below.

Evaluation of the growth phenotype of Arabidopsis T-DNA mutants of *GAUT6 (At1g06780)* revealed that in comparison to wild-type, *gaut6* mutants had slightly reduced stem length, decreased amounts of water-extractable mucilage, and increased phenolic content in the seeds (Caffall, 2008). Preliminary analysis of *gaut6* total cell wall extracted from leaves, inflorescences, and siliques, revealed a significant reduction in GalA content accompanied by increases in Ara, Rha, and Xyl residues. A similar trend was also observed in the water-extracted seed mucilage. Furthermore, a considerable increase in the ratio of GalA:Rha was observed in the pectin-enriched wall fractions. Taken together, these results suggest a lower total pectin content, with an increased proportion of HG versus RG-I in the mutant walls in comparison to wild-type. Preliminary data from heterologous expression of GAUT6 indicates that it encodes a HG:GalAT, which may explain the above reduction in pectin content in the

mutant (Caffall, 2008). Microsomal membranes extracted from *Nicotiana benthamiana* leaves transiently transfected with GAUT6 constructs have higher HG:GalAT activity in comparison to extracts from *Nicotiana* leaves transfected with control vectors. Significantly, stable expression of truncated constructs of GAUT6 in *E. coli* also yielded a detectable level of HG:GalAT activity. Verification of these preliminary results would confirm GAUT6 identity as a HG:GalAT, and will represent the first demonstration of successful heterologous expression of a pectin biosynthetic glycosyltransferase in *E.coli*.

As mentioned above, GAUT7 is the other glycosyltransferase, besides GAUT1, that was identified in GalAT activity-containing, partially purified Arabidopsis solubilized membrane protein preparations. Arabidopsis GAUT7 (At2g38650) contains 619 a.a., with a predicted molecular mass of 69.7 kDa, a predicted pI of 8.6, and 36% a.a. sequence identity and 60% sequence similarity to GAUT1 (Sterling et al., 2006). Arabidopsis GAUT7, along with two poplar and five rice orthologs, form a clade separate from other Arabidopsis GAUTs (Figure 1.3) (Caffall et al., 2009), which may suggest a specific function of GAUT7 that is not shared by any other GAUT. It is noteworthy that the larger number of GAUT7 orthologs in rice may indicate a more critical GAUT7 function in monocots compared to dicots. In contrast to GAUT1, however, transient expression of GAUT7 in HEK293 cells did not yield GalAT activity (Sterling et al., 2006). Intriguingly, a recent study of the predicted tertiary structure of GAUT1 (see below) provides an indication that GAUT7 may be non-catalytic due to the absence of a proposed catalytic amino acid residue that may serve as a nucleophilic acceptor (Yin et al., 2010). Whether GAUT7 has a catalytic activity that is undetectable in the GalAT assay or, conversely, has no catalytic activity, remains to be determined. Unfortunately, analyses of two homozygous *GAUT7* T-DNA mutant lines did not reveal any difference in growth phenotype or wall glycosyl

residue composition compared to wild-type. This could be due to the fact that the *gaut7* mutants used were knock-down rather than knock-out lines (Caffall et al., 2009). Interestingly, it was also reported that RT-PCR of *GAUT7* resulted in two bands: one of expected size and the other of a smaller size. Whether the smaller band represents a splice variant of GAUT7 remains to be determined. Although *gaut7* mutants had no growth phenotype and GAUT7 did not exhibit GalAT activity upon heterologous expression, the fact that GAUT7 was co-identified with GAUT1 was noted and therefore research presented in this thesis investigates the possibility that the two proteins might work together in a protein complex (see Chapter 3).

Multiple published studies of T-DNA insertion mutants of *QUASIMODO1 (QUA1) / GAUT8 (At3g25140)* have implicated *GAUT8* in HG synthesis. Allelic mutants *qua1-1* and *qua1-2* display dwarfed growth and reduced cell-cell adhesion. This phenotype is more severe in the null mutant *qua1-2* and leads to a high frequency of seed abortion (Bouton et al., 2002). The *qua1* mutants also manifest weaker stems and have more friable calli than wild-type (Leboeuf et al., 2005; Orfila et al., 2005). Analyses of *qua1-1* whole seedlings, leaves, stems, and suspension-cultured cells revealed a 12-30% reduction in total uronic acid and/or GalA content in comparison to wild-type, suggesting a deficiency in pectin HG in the mutant (Bouton et al., 2002; Leboeuf et al., 2005; Orfila et al., 2005). This conclusion was supported by the observation that compared to wild-type: 1) *qua1-1* showed less immunolabeling by antibodies JIM5, JIM7, and 2F4, which recognize low-methyl esterified, highly-esterified, and Ca^{2+}-associated HG, respectively (Bouton et al., 2002; Leboeuf et al., 2005; Orfila et al., 2005), and 2) lower GalAT activity was present in *qua1-1* stem microsomal and detergent-solubilized membrane preparations (23% and 33% reduction compared to wild-type, respectively) (Orfila et al., 2005). Additionally, *qua1-1* showed defects in root tip border-like cells, which were released

as single cells, in contrast to wild-type where they were attached together (Durand et al., 2009). The *qua1-1* mutant phenotype was associated with decreased cell wall HG in the mutant border-like cells and with increased secretion of mucilage rich in xylogalacturonan and arabinogalactan-protein, as determined by immunocytochemistry using JIM5 antibody and Fourier transform infrared microspectroscopy (Durand et al., 2009). A marked reduction (~40% less than wild-type) was also observed in β-1,4-D-xylan synthase activity in *qua1-1* stem microsomes (Orfila et al., 2005). The phenotype was more complex, however, in inflorescence stem where *qua1-1* cell walls had an ~30% reduction in GalA, but only an ~7% reduction in Xyl content. Immunolabeling of stem sections using LM10 (a monoclonal antibody against unsubstituted xylan) did not showed any difference between *qua1-1* and wild-type (Orfila et al., 2005). While a connection between HG and xylan syntheses is clearly evident from these results, the more comprehensive effects on HG suggest that the *qua1-1* mutation has a more direct effect on HG rather than on xylan synthesis (Orfila et al., 2005). A demonstration of the enzymatic activity, in purified or heterologously expressed protein, should serve to clarify the biochemical function of QUA1/GAUT8.

A seed mucilage phenotype was observed in an Arabidopsis mutant of *GAUT11 (At1g18580)* (Caffall, 2008; Caffall et al., 2009). Upon staining with ruthenium red solution, which binds to the negative charge in the GalA residues in pectin, ~30% of *gaut11-2* seeds displayed substantially thinner mucilage than wild-type. The remainder of the seeds (~70%) had almost negligible mucilage. Mucilage extracted from *gaut11-2* with hot water had a similar amount of total carbohydrate compared to wild-type, but was significantly reduced in total uronic acid, GalA, and Xyl, slightly reduced in Rha, and increased in Man and Gal content. These observations suggest a role for GAUT11 in seed mucilage production and extrusion.

Another example of the connection between HG and xylan syntheses is demonstrated by studies of the IRREGULAR XYLEM 8 (IRX8) / GAUT12 (At5g54690). *IRX8/GAUT12* has been implicated in secondary cell wall biosynthesis due to its co-expression with the secondary cell wall cellulose synthases (Arabidopsis *CesA4, 7*, and *8*) (Brown et al., 2005; Persson et al., 2005). Transcript analyses by RT-PCR, *in situ* hybridization, and promoter:*GUS* expression showed a high level of *IRX8/GAUT12* expression especially in tissues undergoing secondary wall formation, such as the vascular tissues of stems, roots, leaves and inflorescences, and more specifically in the developing xylem cells in elongating internodes, in xylem and interfascicular fiber cells in stems, and in developing secondary xylem in roots (Pena et al., 2007; Persson et al., 2007). Mutant plants are dwarfed with collapsed xylem vessels due to a severely reduced thickness of the secondary walls surrounding the xylem vessels (Pena et al., 2007; Persson et al., 2007). Homozygous mutants are sterile, thus necessitating the isolation of homozygous seeds from heterozygous plants (Persson et al., 2007). Glycosyl residue composition and linkage analyses of *irx8/gaut12* mutant walls indicate a marked reduction in glucuronoxylan content, as well as a notable decrease in the GalA content of endogalacturonase-soluble, pectin-rich wall fractions, suggesting a loss of a subfraction of HG (Persson et al., 2007). Furthermore, *irx8/gaut12* appears to produce fewer numbers of glucuronoxylan chains, most of which lack a pentasaccharide sequence [4-β-D-Xyl-1,4-β-D-Xyl-1,3-α-L-Rha-1,2-α-D-GalA-1,4-D-Xyl] that is present at the reducing end of the glucuronoxylan chain and is hypothesized to function either as a primer or a terminating cap sequence for xylan biosynthesis (Pena et al., 2007). IRX8/GAUT12 may be involved in xylan biosynthesis by catalyzing the incorporation of GalA residues into the pentasaccharide sequence, and/or may be involved in synthesizing a subfraction

of HG (Pena et al., 2007; Persson et al., 2007). A demonstration of IRX8/GAUT12 enzymatic activity will identify the biochemical function of the protein.

A predicted tertiary structure of Arabidopsis GAUT1

An *in silico* prediction of the tertiary structure of Arabidopsis GAUT1 was recently reported (Yin et al., 2010). Amino acid sequence positions 334-647 of GAUT1 were threaded using nine different methods onto the only two solved crystal structures of the GT8 family, namely *Neisseria meningitidis* α-1,4-galactosyltransferase (LgtC; PDB-ID: 1GA8) involved in lipooligosaccharide synthesis, and rabbit muscle glycogenin (PDB-ID: 1LL2). The predicted GAUT1 structure, superimposed on the structure of LgtC, revealed a putative structural binding pocket with more conserved residues in the inside of the pocket and less conserved residues on the outside of, and away from, the pocket. The previously identified DxD and HxxGxxKPW motifs, which are shared among the GT8 family members, were located close to each other in the pocket, suggesting that both motifs may be involved in coordinating substrates during catalysis. Another motif present in the binding pocket is the LPP motif that is conserved within the GAUT and GATL families. This motif is replaced by DQD and DQG in the LgtC and glycogenin structures, respectively, of which the Gln (Q) residue is thought to act as a nucleophilic acceptor that receives monosaccharide from the NDP-sugar donor to form a glycosyl-enzyme intermediate and subsequently transfers it to the acceptor substrate. While the Leu and Pro in the LPP motif are unlikely to play the role as a nucleophile, a Thr or Ser residue just upstream of the LPP motif in the GAUTs and GATLs, respectively, were speculated to be the nucleophilic acceptors due to their strategic positions in these proteins based on the predicted GAUT1 structure. Intriguingly, GAUTs 5, 7, and 15 do not have the Thr residue, but instead have a Gly

or Ala in the corresponding positions that may render these proteins non-catalytic. While biochemical proof for this speculation is needed, it may explain the above-mentioned lack of GalAT activity from GAUT7 expressed in HEK293 cells.

A possible acceptor binding site was predicted to reside in a flexible region of the proposed GAUT1 structure, supported by the presence in this area of a patch of positively charged residues that may interact with the negatively charged HG acceptor (Yin et al., 2010). A corresponding flexible region is involved in dimer formation in rabbit glycogenin (Gibbons et al., 2002), raising the possibility that this region in GAUT1 may function similarly in the formation of the GalAT enzyme complex.

MODIFICATIONS OF HOMOGALACTURONAN DURING SYNTHESIS

Methylesterification of the HG backbone by homogalacturonan:methyltransferases (HG-MTs)

It is well accepted that HG is deposited in the wall in a highly methylesterified form, due to the action of HG-MTs during HG biosynthesis (Kauss and Swanson, 1969; Sherrier and VandenBosch, 1994; Staehelin and Moore, 1995; Dolan et al., 1997; Lennon and Lord, 2000; Dardelle et al., 2010). This modification neutralizes the negative charges on the carboxyl groups in the GalA residues. In the wall, removal of the methyl esters by pectin methylesterases (PMEs), or a lack thereof when hampered by pectin methylesterase inhibitors (PMEIs), determines the pattern and degree of methylesterification of HG, which in turn affects the properties of pectin and of the wall (Willats et al., 2001; Wolf et al., 2009). In particular, un-methylesterified stretches of HG can engage in a gel-forming calcium cross-linking, which serves to strengthen the wall (Jarvis, 1984), or otherwise renders the HG chains accessible to degrading enzymes (e.g. polygalacturonases, pectin/pectate lyases) (Wolf et al., 2009) such as

during fruit ripening. A lower degree of methylesterification (DM) has been correlated with increased viscosity and calcium-induced gelling of pectin (Ralet et al., 2003), increased cell:cell adhesion (Liners et al., 1994), and reduced cell elongation (Derbyshire et al., 2007). On the other hand, both the pattern (e.g. blockwise versus random distribution resulting from the actions of plant and fungal PMEs, respectively) and the degree of methylesterification play roles in determining the strength, elasticity, water holding capacity, and porosity of calcium-pectin gels (Willats et al., 2001). It was recently demonstrated that overexpression of a PMEI in various plant species decreased the content of un-methylesterified HG, which led to reduced recalcitrance and increased enzymatic saccharification efficiency of plant biomass for biofuel production (Lionetti et al., 2010).

Pectin methyltransferase (PMT; categorized under polysaccharide *O*-methyltransferase, EC 2.1.1.18) catalyzes the transfer of a methyl group (CH_3) from the donor substrate *S*-adenosyl-L-methionine (SAM) onto the C-6 carboxyl group of α-1,4-linked GalA residues in pectin acceptors. Membrane-associated PMT activities have been demonstrated in seedling hypocotyls of flax and soybean, in seedling epicotyls of pea, and in suspension culture cells of flax and tobacco (summarized in Table 1.2). However, despite efforts to isolate, solubilize, purify, and characterize this activity, no gene encoding the enzyme has yet been definitively identified.

PMT activity was first described in a microsomal preparation from 3-4 days old mung bean etiolated seedlings (Kauss and Hassid, 1967; Kauss et al., 1967). The incubation of microsomes from mung bean and other species with [^{14}C]SAM (Table 1.2) resulted in incorporation of radiolabeled methyl groups onto endogenous acceptors. The PMT reaction products were shown to be water soluble, extractable by 0.5% ammonium oxalate, and precipitable by 75% EtOH, suggesting that the acceptor substrates were high molecular weight

Table 1.2. Pectin methyltransferase activity identified in various plant species.

([a]) Enzyme sources tested were microsomal membrane preparations (microsomes), freeze-thawed microsomes (freeze-thawed), detergent-permeabilized microsomes (permeabilized), detergent-solubilized microsomes (solubilized), intact Golgi vesicles (Golgi), or permeabilized Golgi vesicles (permeabilized Golgi). ([b]) PGA – polygalacturonic acid; DE – degree of esterification. ND – not determined. SAM – *S*-adenosyl methionine.

Plant species	Enzyme source[a]	pH optimum	Temperature optimum	Apparent Km for SAM (μM)	Vmax for SAM (pmol/min/mg)	Type of acceptor[b]	Reference
Mung bean (*Vigna radiata*)	microsomes	6.8	ND	59 (40-70)	2.7	endogenous	(Kauss and Hassid, 1967; Kauss et al., 1967; Kauss and Swanson, 1969)
Flax (*Linum usitatissimum* L.)	microsomes	7.1	ND	10-30	ND	endogenous	(Vannier et al., 1992; Schaumann et al., 1993; Bruyant-Vannier et al., 1996)
	freeze-thawed	7.0	ND	20	ND	exogenous-high DE pectin	(Bourlard et al., 1997)
		5.5	ND	20	ND	exogenous-low DE pectin	
		5.5	ND	20	ND	exogenous-PGA	
	solubilized	7.1	ND	0.5	ND	exogenous-PGA	(Bruyant-Vannier et al., 1996)
	solubilized	6.5	ND	ND	ND	exogenous-high DE pectin	(Bourlard et al., 2001)
		5.0	ND	ND	ND	exogenous-low DE pectin	
	solubilized	7.0-7.5	ND	ND	ND	exogenous-RG-I	(Bourlard et al., 1997)
		6.5-8.0	ND	ND	ND	exogenous-RG-II	

Tobacco (*Nicotiana tabacum* L.)	microsomes	7.8	25°-40°C	38	48.6	endogenous	(Goubet et al., 1998)
	solubilized	7.8	25°-30°C	18	7.26	exogenous-HG and low DE pectin	(Goubet and Mohnen, 1999)
Soybean (*Glycine max* Merr.)	permeabilized	6.8	35°-40°C	230	1360	exogenous-low DE pectin	(Ishikawa et al., 2000)
Pea (*Pisum sativum*)	Golgi permeabilized Golgi	ND	ND	ND	ND	exogenous-PGA	(Ibar and Orellana, 2007)

pectin (Kauss and Hassid, 1967; Kauss et al., 1967; Vannier et al., 1992; Goubet et al., 1998). The PMT reaction products were also largely digestible by endopolygalacturonase, co-eluted with commercially available pectin and PGA upon size-exclusion chromatography, and shown by glycosyl residue composition analysis to be relatively GalA-rich with GalA/Rha ratios >1 (Vannier et al., 1992; Goubet et al., 1998). These characteristics are consistent with the acceptor substrates being HG-containing pectin. Furthermore, treatment of the PMT reaction products with base or pectin methylesterase (PME) released the methyl groups as methanol, as identified by gas chromatography, demonstrating that the methyl groups were incorporated onto HG in an ester, rather than ether, linkage (Kauss and Hassid, 1967; Kauss et al., 1967; Goubet et al., 1998). It was observed that highest PMT activity corresponded with the onset of the growth phase of cell cultures and with the cell elongation stage of growth in seedlings (Kauss and Hassid, 1967; Schaumann et al., 1993; Goubet et al., 1998).

Studies with flax, tobacco, and pea microsomes demonstrated that the PMT activity is Golgi-localized and that the catalytic domain faces the Golgi lumen. PMT activity co-fractionated with the Golgi-specific enzyme markers (β-glucan synthase I (β-GS I), latent IDPase, and UDPase) upon membrane fractionations of flax and tobacco microsomes using a linear sucrose gradient centrifugation (Vannier et al., 1992; Goubet et al., 1998). Treatment of intact Golgi vesicles from tobacco and pea with proteinase K or trypsin diminished the PMT activity only in the presence of detergent, indicating that the PMT catalytic domain resides within the vesicles (Goubet et al., 1998; Ibar and Orellana, 2007). Moreover, inclusion in the PMT reaction of exogenous pectic acceptor substrates, which cannot pass through intact Golgi membranes, affected the PMT activity only when the membrane integrity was compromised, e.g. by freeze-thawing or by the addition of detergent which permeabilized the membranes (Bourlard

et al., 1997; Goubet et al., 1998; Ishikawa et al., 2000; Ibar and Orellana, 2007). SAM synthesis occurs in the cytosol (Wallsgrove et al., 1983), and the transport of SAM into the lumen of the Golgi and its subsequent use in pectin methylation have been demonstrated (Ibar and Orellana, 2007). The subcellular localization and membrane topology of the PMT activity agrees with the previous immunocytochemistry observation suggesting the location of pectin methylation in the medial-Golgi (Zhang and Staehelin, 1992; Staehelin and Moore, 1995), and is similar to the reported topology of HG:GalAT activity (Sterling et al., 2001). It is noteworthy that with mung bean and tobacco microsomes, addition of UDP-GalA in the PMT reaction mixture increased the incorporation of [^{14}C]-methyl groups onto endogenous pectic acceptors (Kauss and Hassid, 1967; Goubet et al., 1998). However, inclusion of SAM in the HG:GalAT reaction mixture did not stimulate HG synthesis (Kauss and Swanson, 1969; Doong et al., 1995), indicating that the polymerization of HG precedes its methylesterification.

Contrary to the strict divalent cation requirement of HG:GalAT, the presence of divalent cations Mn^{2+} or Mg^{2+} gave variable responses with regards to the PMT activity. For example, there was no effect observed with Mn^{2+} or Mg^{2+} on the activity of mung bean and tobacco microsomes (Kauss and Hassid, 1967; Goubet et al., 1998); Mn^{2+} was activating at 10-20 mM and inhibitory at >50 mM to tobacco solubilized enzymes (Goubet and Mohnen, 1999); while Mg^{2+} was inhibitory at 1 mM in tobacco solubilized (Goubet and Mohnen, 1999) and at 5 mM in soybean permeabilized enzymes (Ishikawa et al., 2000). The presence of KCl and other metal ions (Ca^{2+}, Ba^{2+}, Hg^{2+}, Cu^{2+}, Al^{3+}, Co^{2+}, Fe^{2+}, Fe^{3+}, Zn^{2+}, and Ni^{2+}) were found to be inhibitory (Kauss and Hassid, 1967; Goubet et al., 1998; Ishikawa et al., 2000). PMT appears to use SAM as the sole donor substrate, as the use of 5-*N*-methyltetrahydrofolate did not result in activity (Kauss and Hassid, 1967; Goubet et al., 1998). AMP, ADP, ATP, and most significantly *S*-

adenosyl-L-homocysteine (SAH), which is the by-product of the PMT activity, inhibit the activity of the soybean enzyme (Ishikawa et al., 2000).

PMT activities have been successfully solubilized from flax, tobacco, and soybean microsomes using Triton X-100 or CHAPS (Bruyant-Vannier et al., 1996; Goubet and Mohnen, 1999; Ishikawa et al., 2000). In contrast to activity in intact microsomes, the PMT activities of detergent-permeabilized or solubilized membrane preparations, as well as those of freeze-thawed microsomes, are not activated by addition of UDP-GalA, and are stimulated by the presence of exogenous acceptor substrates. Solubilized HG-MT activity has been reported when pectin (HG-enriched with various degree of methylesterification (DE)), PGA, and OGA were used as the exogenous acceptor substrates (Bourlard et al., 1997; Goubet and Mohnen, 1999; Ishikawa et al., 2000). In general, it appears that pectin with low DE (~10-30%) serves as a better acceptor than PGA or pectin with high DE, while PGA appears to be a better acceptor than OGA. In addition to HG, the GalA residues on RG-I and RG-II backbones can also be methylesterified, and PMT activities that act on RG-I and RG-II (RG-I-MT and RG-II-MT, respectively) have been described in flax solubilized enzymes (Bourlard et al., 1997).

Microsomal PMT activities have a neutral pH optimum in the absence of exogenous acceptors (Table 1.2). However, studies of flax freeze-thawed microsomes and solubilized enzymes identified two HG-MT isoforms, designated PMT5 and PMT7, respectively, with different pH optima and acceptor substrate preferences. PMT5 had a pH optimum of 5.5 in the presence of exogenous low-methylated pectin (LMP; DE = 10%), and was associated with both high (1.12-1.14) and low (1.08-1.11) density membranes upon sucrose gradient centrifugation (Bourlard et al., 1997). Less active than PMT5, PMT 7 had maximum activity at pH 7.0 in the presence of highly-methylated pectin (HMP; DE = 50%) and was found to reside only in the high

density membrane fraction (Bourlard et al., 1997). A three-step chromatography of the flax solubilized HG-MT activity over cross-linked pectate (CLPA) gel, α-SAH gel, and size exclusion columns, respectively, purified PMT5 and PMT7 to apparent homogeneity (Bourlard et al., 2001). PMT5 corresponds to a ~40 kD protein, as determined by a sucrose gradient centrifugation, with a pI of 5.8-6.5, while PMT7 was determined to be ~110 kD with a pI of 8.7-9.2 (Bourlard et al., 2001). Interestingly, independent size exclusion chromatography of PMT5 and PMT7 also identified a common, additional peak designated PMT18, which corresponded to an ~18 kD protein with an acidic pI (4.0-4.5). PMT18 appears to be the lone subunit of multimeric PMT5 and PMT7 based on two observations: 1) SDS-PAGE of PMT5 and PMT7 yielded only a single protein band of ~18 kD that was able to bind [^{14}C]SAM, and 2) re-subjection of the PMT18 peak collected from the size exclusion chromatography onto the same column resulted in detection of HG-MT activity peaks corresponding to those of PMT5, PMT7, and PMT18, indicating re-association of PMT18 molecules to form the higher molecular weight PMTs (Bourlard et al., 2001). These results provide evidence for protein complex involvement in the methylesterification of pectin. Unfortunately, neither the identity of the PMT18 polypeptide, nor those of PMT5 and PMT7, has been determined.

An Arabidopsis putative HG-MT and the QUA2/TSD2/OSU1 gene family

Analyses of the Arabidopsis putative HG-MT QUASIMODO 2 / TUMOROUS SHOOT DEVELOPMENT 2 / OVERSENSITIVE TO SUGAR 1 (QUA2/TSD2/OSU1) mutants identified the mutated gene as a putative PMT gene (Krupkova et al., 2007; Mouille et al., 2007; Gao et al., 2008). The ethyl methane sulfonate (EMS)-induced mutants *qua2-1* and *tsd2-1* display phenotypes similar to those observed in *qua1*, a mutant of the GAUT-gene family

QUA1/GAUT8, including short seedling hypocotyls when grown in the dark, dwarfed mature plants, and defects in cell:cell adhesion (Krupkova et al., 2007; Mouille et al., 2007). The *qua2-1* mutant also exhibits an abnormal root tip border-like cell phenotype that was also seen in more severity in *qua1* (see above) (Durand et al., 2009). Using the promoter:*GUS* expression technique, it was shown that the area of expression of the shoot meristem marker genes *KNAT1* and *KNAT2* were expanded in the *tsd2-1* mutant compared to the wild-type, which may have contributed to the disorganized tumor-like callus formation in the mutant seedlings *in vitro* (Krupkova et al., 2007). Analysis of *osu1-1*, another mutant of *QUA1/TSD2/OSU1*, indicated a role of the gene in the regulation of plant response to changes in carbon (C) and nitrogen (N) nutrient balance (Gao et al., 2008). *osu1-1* seedlings showed wild-type appearance when grown on media with balanced C/N ratios (low C/low N or high C/high N), but demonstrated anthocyanin accumulation and reduced root length under imbalanced C/N conditions (low C/high N or high C/low N) (Gao et al., 2008). Cell wall analysis of *qua2-1* indicated a 13% and 50% reduction in total GalA and HG content in the wall, respectively, but no changes in the RG-I content, compared to wild-type (Mouille et al., 2007). Interestingly, the HG remaining in the mutants *qua2-1* and *tsd2-1* was shown to have similar size distribution and degree of methylesterification as that of wild-type (Krupkova et al., 2007; Mouille et al., 2007; Ralet et al., 2008). Nevertheless, all the observed phenotypes are consistent with a defect in HG synthesis.

Positional cloning of the mutations and complementation studies identified Arabidopsis locus *At1g78240* as *QUA2/TSD2/OSU1* (Krupkova et al., 2007; Mouille et al., 2007; Gao et al., 2008). The gene encodes a 684 a.a. protein with a molecular mass of 77.9 kDa and a pI of 7.34 (Krupkova et al., 2007). The protein is predicted to have a type II transmembrane protein topology and to contain a putative methyltransferase domain (a.a. 155-676; pfam03141; Domain

of Unknown Function (DUF) 248), within which resides a SAM-dependent methyltransferase motif (a.a. 230-372) (Krupkova et al., 2007; Mouille et al., 2007). The mutations in *qua2-1*, *tsd2-1*, and *osu1-1* were determined to cause Arg627STOP, Trp210STOP, and Asn560Tyr codon changes, respectively, the later being a highly conserved residue within the *QUA2/TSD2/OSU1* gene family (see below) (Krupkova et al., 2007; Mouille et al., 2007; Gao et al., 2008). *QUA2/TSD2/OSU1* expression was observed ubiquitously in all tissues, based on microarray data, RT-PCR, and promoter:*GUS* analyses, and expression of a *QUA2::GFP* construct localized the encoded protein to the Golgi apparatus (Krupkova et al., 2007; Mouille et al., 2007; Gao et al., 2008), consistent with the reported location of HG synthesis and methylesterification (Zhang and Staehelin, 1992; Staehelin and Moore, 1995).

A BLAST search of QUA2/TSD2/OSU1 revealed a QUA2/TSD2/OSU1 gene family consisting of 29 members (Figure 1.4), all of which contain the putative methyltransferase domain (Krupkova et al., 2007; Mouille et al., 2007). It is intriguing that analysis of microarray databases identified at least three sets of co-expression pairing between members of the GAUT and the QUA2/TSD2/OSU1 gene families: 1) *QUA1/GAUT8* with *QUA2/TSD2/OSU1* and *At3g23300*; 2) *GAUT1/At3g61130* with *At4g14360* and *At1g26850*; and 3) *GAUT9/At3g02350* with *At3g23300* and *At1g26850* (Figure 1.4) (Mouille et al., 2007). It is possible that different combinations of GalATs and HG-MTs are required to synthesize diverse subpopulations of HG or pectin in general, in order to build the developmentally- and spatially-regulated complex pectin structures in the wall.

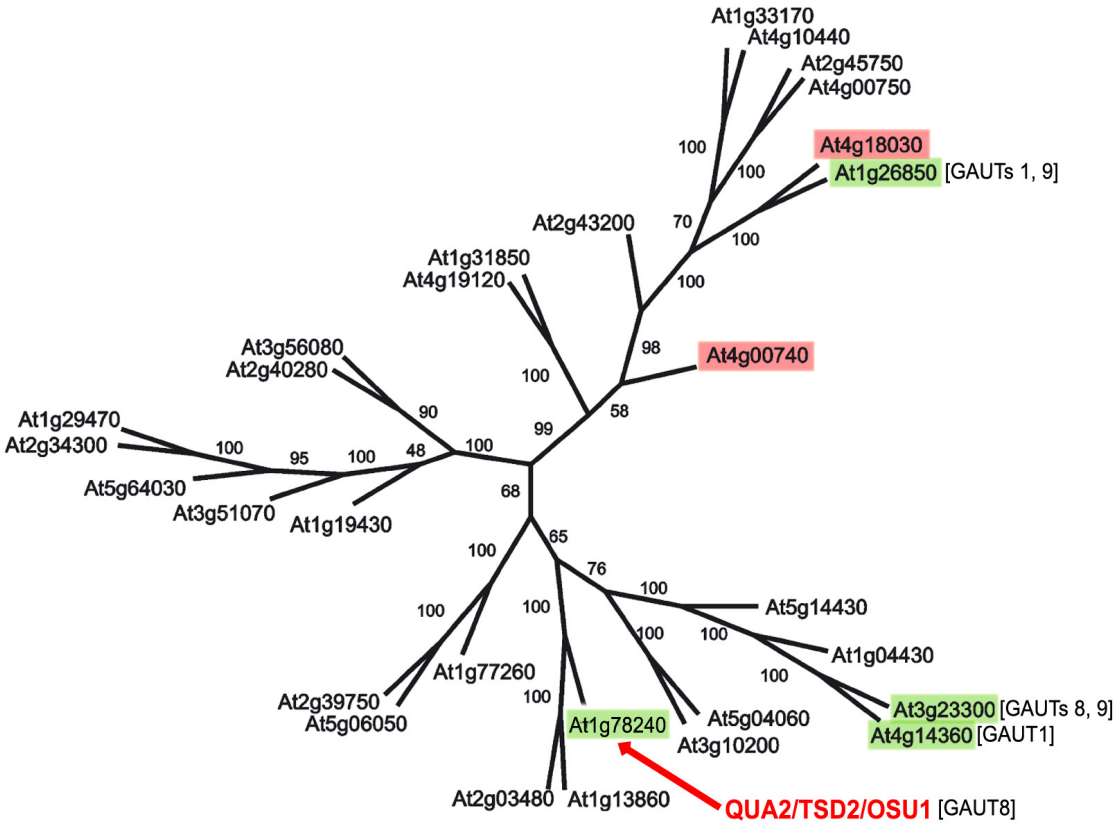


Figure 1.4. A phylogenetic tree of the QUA2/TSD2/OSU1 putative methyltransferase gene family with 29 members. Highlighted in green are the family members found to co-express, based on analysis of microarray databases, with the GAUT gene family members, which are denoted in brackets. Highlighted in pink are the putative methyltransferases identified by proteomics analysis as putative GAUT1:GAUT7 HG:GalAT complex-interacting proteins (see Chapter 3) (Adapted from Krupkova et al., 2007).

O-Acetylation of the HG backbone by acetyltransferases

O-acetylation of the hydroxyl groups at the C-2 and/or C-3 of GalA residues has been reported in both HG and RG-I, but not in RG-II (Whitcombe et al., 1995; Ishii, 1997; Perrone et al., 2002). The presence of acetyl groups in pectin and especially in HG affects intercellular adhesion (Liners et al., 1994), inhibits degradation of HG by endo-polygalacturonase (Renard and Jarvis, 1999), and hinders formation of both calcium-pectin and pectin-sucrose-acid gels (Pippen et al., 1950; Ralet et al., 2003).

To date, only one report has demonstrated the occurrence of pectin acetyltransferase activities (Pauly and Scheller, 2000). Incubation of a microsome preparation from potato suspension cultured cells with [^{14}C]acetyl-CoA resulted in incorporation of the [^{14}C]acetyl groups onto endogenous acceptors. Analysis of the ^{14}C-acetylated, ethanol-precipitable reaction products by various enzymatic digestions showed that ~23% of the incorporated radioactivity could be released by a EPG/PME treatment, and ~33% was released by a rhamnogalacturonan acetylesterase (RGAE) treatment. These data suggest that at least half of the total acetyltransferase activity transferred the [^{14}C]acetyl group onto endogenous pectic acceptors, and provide evidence for membrane-bound pectin HG- and RG-I-acetyltransferase (HG-AT and RG-AT, respectively) activities.

Further characterization of the RG-AT activity, and of the total microsomal acetyltransferase activity, indicated a pH optimum of 7.0 and temperature optimum of 30° for maximal [^{14}C]acetyl incorporation, with highest activities observed during the linear growth phase of the cell culture (Pauly and Scheller, 2000). The addition of divalent cations into the reaction mixture stimulated both RG-AT and total acetyltransferase activities by ~1.5-fold when >1 mM Mg^{2+} was used and by at least 10-fold in the presence of 5 mM Mn^{2+} (Pauly and

Scheller, 2000). These characteristics are very similar to those of HG:GalAT and, at least for the pH and temperature optima, of HG-MT. The apparent K_m and V_{max} for [^{14}C]acetyl-CoA of RG-AT and of total acetyltransferase activities were 35 μM and 0.9 pkat/mg protein, and 95 μM and 10.4 pkat/mg protein, respectively (Pauly and Scheller, 2000). The use of acetate, another potential donor substrate, did not result in acetyltransferase activity, and addition of exogenous substrate also had no effect, suggesting that the AT activities may be enclosed within vesicles. The molecular mass of the total *O*-acetylated products were estimated by size exclusion chromatography to be >500 kDa, which upon digestion with a rhamnogalacturonan lyase resulted in presumably RG-I-derived peaks of ~20 kDa and ~8 kDa. Based on these data, a further characterization of the HG-AT activity is warranted.

CONCLUSIONS AND RELEVANCE

GAUT1 is the first pectin biosynthetic enzyme for which the encoding gene has been identified and the enzymatic function verified. The identification of GAUT1 was a major step forward in our understanding of the biosynthesis of homogalacturonan in particular and of pectin in general, and provides biochemical and molecular tools for further research.

The precise role of GAUT1 in the synthesis of homogalacturonan still needs to be determined. It is of interest to determine the biochemical function of GAUT7, since GAUT1 and GAUT7 are related to each other, and since both were present in the same partially purified Arabidopsis solubilized enzyme preparation used in the proteomics approach employed to identify GAUT1. Indeed, it has been suggested that GAUT1 and GAUT7 may be part of a pectin biosynthetic complex (Sterling et al., 2006). It also remains unclear how the

polymerization of the homogalacturonan backbone is connected to its modification by methyl- and acetyltransferase activities.

The project described in this dissertation was aimed at further clarifying the role of GAUT1 and elucidating the function of GAUT7 in homogalacturonan synthesis. Significantly, the interaction between GAUT1 and GAUT7 in a pectin biosynthetic homogalacturonan:galacturonosyltransferase enzyme complex is demonstrated and characterized, and two putative methyltransferases have been identified as interacting with the GAUT1:GAUT7 complex.

CHAPTER 2

LITERATURE REVIEW:

EUKARYOTIC POLYSACCHARIDE BIOSYNTHESIS -

MECHANISM AND THE ROLE OF PROTEIN COMPLEXES

INTRODUCTION

Polysaccharides are a group of biomolecules indispensable for life that are collectively the most abundant organic biomass on earth. Polysaccharides serve as a major energy reserve for living organisms, and make up a large portion of the extracellular matrices that surround the plant and animal cells, playing important roles in cell-to-cell adhesion, interaction, and for protection. While nucleic acids and proteins are linear polymers, each with only one specific type of linkage connecting the building blocks, polysaccharides exist in both linear and branched forms, and the glycosyl units can be linked to each other in more than one way. The ability of carbohydrates to link in branched structures, in pentose and hexose ring forms, and in α and β configurations, contributes to the vast diversity and complexity of polysaccharide structures.

In contrast to nucleic acids and proteins, the biosynthesis of polysaccharides is template-independent. The plethora of polysaccharide structural architectures found in nature is thus largely due to the orchestrated actions of polysaccharide synthesizing, modifying, and degrading enzymes. With regards to the synthesizing enzymes, glycosyltransferases (GTs) comprise 1-2% of the genome of at least 500 species (Rini et al., 2008), and can be classified into more than 90 GT families based on protein structures (Cantarel et al., 2009). GTs catalyze the transfer of a

glycosyl residue from a glycosyl donor onto diverse acceptors, including oligosaccharides, peptides, and lipids. Research to elucidate the biochemical/catalytic functions of GTs has been progressing slowly, especially in comparison to that of the polysaccharide degrading enzymes. This, in most part, is because of the difficulty in working with the GTs that are mostly membrane-bound and in very low abundance, and due to the lack of availability of suitable assay methods, sugar donors, and acceptor substrates (Wagner and Pesnot, 2010).

In recent years, protein complexes have been increasingly implicated in polysaccharide syntheses. Protein associations, at least for syntheses that occur in the Golgi apparatus, have been suggested to function in maintaining Golgi localization and, more significantly, in enhancing synthesis efficiency by coupling together two or more enzymes that catalyze sequential reactions (de Graffenried and Bertozzi, 2004). Furthermore, many GTs have been shown to form gene families with multiple members. It is possible that different members of a gene family are co-expressed in the same cells or tissues, and form complexes with different subunit compositions to facilitate the production of the diverse array of polysaccharide structures, such as those present in plant cell walls. One example of this, as presented in this dissertation, are Arabidopsis GAUT1 and GAUT7, which are members of the GAUT gene family, are co-expressed in many cells/tissues, and interact with each other to form a pectin biosynthetic enzyme complex.

The goal of this literature review is to summarize the biosyntheses of several different polysaccharides, with an emphasis on the role of protein complexes in polysaccharide synthesis. The polysaccharides covered here, namely cellulose, starch, hyaluronan, and heparan sulfate, are selected to represent both plant and animal polymers, with biosyntheses that are located at the plasma membrane, in stroma of plastids, and in the Golgi apparatus. Also included is the N-

glycan biosynthesis of glycoproteins, but with a specific emphasis on the oligosaccharyltransferase activity, which is catalyzed by a large enzyme complex. The aim of this review is to explore several different mechanisms of polysaccharide synthesis, and to describe how the protein complexes contribute to the process. This topic is relevant to the research presented in this dissertation, which is the first demonstration of a pectin biosynthetic enzyme complex, and serves as a knowledge base to compare and interpret the research results regarding pectin biosynthesis, as well as to form hypotheses for future work.

CELLULOSE BIOSYNTHESIS IN HIGHER PLANTS

Cellulose structure and functions

Cellulose is the most abundant biopolymer in nature, with at least 10^{11} tons being synthesized on earth annually (Brown, 1996), mostly by vascular plants. Besides this major group of cellulose producers, a broad spectrum of organisms, including algae, the slime mold *Dictyostelium*, several bacterial species, and tunicates, also synthesize cellulose as an extracellular polysaccharide (Saxena and Brown, 2005). Nevertheless, it is only in plants, algae, and *Dictyostelium* that cellulose constitutes part of the cell wall (Saxena and Brown, 2005).

Cellulose is composed of glucose molecules in β-1,4-linkages that form linear polymer chains, in which each glucosyl residue is rotated 180° with respect to its neighbors (Figure 2.1). The resulting flat ribbons are held together by van der Waals forces and hydrogen bonds, forming crystalline microfibrils (Cousins and Brown, 1995). Native cellulose is grouped into cellulose I and II, in which glucan chains are arranged parallel and anti-parallel, respectively (Gardner and Blackwell, 1974; Kolpak and Blackwell, 1976). Two sub-allomorphs of cellulose I, designated Iα and Iβ, are different in their crystal packing, molecular conformation, and

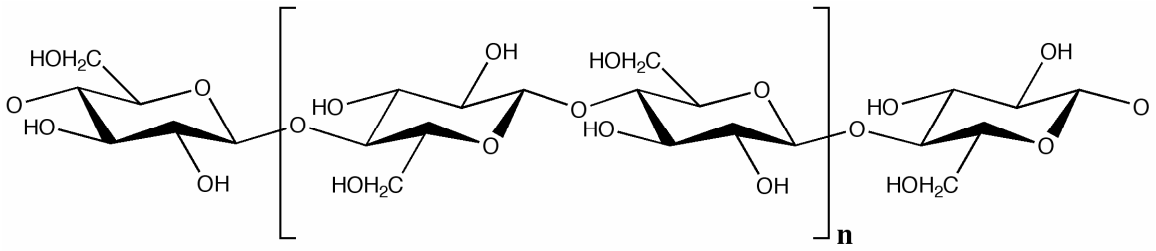


Figure 2.1. The basic structure of cellulose, with the repeating unit cellobiose in brackets.

hydrogen-bonding that may contribute to the difference in stability and properties of the cellulose fibers (Sugiyama et al., 1991; Nishiyama et al., 2002; Nishiyama et al., 2003). Cellulose Iα and Iβ are found in varying proportions depending on the source of the material (Brett, 2000). It was reported that in conifer tracheid cells, cellulose Iα and Iβ are predominant in the primary and secondary walls, respectively (Kataoka and Kondo, 1998). On the other hand, cellulose II is rare and only found naturally in a few algal and bacterial species (Brett, 2000; Brown, 2004), but can be derived synthetically from cellulose I through mercerization or regeneration processes (O'Sullivan, 1997). Cellulose II is more stable thermodynamically than cellulose I due to the anti-parallel chain packing that allows for additional hydrogen bonds in cellulose II (O'Sullivan, 1997; Saxena and Brown, 2005).

The length and width of the cellulose microfibrils vary greatly depending on the stage of cell wall development and the source. The degree of polymerization (DP) of the glucosyl residues in cellulose has been reported to range from ~500 to 8,000 in primary walls, and from ~14,000 to 15,000 in secondary walls (Brett, 2000; Brown, 2004). Microfibrils are approximately 5-10 nm in diameter and are believed to contain up to 90 glucan chains in most plants (Ha et al., 1998; Brett, 2000). It has been shown, at least in higher plants, that these microfibrils are made up of thinner crystalline fibrils ~2 nm in diameter, which are estimated to contain 10-15 glucan chains each (Ha et al., 1998).

In higher plants, cellulose comprises a major portion of the cell wall, ~20% of the primary wall and ~40% of the secondary wall (Harris and Stone, 2008). Cellulose functions primarily as a protective layer from the environment for protoplasts and as a source of strength that allows plants to grow upright. For example, in the primary wall of dicots and non-graminaceous monocots, cellulose microfibrils are hydrogen-bonded with the hemicellulose

xyloglucan, forming the load-bearing network that has been shown to contribute to ~70% of the tensile strength of the tissue (Shedletzky et al., 1992). Through the deposition site and pattern of microfibril arrangement in the wall, cellulose also plays a determining role as the skeleton in controlling cell growth and elongation and, in turn, organ formation and overall plant architecture (Brown, 2004). Utilization of cellulose is vital in human daily life, e.g. wood for housing and fibers for clothing. In the past several years, cellulosic material has also gained significant attention from both the research and industrial communities, as its natural abundance makes it a sustainable biomass source for biofuel production (Ragauskas et al., 2006).

Cellulose synthesis in higher plants by cellulose synthase complexes in rosette structures

Multiple lines of evidence demonstrate that the cellulose biosynthetic machinery is located at the plasma membrane, with the exception of some algae that synthesize cellulosic scales in the Golgi apparatus (Brown et al., 1970). Pulse-chase experiments using the radioactive nucleotide-sugar substrate UDP-glucose coupled with cell-fractionation demonstrated co-fractionation of a high percentage of radioactive β-glucan, presumably cellulose, with the wall fraction instead of with the membrane fraction, indicating that cellulose is not made within the cell (Bowles and Northcote, 1972). Freeze-fracture electron microscopy of the outer leaflet of the plasma membrane of the green algae *Oocystis* provided direct visualization of enzyme complexes associated with the ends of cellulose microfibrils (Brown and Montezinos, 1976). Similar enzyme complexes, named terminal complexes (TCs) were soon also observed in bacteria, other algae, higher plants, tunicates, and *Dictyostelium* (Brown et al., 1976; Willison and Brown, 1978; Mueller and Brown, 1980; Grimson et al., 1996; Kimura and Itoh, 1996).

TCs in higher plants are arranged in six-fold symmetry called 'rosettes', which are ~25 nm in diameter and composed of six 8 nm cellulose synthase complex (CSC) subunits (Figure 2.2C) (Mueller and Brown, 1980). Each TC produces cellulose 5-10 nm thick that consists of up to 90 glucan chains (Ha et al., 1998). Besides the plasma membrane, rosette TCs have also been observed in the Golgi apparatus and in Golgi-derived vesicles from mesophyll cells of *Zinnia elegans*, indicating that TCs may be assembled in the Golgi, or maybe even in the endoplasmic reticulum, and then transported to the plasma membrane (Haigler and Brown, 1986). The more dispersed distribution of, and absence of associating cellulose microfibrils with, the rosettes in intracellular compartments suggest that TCs remain inactive until they reach the plasma membrane (Haigler and Brown, 1986), possibly serving as a regulatory mechanism for cellulose synthesis.

The isolation of cellulose synthase activity has been very challenging due to the instability of the membrane-bound activity upon extraction from the plasma membrane and to the presence of callose (β-1,3-glucan) synthase activity in the same enzyme preparations (Guerriero et al., 2010). Cellulose synthase activity in detergent extracts has been demonstrated in cotton (*Gossypium hirsutum*) (Okuda et al., 1993; Kudlicka et al., 1995; Peng et al., 2002), *Lolium multiflorum* (Meikle et al., 1991), mung bean (Kudlicka and Brown, 1997), blackberry (*Rubus fruticosus*) (Lai-Kee-Him et al., 2002), hybrid aspen (Colombani et al., 2004; Bessueille et al., 2009), and tobacco (Cifuentes et al., 2010). However, it is only in mung bean that physical separation between cellulose and callose synthase activities has been achieved, proving that the activities are catalyzed by different enzymes. Interestingly, both cellulose and callose synthase activities were shown to reside in detergent-resistant plasma membrane microdomains (lipid

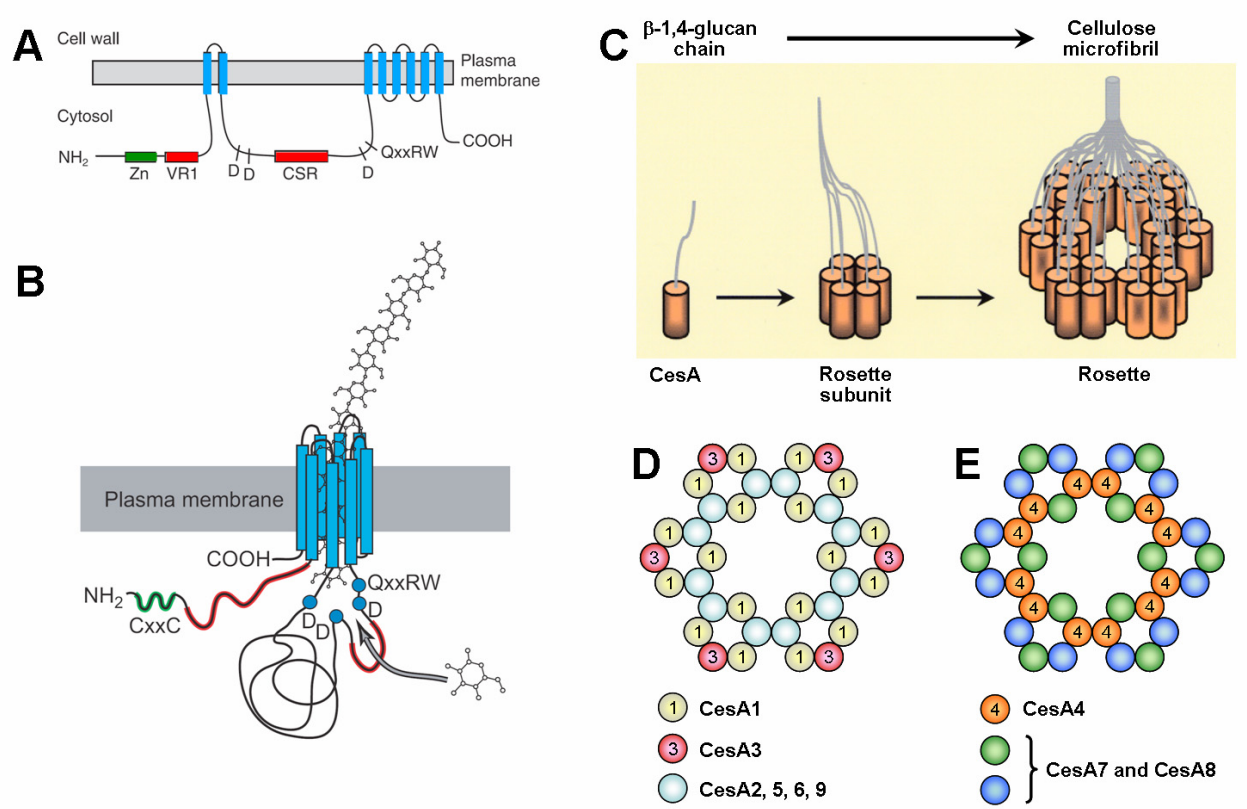


Figure 2.2. Cellulose synthases (CesAs) are the building blocks of the supramolecular cellulose biosynthetic machinery.

(A) The predicted topology of a CesA protein. Transmembrane spanning domains are colored blue. The zinc finger domain (Zn) is shown in green. The variable region (VR1) and Class-Specific region (CSR) are colored red. The approximate positions of D, D, D, and QxxRW motifs are indicated. (Adapted with modification from Taylor, 2008).

(B) A model of the three-dimensional structure of a CesA protein. During polymerization, the growing glucan chain is extruded through a pore formed by the eight transmembrane domains of the CesA protein. Regions are colored as decribed in (A). (Adapted from Richmond, 2000).

(C) A model of CesA assembly to form the rosette structure. Six CesA proteins, each of them synthesizing a glucan chain, form a cellulose synthase complex (CSC), and six CSCs further

assemble into the rosette structure that produces a cellulose microfibril. (Adapted from Doblin et al., 2002).

(D) Schematic of the proposed organization of primary wall CesAs into a rosette structure. The numbering is that of Arabidopsis CesAs. (Adapted with modification from Mutwil et al., 2008).

(E) Schematic of the proposed organization of secondary wall CesAs into a rosette structure. The numbering for CesAs in Arabidopsis is shown. (Adapted with modification from Timmers et al., 2009).

rafts) from hybrid aspen, which may explain the difficulty in separating the two enzyme activities (Bessueille et al., 2009).

The available literature indicates the need to empirically determine the conditions for isolation and assay of the cellulose synthase activity. The choice of detergents for the successful solubilization of active CSCs appears to be species-specific, probably due to the different lipid compositions surrounding CSCs from different plant species (Cifuentes et al., 2010). For example, digitonin was the most efficient detergent to extract active CSCs from cotton fibers, hybrid aspen, and tobacco BY-2 cells, while blackberry CSCs were extracted most efficiently using taurocholate or Brij58. The cation requirement also varies for the CSC extracts, even from the same species, that are prepared using different detergents, suggesting a possible isoform-specific cation preference (Cifuentes et al., 2010). The highest cellulose synthase activity *in vitro* was observed in the absence of cations for taurocholate extracts and in the presence of 8 mM Mg^{2+} for Brij58 extracts from blackberry (Lai-Kee-Him et al., 2002). Other preparations require different mixtures of cations for optimal cellulose synthesis *in vitro*, as shown with digitonin extracts from hybrid aspen (1 mM Ca^{2+}, 8 mM Mg^{2+}) and tobacco (8 mM Ca^{2+}, 8 mM Mg^{2+}) (Colombani et al., 2004; Cifuentes et al., 2010). With these digitonin extracts, which were obtained from suspension cultured cells harvested at stationery phase, the β-glucan polysaccharide products were shown to contain up to >50% cellulose with the rest being callose.

Cellulose synthase is an inverting enzyme that catalyzes the polymerization of β-1,4-glucan using UDP-Glc as the donor substrate. Addition of the new Glc residues occurs at the non-reducing end of the growing chain, as observed in *Acetobacter aceti* (Koyama et al., 1997) and blackberry (Lai-Kee-Him et al., 2002). It is still not clear if cellulose synthesis requires a primer. Cellobiose is typically added to the cellulose synthase reaction, however it is thought to

be an effector of the activity instead of a primer (Guerriero et al., 2010). It was proposed that sitosterol-β-glucoside acts as a primer for β-1,4-glucan chain elongation by cellulose synthase, since incubation of cotton fiber membranes or heterologously-expressed cDNA of a cotton cellulose synthase catalytic subunit (GhCesA1, see next section on CesA gene family) in the presence of UDP-Glc and this glycolipid, resulted in the production of sitosterol-cellodextrins (Peng et al., 2002). It was argued, however, that cellulose synthase activity may only use sitosterol-β-glucoside as an acceptor in *in vitro* reactions (Somerville, 2006; Guerriero et al., 2010). The potential function of the sitosterol-β-glucoside as a primer for cellulose biosynthesis still needs to be demonstrated *in vivo*.

Transmission electron microscopy of *in vitro* synthesized cellulose following treatment with Updegraff reagent (to remove callose) revealed microfibrils of 2-3 nm in diameter and 0.2-5 μm in length, that are clustered into small bundles of 3-10 units each (Lai-Kee-Him et al., 2002; Cifuentes et al., 2010). Comparable structures were observed with Updegraff reagent-treated, endogenous cellulose isolated from blackberry cell walls (Lai-Kee-Him et al., 2002). It is noteworthy, however, that the Updegraff reagent-treated *in vitro* cellulose appears to be much longer (degree of polymerization (DP) >100, up to 5 μm in length) than the Updegraff reagent-treated endogenous cellulose (DP ~20, 0.2-0.3 μm in length), and has a much higher crystallinity and larger crystal size than endogenous cellulose (Lai-Kee-Him et al., 2002). It was suggested that *in vitro* cellulose synthesis may occur in a less-constraining environment in the absence of already-laid-down wall, which allows the synthesis of microfibrils with fewer defects than the *in vivo* counterparts.

Higher plant cellulose synthase genes

Higher plant cellulose synthase genes were first identified through the screening of a cDNA library of cotton fiber mRNA harvested from cotton bolls at the onset of secondary cell wall cellulose synthesis (Pear et al., 1996). The mRNAs of two cotton genes, *GhCesA1* and *GhCesA2,* are up-regulated during the time of secondary cellulose synthesis in developing cotton fibers. The encoded proteins show homology to the previously described bacterial cellulose synthase catalytic subunits encoded by the *CelA* genes in *Acetobacter xylinum*, *Agrobacterium tumefaciens*, and *E.coli*, including subdomains that are conserved and predicted to be critical for UDP-Glc binding and catalysis (Pear et al., 1996). Heterologous expression of a portion of the *GhCesA1* gene resulted in a polypeptide containing the conserved subdomains that was able to bind UDP-Glc. Soon after, Arioli and co-workers identified the Arabidopsis cellulose synthase gene *AtCesA1*, which encodes a protein with homology to those of *GhCesAs*, through analysis of an Arabidopsis cellulose synthesis mutant *cesa1/rsw1* (Arioli et al., 1998). Under non-permissive temperature, the mutant plants showed a significantly reduced cellulose content, disassembly of the rosette structures into particle subunits, and an accumulation of non-crystalline β-1,4-glucan. The wild-type phenotype was restored upon complementation with the cloned *AtCesA1* gene. These results, together with successful immunolabeling of the rosette structure in azuki bean (*Vigna angularis*) using polyclonal antibodies raised against the catalytic domain of cotton cellulose synthase (Kimura et al., 1999), demonstrated that the proteins encoded by these cellulose synthase genes were essential components of the rosettes.

The availability of genome sequences reveals multiple-member *CesA* gene families in higher plants, consisting of 10 *AtCesA* genes in Arabidopsis, and of gene families of at least 12, 8, 10, and 17 genes in maize, barley, rice, and poplar, respectively (Richmond, 2000; Tanaka et

al., 2003; Appenzeller et al., 2004; Burton et al., 2004; Kumar et al., 2009). Plant CesAs are membrane-bound proteins of 985-1088 a.a. in length, and contain eight transmembrane domains (TMDs), of which two are located near the N-terminus and the other six are clustered near the C-terminus (Figure 2.2A and B) (Richmond, 2000). The cytoplasmic N-terminus contains a cysteine-rich domain with a RING-type zinc finger motif that has been implicated to mediate protein-protein interactions in various protein complexes (Saurin et al., 1996), followed by a hypervariable region ~150 a.a. in length (Richmond, 2000). The N-terminal domain of GhCesA1 and GhCesA2 can form homo- or heterodimers, as shown by the yeast two-hybrid system and pull-down assays (Kurek et al., 2002). Further investigations indicated that the GhCesA1 N-terminal zinc-finger domain (and also the full length protein) dimerizes via intermolecular disulfide bonds under oxidative conditions and, in this dimerized form, is resistant to *in vitro* degradation by cysteine protease activity (Kurek et al., 2002; Jacob-Wilk et al., 2006). In contrast, under reducing conditions the domain binds zinc, which is presumably facilitated by interaction with metallothionein (GhMet-1), and does not dimerize, which leads to susceptibility to proteolytic degradation. It was proposed that this redox state-dependent dimerization might contribute to rosette assembly and function, as well as to its degradation and turnover (Kurek et al., 2002; Jacob-Wilk et al., 2006).

The large, central domain of the CesA proteins is ~530 a.a. flanked by the two clusters of TMDs, and thought to contain the catalytic site (Figure 2.2A and B) (Delmer, 1999). The sequence is highly conserved among all CesA proteins, except for a region of ~64-91 residues named Class-Specific Region that is hypervariable among CesAs from the same species but relatively conserved between CESAs from different species (Vergara and Carpita, 2001). It also contains the D,D,D,QXXRW motif, a signature of processive β-glycosyltransferases, of which

the first two Asp residues have been shown by site-directed mutagenesis to be essential for enzyme activity (Saxena and Brown, 1997). The central globular domain is thought to be cytoplasmic (Delmer, 1999), which is in keeping with the visualization of rosettes in plasma membrane patches of tobacco BY-2 protoplasts. In this study, the cytoplasmic domains of the rosettes were seen as hexagonal structures 40-45 nm in diameter and 30-35 nm in height, i.e. about twice as wide as the size of the membrane bound portion of rosette TCs observed previously using the freeze fracture method (Bowling and Brown, 2008). The hypervariable regions in both the N-terminal and globular domains have been shown to be phosphorylated in many CesA proteins, which has been suggested to play a role in targeting of the CesAs to proteosomal degradation and in interaction with microtubules (Nuhse et al., 2004; Taylor, 2007; Chen et al., 2010).

Organization of the plant cellulose synthases into CSCs, the subunits of rosettes

Distinct sets of CesA proteins are apparently required for cellulose synthesis in primary and secondary cell walls. Analyses of Arabidopsis plants mutated in *AtCesA* genes and those expressing antisense *AtCesA* constructs, pointed to AtCesA1/RSW1, AtCesA3/IXR1, and a group of AtCesA6-related proteins as specific for the cellulose synthesis in the primary cell walls (Arioli et al., 1998; Burn et al., 2002; Desprez et al., 2007; Persson et al., 2007). Diverse phenotypes observed in non-lethal mutant alleles of the primary wall *AtCesAs* include reduced cellulose content, defects in cell expansion rather than division resulting in radial swelling, altered cell shapes, shorter seedling hypocotyls and roots, reduced fertility, ectopic lignin deposition, and resistance to the cellulose synthesis inhibitor isoxaben (Arioli et al., 1998; Fagard et al., 2000; Scheible et al., 2001; Burn et al., 2002; Cano-Delgado et al., 2003; Wang et al.,

2006; Chu et al., 2007). Some of these effects are only observed under non-permissive conditions, e.g. higher growing temperature or in the absence of light. In contrast, Arabidopsis plants mutated in the secondary cell wall-specific AtCesAs (AtCesA4/IRX5, AtCesA7/IRX3, and AtCesA8/IRX1) display significantly reduced cellulose content in stem tissues (~20-30% of that in wild-type) and a characteristic collapsed xylem cell phenotype referred to as *irregular xylem* (*irx*) (Turner and Somerville, 1997; Taylor et al., 1999; Taylor et al., 2000; Taylor et al., 2003). Mutations in the rice homologs of the secondary cell wall CesAs are also dramatically reduced in cellulose content (~10-25% of wild-type level), have thinner cortical fiber cell walls, and exhibit brittle culm phenotype (Tanaka et al., 2003).

In support of the above mutant data suggesting two groups of CesAs, i.e. the primary wall CesAs (CesAs 1, 3, and 6 or 2 or 5 or 9) and the secondary wall CesAs (CesAs 4, 7, and 8), two independent analyses of publicly available microarray data revealed that the primary and secondary wall sets of *AtCesA* genes are co-expressed in tissues undergoing the respective developmental stages of growth (Brown et al., 2005; Persson et al., 2005). Similar patterns of transcript co-expression were also observed with *CesA* homologs in maize and barley (Dhugga, 2001; Burton et al., 2004), suggesting that the specialization of CesA isoforms is conserved among higher plants.

It has been established that the primary wall CSC is composed of three distinct AtCesAs subunits, i.e. AtCesA1, AtCesA3, and AtCesA6 (or AtCesA6-related proteins, which are AtCesA2, 5, and 9). Each of the three AtCesA subunits is essential and non-redundant, as exemplified by the lethality of Arabidopsis null mutants *cesa1* and *cesa3*, as well as double and triple mutants *cesa5/cesa6*, *cesa2/cesa6/cesa9, cesa2/cesa5/cesa6* (Desprez et al., 2007; Persson et al., 2007). Wang and co-workers attempted to rescue the mutant phenotype of *cesa1*^{rsw1} and

cesa3^{rsw5} using two recombinant proteins made by N-terminal domain swapping between AtCesA1 and AtCesA3 (Wang et al., 2006). However, one of the chimeric proteins could only partially restore *cesa1*^{rsw1} phenotype while the other showed a dominant negative effect on *cesa3*^{rsw5}, confirming that AtCesA1 and AtCesA3 are each indispensable. Promoter:*GUS* analyses verified the transcript co-expression between *AtCesA1* and *AtCesA3* (Scheible et al., 2001), and the partial redundancy among *AtCesA2*, *5*, *6*, and *9* (Desprez et al., 2007; Persson et al., 2007). Furthermore, AtCesA1, 3, and 6 co-immunoprecipitate, and all three proteins can form homo- and heterodimers with each other *in vivo*, as shown by bimolecular fluorescence complementation (BiFC) (Hu et al., 2002) analysis (Desprez et al., 2007; Wang et al., 2008). Blue native polyacrylamide gel electrophoresis (BN-PAGE) of Triton X-100 solubilized microsomes indicated that AtCesA1, 3, and 6 are parts of an ~840 kDa protein complex, which was suggested to correspond to the CSC subunit of rosette TCs and to possibly contain six CesA proteins (Wang et al., 2008). In contrast, when using extract from dark-grown seedlings of the premature stop codon *cesa6*$^{prc1\text{-}19}$ mutant line, BN-PAGE showed only protein bands corresponding to CesA monomers instead of the ~840 kDa protein complex, and no co-immunoprecipitation of CesA proteins was observed. Based on all these data, a model for the organization of CesAs in primary wall CSCs has been proposed (Figure 2.2D) (Mutwil et al., 2008).

The secondary wall CSC appears to have an organization similar to that of the primary wall CSC. The substantially reduced cellulose content in mutants of AtCesA4, 7, and 8 indicates that the three proteins are not redundant (Taylor et al., 2003). AtCesA4, 7, and 8 proteins are co-expressed in the exact same xylem and interfascicular region cells at the same time, as judged by tissue printing probed with respective polyclonal antibodies for each protein. More specifically,

by immunostaining Arabidopsis seedlings, the three proteins were shown to localize to the plasma membrane of mature xylem vessels with a characteristic banded pattern (Gardiner et al., 2003). This plasma membrane localization appears to be associated with and dependent on the presence of microtubules, and was not observed in *cesa4*$^{irx5-1}$ and *cesa7*$^{irx3-1}$ mutants, in which the respective proteins were undetectable due to premature stop codons. AtCesA4, 7, and 8 proteins can be co-precipitated/co-immunoprecipitated from solubilized stem extracts of a transgenic line expressing His-tagged-AtCesA7 (Taylor et al., 2003). However, no co-immunoprecipitation was observed with extract from *cesa4*$^{irx5-1}$. Furthermore, BN-PAGE analysis of detergent-solubilized stem wild-type microsomes showed that AtCesA4, 7, and 8 can be detected not only as monomers of ~120 kDa, but also as protein complexes of approximately 220, 440, and 720 kDa, which are likely to correspond to dimeric, tetrameric, and hexameric complexes of CesA proteins (Atanassov et al., 2009). The premature stop codon mutants, however, yielded only the ~120 and 240 kDa protein bands of AtCesA7 and 8, and of AtCesA4 and 8, from analyses of similar protein samples from *cesa4*$^{irx5-2}$ and *cesa7*$^{irx3-1}$, respectively, corresponding to the monomers and dimers of the CesA proteins. Two-step purification of solubilized extracts from a line expressing both His- and STREP-tagged AtCesA7, yielded protein samples containing relatively pure CSCs, which were subsequently subjected to LC-MS/MS analyses. The results indicated that the secondary wall CSCs contain only AtCesA4, 7, and 8. Investigations of the effects of the reducing agent DTT, salt (NaCl), and the detergent Sarkosyl on the integrity of the CSCs showed that both disulfide bonds and non-covalent intermolecular interaction play roles in forming the secondary wall CESA complex. Timmers and co-workers used a membrane-based yeast two hybrid system (Fetchko and Stagljar, 2004) and BiFC to demonstrate that AtCesA4, 7, and 8 can form heterodimers with each other

(Timmers et al., 2009). However, homodimer formation was only seen convincingly with AtCesA4, and weakly with AtCesA8, leading to a proposed model of how the CesAs are organized in secondary wall CSCs (Figure 2.2E).

The above data on the primary and secondary wall CSCs clearly demonstrate that all three CesAs are necessary for the formation of CSCs, for normal cellulose synthesis, and, at least for the secondary wall CSCs, for proper localization at the plasma membrane. Nevertheless, the catalytic function of the CesA proteins themselves has not been definitely demonstrated to date.

What is the mechanism of cellulose synthesis?

It is still unclear exactly how the cellulose synthesis reaction takes place. With regards to the catalytic mechanism, cellobiose, instead of glucose, is sometimes considered as the repeating unit of the polymer, since glucosyl residues are rotated 180° to each other in the cellulose chains. A disaccharide synthesis theory has thus been proposed, in which two glucosyl residues are proposed to be added to the growing chain in one reaction (Albersheim et al., 1997; Carpita and Vergara, 1998). This event could be catalyzed by a single enzyme with two active sites, with a possible covalent attachment of glucosyl units to Ser, Thr, or Asp at the active sites, or alternatively, by two enzymes facing each other on opposite sides of the polymer chain. Structural data on the CesAs will be necessary to test these hypotheses.

Cellulose synthesis is thought to occur at the cytoplasmic side of the CSCs, based on the membrane topology of the CesAs. Consequently, the cellulose chains must be transported out of the cell and into the wall, which is suggested to occur by extrusion of each glucan chain via a pore formed by the eight TMDs of CesAs (Richmond, 2000; Doblin et al., 2002). Six hexameric CSCs are believed to form a rosette structure (Figure 2.2C), from which 36 extruded glucan

chains, each from one CesA subunit, coalesce and instantaneously form the crystalline cellulose microfibril through hydrogen bonding. By using fluorescent proteins fused to the CesA proteins, fluorescently labeled primary and secondary CSCs have been observed moving at the plasma membranes guided through association with microtubules (Gardiner et al., 2003; Paredez et al., 2006; Wightman and Turner, 2008; Wightman and Turner, 2010). These CSC movements are presumably enforced by the space limitation between the plasma membranes and the previously laid-down walls. The cellulose polymerization and crystallization have been suggested as the driving forces propelling the polymerizing CSCs through the membranes, leaving behind the cellulose microfibrils in their paths (Diotallevi and Mulder, 2007; Lindeboom et al., 2008). In addition, fluorescently labeled CesAs have also been observed associated with intracellular, doughnut-shaped vesicles, which probably correspond to Golgi vesicles (Crowell et al., 2009; Wightman et al., 2009; Wightman and Turner, 2010). These CesA-containing vesicles move rapidly underneath the plasma membrane and pause at sites where cellulose deposition takes place, apparently guided by actin filaments, suggesting that they may serve to deliver CSCs to the plasma membrane.

Roles of other proteins implicated in cellulose synthesis

To date, CesAs are the only proteins that have definitively been shown as components of the CSC/rosettes. However, a growing number of other proteins are also thought to play roles in the biosynthesis of cellulose microfibrils, including sucrose synthase (SuSy), KORRIGAN 1, COBRA, KOBITO, and Cellulose synthase-interactive protein 1 (CSl1) (Taylor, 2008; Gu et al., 2010). For most of these proteins, association with cellulose synthesis is deduced from studies of mutant plants exhibiting defects in cellulose content, and none of them has their biochemical

function demonstrated. A few of these proteins are discussed briefly below in regards to their putative roles in cellulose synthesis.

Sucrose synthase (SuSy; E.C. 2.4.1.13) catalyzes the conversion of sucrose to UDP-Glc and fructose (Amor et al., 1995). A cotton membrane-bound SuSy has been identified that binds UDP-Glc and localizes in strips at the plasma membrane, parallel to secondary wall microfibrils. SuSy is thought to facilitate cellulose synthesis by channeling UDP-Glc to the CSCs, supported by the observation that incubation of permeabilized cotton fibers with [^{14}C]sucrose results in increased synthesis of cellulose and callose, and that sucrose synthase overexpression in poplar and tobacco results in higher biomass yields (as reviewed in Guerriero et al., 2010).

KORRIGAN 1 (KOR1) is a type II transmembrane protein involved in both primary and secondary wall cellulose synthesis (Nicol et al., 1998; Master et al., 2004; Szyjanowicz et al., 2004). It is probably the most extensively studied cellulose synthesis-related, non-CesA protein, with at least six mutant alleles identified in Arabidopsis (Nicol et al., 1998; Zuo et al., 2000; His et al., 2001; Lane et al., 2001; Sato et al., 2001; Szyjanowicz et al., 2004). Mutant phenotypes include a cell elongation defect, dwarf stature, reduced cellulose content, increase pectin content, and, in more severe mutants, a collapsed xylem phenotype, callus formation in 2-3 week-old plants, deformed cell shape, and incomplete cell walls. KOR1 orthologs from *Brassica napus* and hybrid aspen have been heterologously expressed in soluble forms in *Pichia pastoris*, and the recombinant proteins exhibited endo-β-glucanase activity on soluble cellulosic substrates, but not on small (< 5 glycosyl units) cellooligosaccharides, xylan, xyloglucan, or mixed-linked glucan (Molhoj et al., 2001; Master et al., 2004). KOR1 has been shown to localize to the cell plate during cytokinesis, and to plasma membrane and Golgi (Zuo et al., 2000; Robert et al., 2005). However, KOR1 does not seem to directly associate with primary or secondary wall

CesAs, since it is not co-immunoprecipitated with the CesA proteins (Szyjanowicz et al., 2004; Desprez et al., 2007). The function of KOR1 in cellulose synthesis is unclear, although it has been suggested to cleave glucan chains from lipid-linked primers during polymerization of cellulose, or to release the glucan chain so as to release tension or terminate synthesis (Peng et al., 2002; Somerville, 2006; Taylor, 2008). KOR1 requires N-glycosylation to be catalytically active *in vitro*, which may explain the cellulose deficient phenotypes displayed by mutants defective in several genes involved in the N-glycosylation pathway (Molhoj et al., 2001; Somerville, 2006).

COBRA (COB) (Schindelman et al., 2001; Roudier et al., 2005) and KOBITO (KOB) (Pagant et al., 2002) appear to function in ensuring a proper orientation of cellulose microfibrils. Mutants of both proteins have a disorganized orientation of cellulose microfibrils, which is further manifested by cell swelling, severely decreased cellulose content, and dwarfism. COB is a glycosyl-phosphatidylinositol (GPI)-anchored protein that is located in the cell wall and organized in transverse strips parallel to cortical microtubules. In contrast, KOB is an unknown protein with a broad localization pattern, including location in the cell wall, plasma membrane, and intracellular compartments (Pagant et al., 2002; Lertpiriyapong and Sung, 2003). The exact role of these proteins in cellulose synthesis remains to be determined.

Cellulose synthase-interactive protein 1 (CSl1) is the first Arabidopsis protein shown to associate with the primary wall CSCs, based on yeast two-hybrid assays that indicate positive protein-protein interactions between CSl1 and AtCesA1, 3, and 6 (Gu et al., 2010). This protein contains multiple tandem armadillo repeats and a C2 domain at the C-terminus, both suggested to facilitate protein-protein interactions. CSl1 co-expresses and co-localizes with primary wall CesAs at the plasma membrane. Similar to fluorescently labeled primary wall CesAs,

recombinant CSl1 fused to red fluorescent protein (RFP) also moves bidirectionally on the plasma membrane. Observed mutant phenotypes include dwarfism, significantly reduced amounts of crystalline cellulose, pollen defects, and an ~50% decrease in the velocity of AtCesA3 particle at the plasma membrane, which are similar to the phenotype observed in mutants of AtCesA1 and 3. The function of CSl1 is still unclear although, as with COB and KOB, it is suggested to be involved in organization of cellulose microfibrils.

STARCH BIOSYNTHESIS IN HIGHER PLANTS

Starch structure and functions

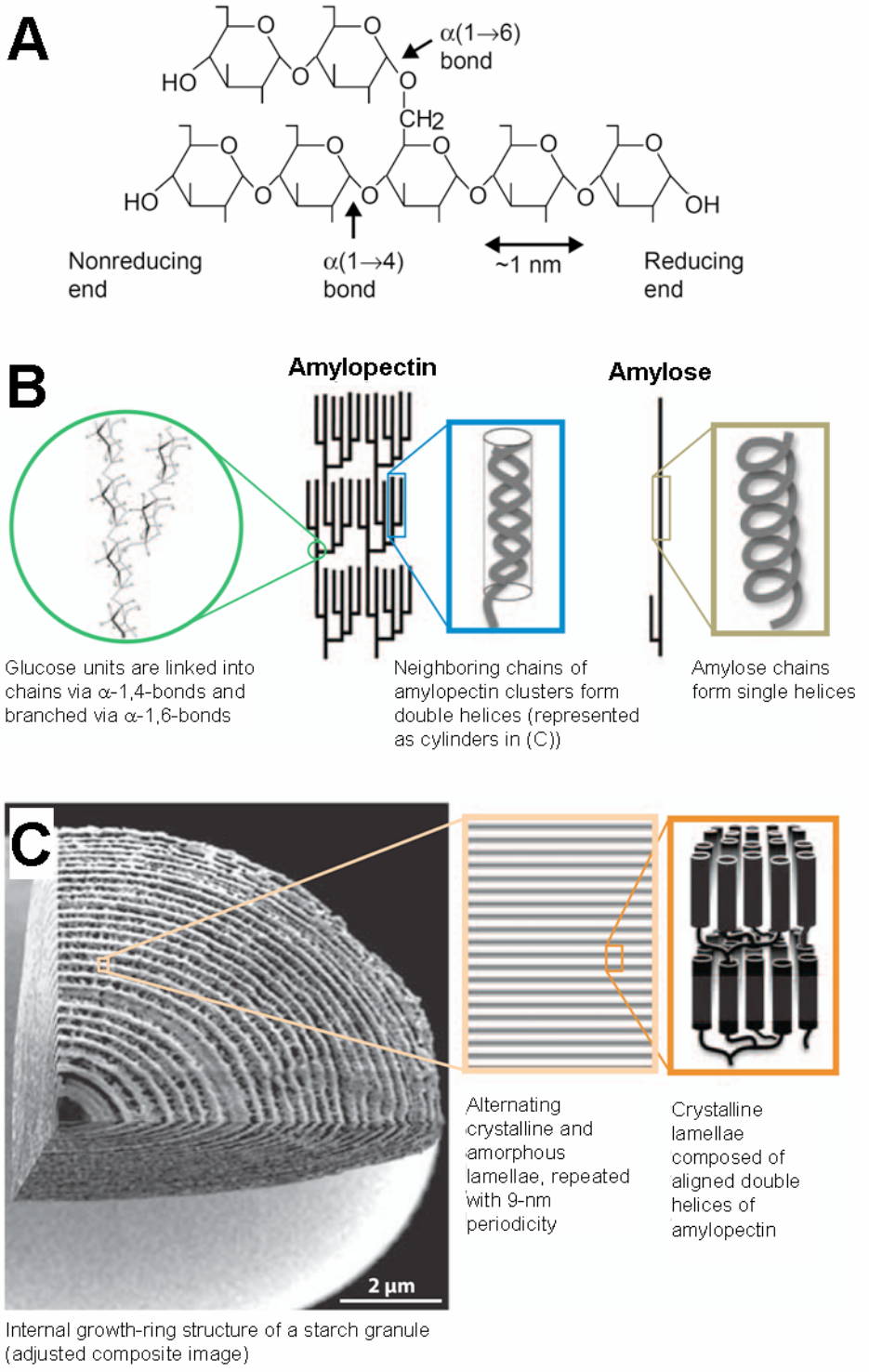
Starch is the major storage polysaccharide produced by photosynthetic eukaryotes such as green algae and land plants, and has been called the second most abundant carbohydrate in nature (Blennow et al., 2002; Ball and Morell, 2003). Starch is a homopolymer of α-D-glucose (Glc) that is made up of two different polysaccharides, amylose and amylopectin, which make up ~20-30% and ~70-80% of starch, respectively (Jobling, 2004). Amylose is a linear polymer of α-1,4-linked glucose residues (Figure 2.3) with DP up to >6,000 in potato tubers and cassava roots (Hizukuri and Takagi, 1984). In contrast, amylopectin is a branched polymer consisting of linear α-1,4-linked glucose chains interspersed every 24-30 Glc residues by α-1,6-linkages (~5% of the total Glc units) (Hizukuri, 1986; Jobling, 2004). The α-1,6-linkages establish branch points at which the short α-1,4-linked Glc chains, each approximately 12-30 residues, are clustered at regular intervals (Figure 2.3A and B) (Hizukuri, 1986; Smith, 2001). The Glc residues of amylopectin, especially in starches obtained from tubers, can be phosphorylated at the C-3 and C-6 positions, the former being a modification unique to starch (Blennow et al., 2002).

Figure 2.3. The composition and structure of starch granules.

(A) The molecular structure of amylopectin showing glucosyl residues connected by α-1,4- and α-1,6-linkages. Amylose, on the other hand, contains only α-1,4-linkages. (Adapted from James et al., 2003).

(B) Organization and secondary structure of glucan chains in amylopectin and amylose. (Adapted from Zeeman et al., 2010).

(C) A starch granule is composed of alternating crystalline and amorphous layers of amylopectin, which correspond to the regions containing clusters and branch points, respectively. (Adapted from Zeeman et al., 2010).


A
α(1→6) bond
HO
O
CH2
OH
Nonreducing end
α(1→4) bond
~1 nm
Reducing end
B
Amylopectin
Amylose
Glucose units are linked into chains via α-1,4-bonds and branched via α-1,6-bonds
Neighboring chains of amylopectin clusters form double helices (represented as cylinders in (C))
Amylose chains form single helices
C
2 µm
Alternating crystalline and amorphous lamellae, repeated with 9-nm periodicity
Crystalline lamellae composed of aligned double helices of amylopectin
Internal growth-ring structure of a starch granule (adjusted composite image)

The clustered chain arrangement of amylopectin allows for efficient packing of starch into semi-crystalline granules (Buleon et al., 1998; Smith, 2001; Jobling, 2004). The adjacent glucan chains within a cluster form left-handed double helices that assemble parallel to each other. This results in the alternating organization of crystalline and amorphous lamellae, corresponding to regions of the clusters and of the branch points, respectively, which form a concentric, semi-crystalline amylopectin layer in the granule (Figure 2.3C). The semi-crystalline layers alternate with amorphous layers, which contain less organized amylopectin, together forming the so-called "growth rings" that build up the granules. Amylose is thought to reside in the amorphous layers, where it has been suggested to disrupt the amylopectin organization (Jenkins and Donald, 1995).

Higher plants synthesize starch in plastids and amyloplasts (Smith, 2001; Tetlow et al., 2004). Plastids in the leaves synthesize transient starch during the day, which is broken down during the night for use as a carbon supply for non-photosynthetic metabolism, while amyloplasts in storage organs, such as tubers and seeds, produce starch as an energy reserve to be used during periods of dormancy and regrowth. Different plant species synthesize starch granules with considerable variations in composition (e.g. amylose:amylopectin ratios; phophate, protein, and lipid contents), size (from 0.1 to 100 μm in diameter), and shape (e.g. round, oval, or irregular) (Ramesh and Tharanathan, 2003; Jobling, 2004; Orzechowski, 2008). These characteristics are determined by the glucan chain length distribution and the degree of branching, and affect the properties of the starch obtained from different sources.

Besides being an important energy source for animals and humans, starch also has multiple industrial applications. Starch is insoluble in cold water, but in warm water the helices in amylopectin melt, resulting in the swelling of the starch granules and an increase in viscosity

of the solution (Jobling, 2004). With further heating and stirring, the granule disintegrates and the starch solubilizes. Cooling the solution allows the linear glucan chains to re-associate into aggregates that precipitate, and the starch solution sets into a gel. Starch with no amylose, such as that of maize *waxy* mutants (Shure et al., 1983), solubilizes into clear pastes that do not set, a characteristic that is enhanced by high phosphate content (Jobling, 2004). This characteristic makes it ideal as a stabilizer, thickener, and emulsifier. In contrast, high-amylose starch forms robust gels, and therefore has good coating, film-forming, and bonding properties (Ramesh and Tharanathan, 2003; Jobling, 2004). Starch and its derivatives are widely used in the food, textile, adhesive, board, paper, and building industries, and serves as a renewable source for bioethanol production (Ramesh and Tharanathan, 2003; Jobling, 2004).

Starch biosynthesis in plants: the key players

The highly organized structure of starch granules results from the concerted actions of a set of enzymatic activities summarized in Table 2.1. These starch biosynthetic enzymes appear to be highly conserved between chloroplasts and amyloplasts, and many of them exist in multiple isoforms with orthologs that are conserved across different plant species such as maize, rice, potato, and Arabidopsis (Smith, 2001; Tetlow et al., 2004; Orzechowski, 2008). In theory, only starch synthases and starch branching enzymes are starch-synthesizing activities. However, other enzymes, some of which were previously thought to be involved in starch degradation, are also needed for normal starch biosynthesis.

Adenosine 5'-diphosphate-Glc pyrophosphorylase (AGPase) catalyzes the conversion of Glc-1-P and ATP into pyrophosphate (PPi) and ADP-Glc, which is the starch biosynthesis nucleotide-sugar donor (Orzechowski, 2008; Jeon et al., 2010). This reaction is the first

Table 2.1. Starch biosynthetic enzymes.

Enzymatic activity / Isoforms	Remarks
ADP-Glc pyrophosphorilase (AGPase; E.C. 2.7.7.27)	Synthesizes ADP-Glc for starch synthesis
Plastidial AGPase	Present in all tissues of all plants
Cytosolic AGPase	Present only in cereal endosperms
Starch synthases (E.C. 2.4.1.21)	Elongates α-1,4-glucan chains
Granule-bound starch synthase (GBSS)	Synthesizes amylose (and extra long amylopectin chains)
GBSSI	Specific to storage tissues
GBSSII	Specific to transient starch-synthesizing tissues
Soluble starch synthase (SS)	Synthesizes amylopectin
SSI	Elongates short glucan chains
SSII	Elongates intermediate glucan chains. Monocots have two isoforms SSIIa and SSIIb.
SSIII	Elongates long glucan chains
SSIV	Thought to be involved in starch granule initiation
Starch branching enzymes (BEs; E.C. 2.4.1.18)	Introduces α-1,6-linkages to form branches
BEI	Acts on low-branching polyglucan
BEII	Acts on high-branching polyglucan. Cereals have two isoforms BEIIa and BEIIb
Debranching enzymes (DBEs)	Hydrolyzes α-1,6-glycosyl linkages
Isoamylase-type (E.C. 3.2.1.68)	Acts on amylopectin and phytoglycogen
ISA1	Forms homo- and hetero- (with ISA2) multimeric complexes
ISA2	Forms heterocomplex with ISA1
ISA3	Thought to be involved in starch degradation
Pullulanase-type / limit dextrinase (E.C. 3.2.1.41)	Acts on amylopectin and pullulan
Plastidial starch / α-glucan phosphorylase (SP; E.C. 2.4.1.1)	Catalyzes reversible transfer of Glc from Glc-1-P to the non-reducing end of α-1,4-glucan chains. The exact function in starch biosynthesis is still unclear

committed, rate-limiting step in starch biosynthesis. AGPase is a heterotetrameric ($\alpha_2\beta_2$) enzyme consisting of two small (ApS) and two large (ApL) subunits, which differ only slightly in molecular masses and are encoded by related, but distinct, genes. For example, the potato tuber ApS and ApL are 50 and 51 kDa, respectively, and share 53% sequence identity (Okita et al., 1990; Nakata et al., 1991). These subunits, both of which contribute to catalysis (Ventriglia et al., 2008; Baris et al., 2009), may exist in multiple isoforms. For instance, there are one ApS and three ApL genes in Arabidopsis (Villand et al., 1993; Kleczkowski et al., 1999), while rice has two ApS and four ApL genes (Jeon et al., 2010, and references therein). AGPase is localized to the plastids in all plant tissues and species, except for the cereal endosperm where both plastidial and cytosolic AGPases exist (Jeon et al., 2010, and references therein). As the first enzyme in the starch biosynthesis pathway, AGPase is highly regulated by allosteric interactions at both the ApS and ApL subunits, as well as at the subunit interfaces (Boehlein et al., 2008, 2009; Georgelis et al., 2009; Boehlein et al., 2010; Kotting et al., 2010). Allosteric activation has been reported by, for example, the glycolytic intermediate 3-phosphoglycerate (3-PGA), fructose-6-P, and Glc-6-P, while allosteric inhibitors include ADP and orthophosphate (Pi). AGPase is also regulated by redox conditions in response to sugar levels in amyloplasts or diurnal cycles in chloroplasts (Kotting et al., 2010). Under oxidizing conditions, the two small subunits become covalently linked through an intermolecular disulfide bond at the N-termini, resulting in a stable dimer within the AGPase heterotetramer and deactivation of the AGPase activity. Reductive activation, thought to be mediated by thioredoxin, breaks the disulfide bond and reactivates the enzyme.

Starch synthases catalyze the transfer of α-1,4-Glc residues from ADP-Glc onto pre-existing glucan chains, thereby elongating the polymer. For a long time it was thought that

starch biosynthesis occurs by addition of Glc at the non-reducing ends of oligosaccharide primers. This was based on the fact that 1) incubation of starch granules with UDP-[^{14}C]Glc or ADP-[^{14}C]Glc resulted in incorporation of the radioactivity into products that were digestible by exo-acting β-amylase to produce radiolabeled maltose (Rongine De Fekete et al., 1960; Frydman and Cardini, 1964), and 2) incubation of starch granules with UDP-[^{14}C]Glc and maltodextrins resulted in incorporation of a single labeled Glc to the non-reducing ends of the maltodextrins (Leloir et al., 1961). However, recently Mukerjea and co-workers used ADP-[^{14}C]Glc pulse-chase experiments to provide evidence that starch synthesis in the granules occurs by addition of the Glc residues onto the reducing ends of the growing polymers by a two-site insertion mechanism, and that the enzymes form covalent, high-energy complexes with both D-glucose and the growing starch chains (Mukerjea et al., 2002; Mukerjea and Robyt, 2005). The same research group also demonstrated that maltodextrins actually serve to inhibit starch biosynthesis, which is contrary to the primer function that should stimulate the process (Mukerjea and Robyt, 2005). Despite this new finding, however, the directionality of starch synthesis remains a controversial topic that requires further research.

Amylose is synthesized by the granule-bound starch synthases GBSSI and GBSSII, while amylopectin is synthesized by soluble starch synthases SSI, SSII, SSIII, and SSIV (Ball and Morell, 2003; Orzechowski, 2008). GBSSI appears to be specifically expressed only in storage tissues, while GBSSII is localized to the chloroplasts in leaves and other transient-starch-synthesizing tissues (Fujita and Taira, 1998; Vrinten and Nakamura, 2000). The *waxy* mutants in cereals are defective in GBSSI and produce low-amylose or amylose-free starches (Nelson and Rines, 1962; Jeon et al., 2010). GBSSI has also been shown to be involved in amylopectin synthesis in rice, to form the extra-long unit chain (ELC) fraction (Hanashiro et al., 2008). SSI,

SSII, and SSIII appear to act sequentially, since SSI synthesizes the shortest amylopectin chains (DP 6-15), SSII synthesizes chains with intermediate length (DP 13-25), and SSIII synthesizes amylopectin chains with DP ≥ 30 (Jeon et al., 2010, and references therein; Tetlow et al., 2004). Monocots have two isoforms of SSII, i.e. SSIIa and SSIIb, which are expressed predominantly in storage and photosynthetic tissues, respectively (Tetlow et al., 2004; Ohdan et al., 2005). SSII and SSIII have been shown to have partially redundant functions. It was shown in Arabidopsis that SSIII could partially function to synthesize the intermediate amylopectin chain in the absence of SSII (Zhang et al., 2008). SSIV, and to some extent SSIII, appear to play a role in starch granule initiation. Chloroplasts in the leaves of Arabidopsis *ssiv* mutants contain only one large granule, in contrast to an average of five granules observed in chloroplasts of wild-type plants (Roldan et al., 2007). On the other hand, the *ssiii/ssiv* double mutant could not synthesize any starch granules although it still contained ~60% of wild-type starch synthase activity, suggesting that SSIII could partially rescue the granule initiation in the *ssiv* mutant (Szydlowski et al., 2009). SSIV was further shown to localize to specific regions associated with the edges of starch granules, supporting the proposed specific function of this SS isoform (Szydlowski et al., 2009).

Starch branching enzymes (BEs) introduce branch points by cleaving the α-1,4-glucan chain and reattaching the cleaved fragment, which contains ≥ 6 glucosyl residues, by an α-1,6-linkage to the same or another α-1,4-glucan chain (Jeon et al., 2010; Zeeman et al., 2010). The two classes of the enzyme, BEI and BEII, differ in substrate preferences and in the length of the produced chains. Maize and rice BEIs have a higher affinity for amylose and less-branched polyglucan, and transfer longer chains (DP > 16) (Guan et al., 1997; Vu et al., 2008; Nakamura et al., 2010). In contrast, BEII prefers to act on highly branched polyglucans and transfers

shorter chains (DP ≤ 12). Cereals also have two BEII isoforms with different expression patterns and catalytic functions. BEIIa is expressed in all tissues, acts on intermediate (DP 17-26) and long (DP ≥ 40) chains, and produces short chains of DP 6-15; while BEIIb is specific to the endosperm, prefers to act on intermediate chains, and creates short chains with smaller length distribution (DP 6-7) (Sawada et al., 2009; Nakamura et al., 2010). Mutations in BEs reduce the degree of branching in amylopectin and elevate the amylose content (Regina et al., 2010, and references therein).

Debranching enzymes (DBEs) cleave the α-1,6-linkages of the polyglucans. They are categorized into isoamylase (ISA), which acts on phytoglycogen and amylopectin, and pullulanase (PUL; also known as limit dextrinase), which acts on pullulan and amylopectin but not on phytoglycogen (Jeon et al., 2010). Plants have three ISA genes (*ISA1, 2, and 3*) and one PUL gene. Mutations of any of the DBEs have been shown to cause a lower amount of insoluble glucan and a higher amount of water-soluble glucan (phytoglycogen) content (Zeeman et al., 2010, and references therein; Streb et al., 2008; Jeon et al., 2010). Compared to amylopectin, phytoglycogen has a similar structure but contains a higher amount of shorter chains (DP < 10) and more branch points that are closer to each other and more randomly placed. This prevents phytoglycogen from forming the ordered structure of a starch granule. It was suggested that DBEs function in starch synthesis by clearing or trimming the random and miss-positioned branches, thus allowing aggregation of amylopectin in the granule. Intriguingly, no starch granules were observed in a quadruple mutant defective in all DBEs (Streb et al., 2008). However, synthesis of starch granules is recovered by introducing additional mutations in AMY3, another chloroplastic α-amylase. These data suggest that the starch granules can be

synthesized in the absence of DBEs, and thus, DBEs may facilitate but not be essential for starch granule synthesis.

Starch phosphorylase (SP) catalyzes the reversible transfer of Glc residues from Glc-1-P onto the non-reducing ends of α-1,4-glucan chains, with a release of orthophosphate (Pi) (Orzechowski, 2008; Jeon et al., 2010; Zeeman et al., 2010). Plant starch phosphorylase Pho1 is localized to the plastids. It has been shown that Pho1 transcript and protein expression, as well as enzyme activity, correlate with the period of starch biosynthesis and with the expression of other starch biosynthetic enzymes (Tetlow et al., 2004, and references therein; Tickle et al., 2009). A rice *pho1* mutant shows a reduction in starch content that manifests only at non-permissive temperatures (20°C) (Satoh et al., 2008). However, gene silencing or mutation in potato and Arabidopsis Pho1 do not result in any abnormal phenotype (Zeeman et al., 2010). The function of Pho1 with regards to starch synthesis therefore remains unclear, although several possible roles have been proposed. For examples, Pho1 was suggested to (1) synthesize glucan chains that serve as primers for starch synthases, (2) serve a similar clearing function as DBEs, or (3) convert maltooligosaccharides released by DBEs into Glc-1-P, which would be used as a substrate by AGPase (Tetlow et al., 2004; Zeeman et al., 2010, and references therein).

Apart from GBSSI and GBSSII, many of the other starch biosynthetic enzymes are found in two forms: as soluble proteins in the stroma and as internal granule-associated proteins. In maize endosperm, for example, SSI, SSIIa, SSIII, BEI, BEIIa, and BEIIb are all also found in granule-associated forms (Grimaud et al., 2008, and references therein). It is not known, however, whether there is a functional significance of having soluble and granule-associated forms (e.g. one form may be more active than the other), or whether granule association is a result of the enzymes getting trapped in the polymer during amylopectin crystallization. It has

been shown with maize SSI, which elongates short chains of amylopectin, that its catalytic activity decreases while the binding affinity increases as the glucan chains gets longer (Commuri and Keeling, 2001). It was suggested that the tight binding of SSI towards its reaction product causes the enzyme to become entrapped as an inactive enzyme within the starch granule. Interestingly, it was shown with the above maize enzymes that granule association of one starch biosynthetic enzyme may be dependent on the presence of other enzymes (Grimaud et al., 2008). For instance, the amount of SSI and BEIIb that are granule-associated is less in the maize SSIIa mutant than in wild-type, suggesting that there may be protein-protein interactions between the three enzymes.

Despite all the above description of the starch biosynthetic enzymes, little is known about how the starch synthesis is initiated. Synthesis of glycogen in yeast and mammals is initiated by the self-glucosylating protein glycogenin (Lomako et al., 2004). Glycogen, which serves as an energy storage molecule in animals, fungi, and bacteria, is also a polymer of glucose in α-1,4- and α-1,6-linkages similar to amylopectin (Zeeman et al., 2010). However, in comparison to amylopectin, glycogen has more branches that are arranged uniformly, instead of in clusters, resulting in a soluble glucan with no higher-order structures such as those observed in starch granules. Arabidopsis plant glycogenin-like protein 1 (PGSIP-1) has been suggested as the starch synthesis initiator in plants, based on the fact that it has sequence similarity to the yeast and mammalian glycogenin, and that its mutation leads to a reduction in starch content in leaves of Arabidopsis (Chatterjee et al., 2005). Arabidopsis PGSIP-1 and five other homologous proteins form a PGSIP gene family, which belongs to the GT8 family together with other glycan protein families, including galactinol synthases and pectin homogalacturonan:galacturonosyltranferases (Yin et al., 2010; Yin et al., In press). The

Arabidopsis PGSIPs, however, do not have conserved Tyr residues, which play a role in glycogenin as the sites of self-glucosylation, and only two Arabidopsis PGSIPs (AtPGSIP-1 and AtPGSIP-3) contain a predicted transit peptide for targeting to chloroplast (Chatterjee et al., 2005; Yin et al., In press). Furthermore, it was recently reported that AtPGSIP-1 and AtPGSIP-3 are involved in xylan biosynthesis (Mortimer et al., 2010). Therefore, no functional data is yet available to support the role of the PGSIPs in initiating starch synthesis in higher plants.

Phosphorylation and protein complexes in starch biosynthesis

Some starch biosynthetic enzymes have been found modified by phosphorylation. Incubation of amyloplasts from developing wheat (*Triticum aestivum*) endosperm with ATP and ADP-Glc significantly increases amylopectin synthesis (Tetlow et al., 2004). Analyses of stromal soluble and granule-associated proteins following the incubation of plastids with γ-^{32}P-ATP, identified BEI, BEIIa, BEIIb, and SP among the stromal soluble phosphoproteins, and BEIIa/b and SSIIa among the granule-associated phosphoproteins in amyloplasts. Phosphoproteins BEI and BEIIa were also identified in the stroma of chloroplasts. However, only the amyloplast and chloroplast stromal soluble (and not the granule-associated) BEIIs were found activated by phosphorylation, as indicated by a decrease in the enzymatic activity of these enzymes following dephosphorylation using alkaline phosphatase. The phosphorylation status of BEI, BEIIb, and SP in the wheat amyloplast stroma also determines their ability to form a multiprotein complex (see below). Other phosphoproteins that have been identified include the small and large subunits of AGPase, SSIII, and PUL in Arabidopsis; and GBSS, BEIIb, and SP in maize (Kotting et al., 2010, and references therein; Grimaud et al., 2008).

In the past few years, multiple protein complexes containing starch biosynthetic enzymes have been identified in chloroplasts and amyloplasts of cereals and other plant species. The involvement of protein complexes has previously been proposed in amylopectin synthesis, to explain the pleiotropic effects seen with cereal mutants defective in SSs, BEs, and DBEs (Ball and Morell, 2003; Tetlow et al., 2004). For example, the maize *du1* mutant is defective in SSIII but also displays a reduction in BEIIa activity (Cao et al., 1999; Cao et al., 2000), while the rice BEIIb-deficient *amylose extender* (*ae*) mutant has a substantial decrease in SSI activity (Nishi et al., 2001). Indeed, the concerted action of different starch biosynthetic enzymes may be coordinated through physical interactions among the enzymes, and a combination of different isoforms in different complexes may lead to the synthesis of starch with a specific structure (Zeeman et al., 2010).

Analyses of wheat amyloplasts demonstrated the co-immunoprecipitation of BEI, BEIIb, and starch phosphorylase from amyloplast lysates, suggesting that the three proteins interact in a protein complex (Tetlow et al., 2004). Interestingly, the complex formation is phosphorylation-dependent, since the co-immunoprecipitation was abolished when the lysate was pre-treated with alkaline phosphatase. The co-immunoprecipitation was not observed with lysates from chloroplasts, despite the fact that BEs are also phosphorylated in the chloroplast stroma (see above). Further research using chemical cross-linking, co-immunoprecipitation, gel filtration chromatography, and mass spectrometry, identified ~260 kDa protein complexes that contain SSI and phosphorylated SSIIa and BEIIa/b isoforms, and an ~180 kDa complex that likely corresponds to BEIIa/b homodimers (Tetlow et al., 2008). It was shown that the branching enzymes in the ~260 kDa complex have a higher affinity (lower K_d values) towards starch and amylopectin, in comparison to the monomeric BEIIa/b isoforms. The protein complexes,

however, were only observed in amyloplast lysates prepared from wheat endosperm in relatively late stages of development (10-15 days after pollination (DAP) or later), and not from early stage endosperm (6-9 DAP). These data clearly indicate that the formation of starch biosynthetic protein complexes has both biochemical and developmental functional significance.

Similar protein complexes were also found in maize amyloplasts. At least one of three methods (i.e. yeast two-hybrid system, affinity purification using immobilized recombinant proteins, or co-immunoprecipitation) showed that SSI can interact with SSIIa, SSIII, BEI, BEIIa, and BEIIb; SSIIa can interact with BEIIa and BEIIb; and SSIII can interact with BEIIa in maize amyloplasts (Hennen-Bierwagen et al., 2008). Gel filtration chromatography of maize amyloplast lysates further identified an ~300 kDa complex that contains SSIIa, BEIIa, and BEIIb, and an ~670 kDa complex that contains all the three proteins and SSIII (Hennen-Bierwagen et al., 2008; Hennen-Bierwagen et al., 2009). These protein-protein interactions and complexes were not abolished by pretreatment of the lysates with amyloglucosidase and α-amylase, indicating that the interactions are not due to binding of different enzymes to a common glucan chain. Mutation of any one of the SSIIa, SSIII, BEIIa, and BEIIb, prevents the formation of the 670 kDa complex, indicating that all four proteins are involved in, and necessary for, the formation of the high molecular weight complex (Hennen-Bierwagen et al., 2009). Mass spectrometry analysis of the partially purified 670 kDa complex identified several additional proteins including pyruvate orthophosphate kinases (PPDK1 and 2), while analysis of the 300 kDa complex identified a sucrose synthase isoform (SUS-SH1) and starch phosphorylase. These findings were confirmed by affinity purification and co-immunoprecipitation results, which verify the protein-protein interactions between SSIII and SSIIa, BEIIa, BEIIb, PPDK1 and 2, SUS-SH1, and also between SSIII and AGPase. Similar to observations in wheat, co-

immunoprecipitation of these proteins is phosphorylation-dependent: it was reduced significantly by pretreatment with alkaline phosphatase and increased in the presence of a phosphatase inhibitor. It was also shown that, in the absence of BEIIb in the maize *amylose extender* (*ae*) mutant, a larger molecular weight complex was formed that contains SSI, SSIIa, BEI, BEIIa, and starch phosphorylase, and that is not affected by phosphorylation/dephosphorylation (Liu et al., 2009). The maize *ae* mutant has amylopectin with less branch points and longer glucan chains. The functional significance of these complexes was supported by the fact that the starch granule-associated proteins include SSI, SSIIa, and BEIIb in wild-type amyloplasts, and include SSI, SSIIa, BEI, BEIIa, and starch phosphorylase in the *ae* mutant amyloplasts. These data suggest that in the absence of one subunit of the complex in the mutant, other starch biosynthetic enzymes are recruited to form an active but functionally altered amylopectin biosynthetic complex.

The debranching enzymes ISA1 and ISA2 also exist as protein complexes. ISA1 has enzymatic activity while ISA2 does not, as judged by protein sequence analysis and heterologous expression in *E. coli* of potato tuber proteins (Hussain et al., 2003). However, both proteins co-purify as a protein complex of ~500 kDa. In support of this finding, Arabidopsis *isa1* and *isa2* mutants, as well as the *isa1/isa2* double mutant, display identical phenotypes, including an absence of isoamylase activity, reduced starch content, and an accumulation of phytoglycogen (Delatte et al., 2005). Homooligomeric ISA1 and heterooligomeric ISA1/ISA2 complexes have been described in rice and maize endosperms (Utsumi and Nakamura, 2006; Kubo et al., 2010). The homomeric ISA1 complex has a molecular weight of 420-480 kDa and ~300 kDa in rice and maize, respectively. Rice has a single heteromeric ISA1/ISA2 complex of 510-550 kDa that is thought to contain five ISA1 and one ISA2 molecule; while maize has two ISA1/ISA2

complexes, both of which are of ~400 kDa. It was further shown in rice that, compared to the homomeric ISA1 complex, the heteromeric ISA1/ISA2 complex has lower K_m values (and thus higher affinities) for amylopectin, glycogen, maize β-limit dextrin, and especially for phytoglycogen (Utsumi and Nakamura, 2006). In maize, two mutant lines were generated, one line has only the homomeric complex (*su-1-P*) while the other line has only one of the two heteromeric complexes (*isa2-339*) (Kubo et al., 2010). Both mutant lines have starch contents almost the same as wild-type, suggesting that both the homomeric and heteromeric complexes function in starch synthesis. However, the homomeric-only line generates more, but smaller, starch granules, indicating that the heteromeric ISA1/ISA2 complex has a specific function that is not replaceable by the homomeric ISA1 complex. Unfortunately, no analyses of the enzymatic activities of the different maize complexes were performed, and thus the functional significance of the complexes remains unclear.

An interaction between SSIII and a 14-3-3 protein of the ε subgroup has been reported (Sehnke et al., 2001). The 14-3-3 proteins comprise a protein family implicated in the regulation of many biological processes through protein-protein interactions (Oecking and Jaspert, 2009). RNA interference of the 14-3-3 protein, which is a starch granule-associated protein, resulted in increased starch content in Arabidopsis leaves (Sehnke et al., 2001). Results from affinity purification using an immobilized 14-3-3 protein column support SSIII as a binding target of the 14-3-3 protein and indicate a protein-protein interaction between the two proteins. It was suggested that the 14-3-3 protein may serve to regulate starch synthesis by binding to, and inactivating, starch synthase enzymes.

HYALURONAN BIOSYNTHESIS IN VERTEBRATES

Structure and functions of the glycosaminoglycan hyaluronan

Hyaluronan (hyaluronic acid; HA) is a multifunctional extracellular polysaccharide that was first isolated in 1934 by Meyer and Palmer from bovine vitreous humor (Meyer and Palmer, 1934). It is a polymer of disaccharides [→4)-β-D-glucuronic acid-(1→3)-β-*N*-acetyl-D-glucosamine-(1→]$_n$ (Figure 2.4A) that consists typically of 10^3-10^4 glycosyl residues, resulting in very high molecular weight, polydisperse, linear molecules (10^6-10^7 Da), although it can also be found as smaller fragments and oligosaccharides under certain condition (DeAngelis and Weigel, 1994; DeAngelis et al., 1997; Pummill et al., 1998; Tammi et al., 2002). HA is a glycosaminoaglycan (GAG), which is a group of connective tissue polysaccharides whose other members include heparan sulfate, heparin, chondroitin sulfate, dermatan sulfate, and keratan sulfate (Lapcik et al., 1998). However, compared to the other GAGs, HA is the only one that is not modified and not covalently attached to peptides to form proteoglycans. Additionally, HA is synthesized at the plasma membrane instead of in the Golgi apparatus, where proteoglycans are synthesized (Kogan et al., 2007).

HA has been detected in every tissue and body fluid of higher animals, with the highest concentration in soft connective tissue and the lowest concentration in blood serum (Laurent and Fraser, 1992). HA in solution is highly solvated and can contain >1,000-fold more water on weight basis than organic matter (Ogston and Stanier, 1951, 1952). At physiological pH and ionic strength, the molecule is polyanionic and takes an extended, stiff random coil conformation in which each disaccharide repeat unit is rotated 180° relative to its neighbors, and with hydrogen bonding between neighboring glycosyl residues (Lapcik et al., 1998; Hascall and Esko, 2009). HA chains in solution are semi-flexible, forming highly viscous matrices that remain

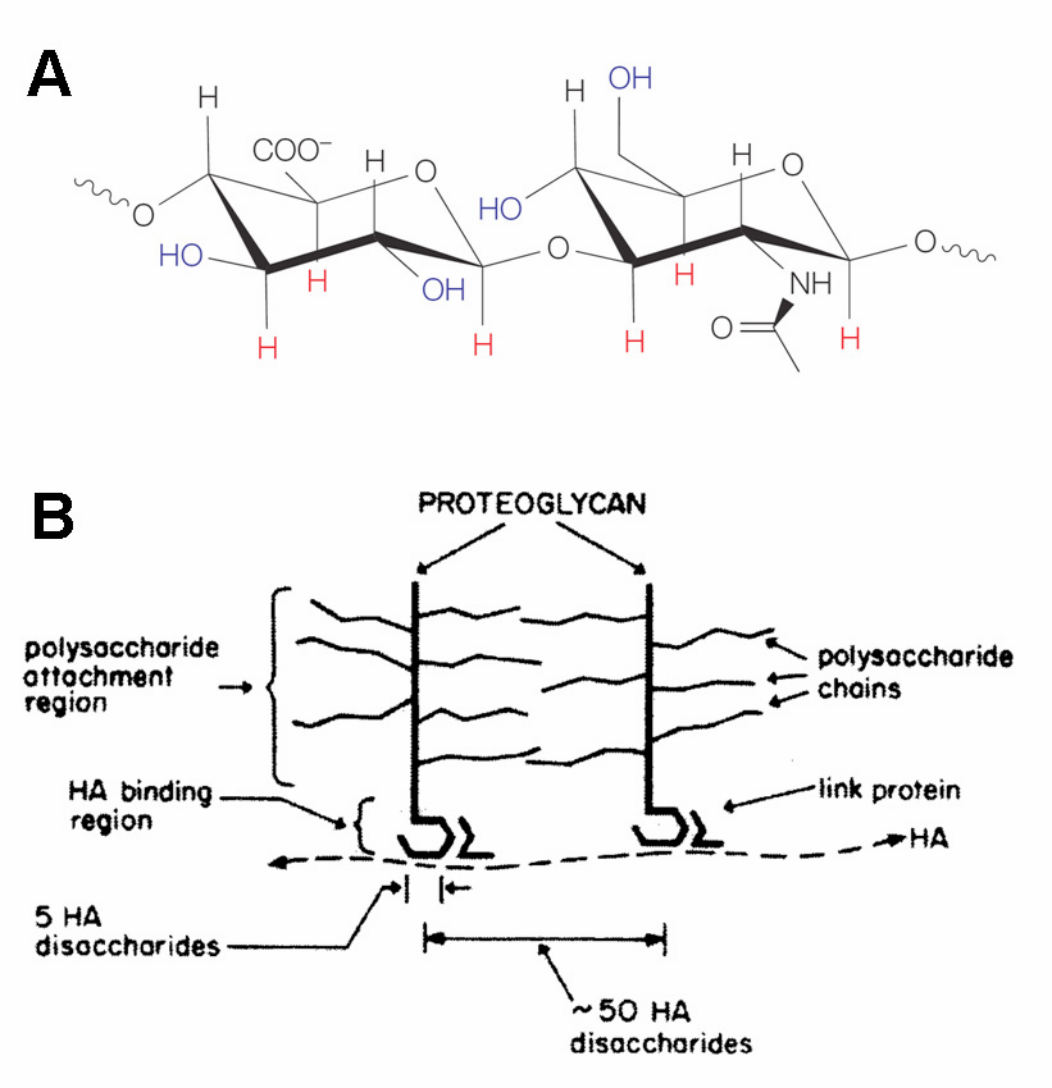


Figure 2.4. HA structure and function.

(A) The molecular structure of a disaccharide (β-1,3-GlcNAc-β-1,4-GlcA) repeat unit of HA. The carboxyl groups of GlcA contribute to the charged face of HA, while the axial hydrogen atoms (in red) contribute to the hydrophobic face of the molecule. (Adapted from Toole, 2004).

(B) Binding of proteoglycans at regular intervals onto HA chains to form a multimolecular aggregate. A proteoglycan binds, via its globular, N-terminal HA binding domain, to a stretch of 5 disaccharide repeats on the HA chain. This proteoglycan-HA interaction is mediated by a link protein. (Adapted from Hascall, 2000).

elastic when the viscosity goes down upon shear stress (Almond, 2007). This physical property makes HA ideal as a space-filler and lubricant, and explains its abundance in skin, skeletal tissues, vitreous fluid of the eye, and synovial fluid (Laurent and Fraser, 1992; Lapcik et al., 1998; Volpi et al., 2009). Interaction of HA with plasma membrane-bound proteins, such as the major HA receptor CD44, anchors HA onto the cell surface. Protruding into the ECM, the long HA chains become substrates onto which proteoglycans, e.g. aggrecan, versican, neurocan, and brevican, can bind and assemble to form huge multimolecular proteoglycan aggregates (Figure 2.4B) that are important for ECM formation and stability (Hascall, 2000; Tammi et al., 2002; Almond, 2007). HA of a certain size is also biologically active as a signaling molecule in certain cell types (Turley et al., 2002; Kogan et al., 2007). As a response to inflammation, tissue injury, or tumorigenesis, high molecular weight HA is fragmented by the actions of hyaluronidases or reactive oxygen species into smaller poly- and oligosaccharides, which then bind to cell-surface HA receptors to activate signaling cascades. Examples of such receptors are CD44 and RHAMM, whose interactions with HA have been implicated in cell adhesion, embryo development, angiogenesis, wound healing, and tumor cell growth and migration (DeAngelis, 1999; Turley et al., 2002; Kogan et al., 2007; Itano, 2008). HA has also been detected intracellularly in various compartments including the nucleus, ER, lysosomes, and Golgi, with potential functions in the inflammatory response (Hascall et al., 2004). Other biomedical and industrial uses of HA include applications in eye surgery, plastic surgery and implants, drug delivery, and as an osteoarthritis theurapetic (Laurent and Fraser, 1992; Lapcik et al., 1998; Kogan et al., 2007).

In addition to vertebrates, HA is also synthesized as a capsular cell coating by certain pathogenic bacteria of the genera *Streptococcus* and *Pasteurella* (DeAngelis, 1999, 2002). The

HA capsules, consisting of the same molecules that the hosts themselves make and thus being non-immunogenic, serve as a camouflage for the pathogens to evade the animal host's immune system by facilitating adhesion of the bacterial cells to the host cells, enhancing virulence of the bacteria, and reducing phagocytic action by the host cells. A functional HA synthase gene has also been identified in many viruses of the family Phycodnaviridae that infect unicellular chlorella-like green algae and cause the infected algal cells to produce HA, although the function of the HA in this case is still unknown (DeAngelis et al., 1997; Graves et al., 1999).

Small but mighty: hyaluronan synthases (HASs) are dual acting enzymes

Hyaluronan synthases (HASs; EC 2.4.1.212) are the enzymes that catalyze the biosynthesis of HA from the nucleotide-sugar precursors UDP-GlcA and UDP-GlcNAc (Weigel et al., 1997). Contradicting the previous view regarding carbohydrate biosynthesis, that the transfer of each sugar residue is catalyzed by a distinct enzyme, HASs are unique since they are the first single glycosyltransferase shown to incorporate two different sugars into a polymer chain (DeAngelis, 1999). HAS activity has been studied since the 1950's in both vertebrates and bacteria, however it was not until the 1990's that the first HAS gene was cloned from *Streptococcus pyogenes* (designated *spHAS*). Following this, several other HAS genes have been cloned including those from the bacteria *Streptococcus equisimilis* (*seHAS*) and *Pasteurella multocida* (*pmHAS*), the algae-infecting *Paramecium bursaria Chlorella* virus (PBCV-1; *cvHAS*), and the vertebrates African clawed frog *Xenopus laevis* (*xlHASs*), murine (*mmHASs*) and human (*hsHASs*) (reviewed in DeAngelis, 1999). In 1999, Paul DeAngelis proposed the classification of the HASs into two classes based on sequence similarities, predicted topology and enzyme characteristics (DeAngelis, 1999). Class I includes the *Streptococcal* and vertebrate

HASs, while the *P. multocida* enzyme (pmHAS) is the only member of Class II. Here, further discussion will focus only on the vertebrate HASs, and particularly on mammalian HASs, with some references or comparisons made to the bacterial HASs when necessary.

The vertebrate HAS gene family has three members, namely HAS1, HAS2, and HAS3, which have been identified in *Xenopus laevis*, mouse, and human (Spicer and McDonald, 1998; Itano and Kimata, 2002). The vertebrate HASs are 551-588 a.a., ~65 kDa in molecular mass, share 55-71% sequence homology, and display a similar predicted topology (Figure 2.5A) (reviewed in Weigel et al., 1997; Itano and Kimata, 2002; Itano, 2008)). All mammalian HASs are integral membrane proteins with seven transmembrane or membrane-associated domains, of which two are at the N-termini and the other five at the C-termini. As a comparison, the *Streptococcal* HASs are ~35% smaller and have only two and three membrane domains at the N- and C-termini, respectively. Located between the two clusters of membrane domains is a large cytoplasmic central domain, which is of 307-328 a.a. in vertebrate HASs, and comprises 54-56% of the total protein and up to ~88% of the intracellular portion of the protein. This central domain is relatively highly conserved among HAS orthologs and believed to contain the catalytic domain. Specifically, it contains several HA binding motifs $B(X_7)B$ identified previously in the HA receptor RHAMM, multiple phosphorylation consensus sequences, the D,D,QXXRW motif specific to processive β-glycosyltransferases, and several conserved residues also found in UDP-glucuronosyltransferases (UDP-GTs) (Yang et al., 1994; Saxena et al., 1995; Spicer et al., 1996; Spicer et al., 1997). While the predicted topology of the vertebrate HASs remains to be experimentally proven, a topological study on SpHAS indicated that the N-terminus, large central domain, and C-terminus of SpHAS are indeed located intracellularly (Heldermon et al., 2001), suggesting the accuracy of the vertebrate HASs topology prediction.

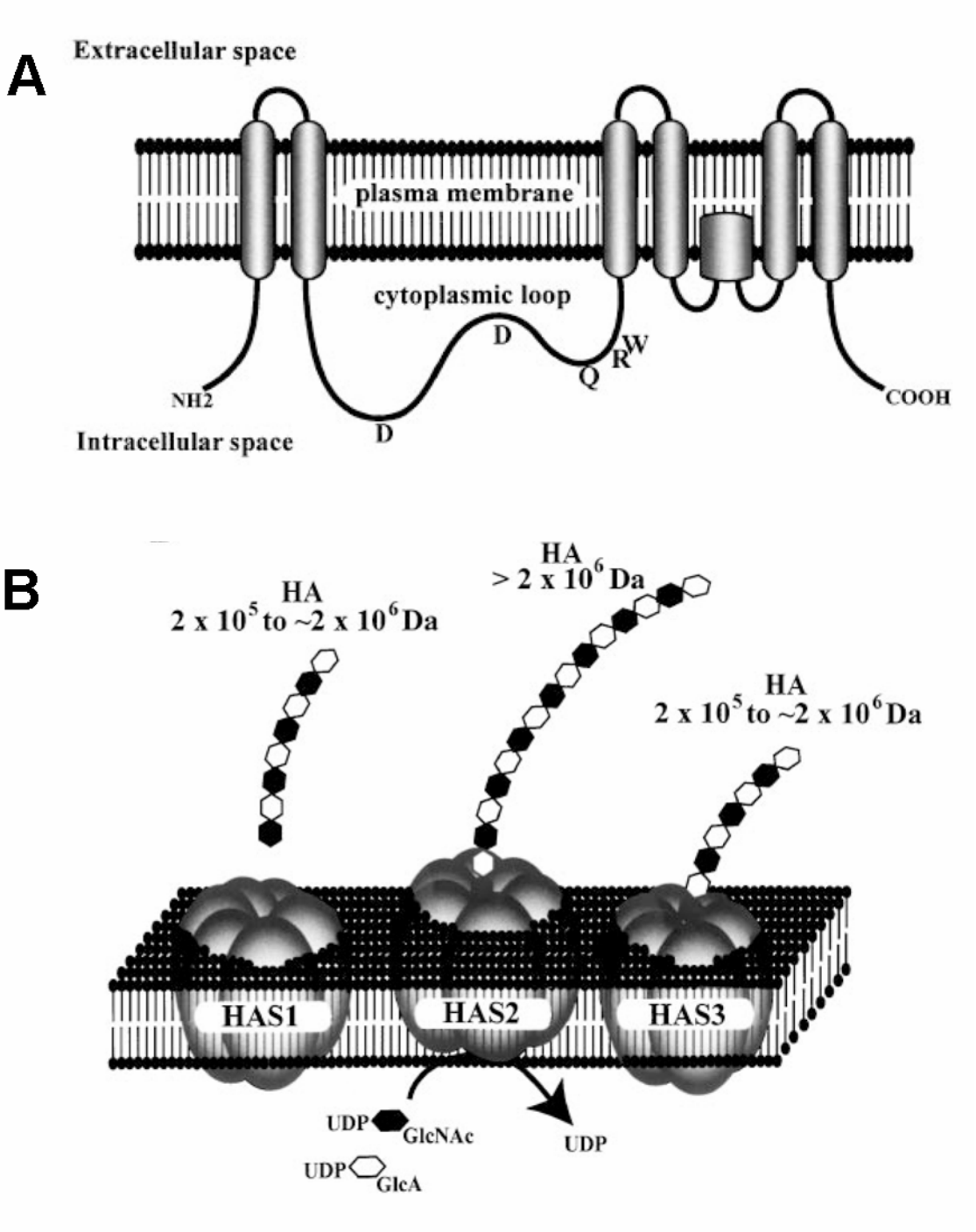


Figure 2.5. A topological prediction of a vertebrate HAS (A) and a hypothetical model of HA synthesis (B). (Adapted from Itano and Kimata, 2002).

(A) Vertebrate HASs are integral membrane proteins with seven putative membrane domains (in contrast to *Streptococcal* HASs, which do not have the last two transmembrane domains). The

N-terminus, a large central domain, and the C-terminus are located cytoplasmically. The D,D,QXXRW motif common to processive β-glycosyltransferases is indicated.

(B) HASs synthesize HA chains by alternating incorporation of GlcA and GlcNAc from sugar donors UDP-GlcA and UDP-GlcNAc. The three HAS isoforms are different in regards to the size of the HA produced (HA sizes shown are of *in vivo* synthesized HA), and in the enzyme characteristics (see text for more details). The model shows that the transmembrane domains of HASs form pores through which the HA chain is thought to be extruded. This model may no longer be accurate due to the identification of HA exporting ABC-transporters, as described in the text.

Individual expression of the vertebrate HASs in mammalian cells resulted in HA synthesis, suggesting that each of the HAS1, 2, and 3 enzymes function independently and are enzymatically active (Meyer and Kreil, 1996; Spicer and McDonald, 1998; Itano et al., 1999). Heterologous expression of mmHAS2 in *Drosophila* also resulted in a massive production of HA in this organism that does not normally synthesize HA (Takeo et al., 2004). All the class I HASs require Mg^{2+} for maximal activity, in contrast to the class II pmHAS that prefers Mn^{2+} (DeAngelis, 1996). The HASs specifically use UDP-GlcA and UDP-GlcNAc as sugar donors, and do not use UDP-Glc, UDP-Gal, or UDP-GalNAc, as shown with xlHAS1 (Pummill et al., 1998). As reflected by the K_m values, the HASs have higher affinity for UDP-GlcA than for UDP-GlcNAc (Pummill et al., 1998; Itano et al., 1999; Yoshida et al., 2000). Interestingly, when supplied with UDP-GlcNAc alone, recombinant mmHAS1 synthesizes chito-oligosaccharides (Yoshida et al., 2000). Site directed mutagenesis on the Asp_{242}, Asp_{344}, Gln_{380}, and Trp_{384} residues of the D,D,QXXRW motif in mmHAS1 resulted in significant decreases in both HA synthase and chito-oligosaccharide synthetic activity. On the contrary, mutagenesis of Leu_{314}, which is conserved in both HASs and UDP-GTs (see above), reduced only the HA synthase activity. These data suggest that both the D,D,QXXRW motif and Leu_{314} are involved in HA synthase activity, however only the D,D,QXXRW motif is required for the chito-oligosaccharide synthesis/β-1,4-GlcNAc transferase activity. The fact that Leu_{314} is located within the D,D,QXXRW motif led to suggestions that the catalytic sites for the GlcNAc and Glc transferases may be close to each other, and that the active site can accommodate both donor binding and catalysis. With regards to the direction of chain elongation, it has been demonstrated that the addition of new sugar units occurs at the non-reducing end for the

vertebrate xlHAS1, and at the reducing end for the bacterial spHAS (Bodevin-Authelet et al., 2005).

Analyses of transcript expression of the HASs, and of HAS activity in transfected mammalian cells, indicated that HASs 1, 2, and 3 each have a distinct expression pattern in both developing embryo and in adults, and suggests that each serves a developmental stage-, tissue-, and even cell-specific biological function (Spicer and McDonald, 1998). The three isoforms also display different enzymatic properties, as demonstrated in an enzymological study of mmHASs expressed in COS-1 and rat 3Y1 fibroblast cells (Itano et al., 1999). For example, mmHAS1 produces smaller pericellular HA coats in both cell types, has higher K_m values for both UDP-GlcA and UDP-GlcNAc, and loses its activity more quickly, compared to mmHAS2 and 3. On the other hand, mmHAS3 has a significantly lower HA elongation rate, and produces HA with lower average molecular mass both *in vitro* ($1 \times 10^5 - 1 \times 10^6$ Da) and *in vivo* (2×10^5 to ~2×10^6 Da) (Figure 2.5B), in comparison to the other two HASs. Overall, mmHAS2 produces HA with the highest average molecular mass *in vitro* (2×10^5 to ~2×10^6 Da) and *in vivo* (> 2×10^6 Da) (Figure 2.5B). The difference in the HA sizes synthesized by the different HAS isoforms may reflect tissue- or cell-specific size-dependent biological/physiological functions of HA.

HAS activity of plasma membrane and intracellular HASs is regulated by post-translational modifications

The HASs were originally thought to localize only at the plasma membrane. Recent results, however, indicate that this view is inaccurate. Experiments using fluorescently-labeled mmHAS2 and mmHAS3 showed that HASs are also present intracellularly in the ER, Golgi, and endocytic vesicles (Rilla et al., 2005). In support of this, Vigetti and co-workers detected HAS

activity in both plasma membrane and ER/Golgi membrane fractions, but not in the nuclear membrane fraction (Vigetti et al., 2009). Goentzel and co-workers previously showed that hsHAS3 is modified by phosphorylation of Ser residues (Goentzel et al., 2006). The hsHAS3 phosphorylation was observed under normal (non-stimulating) conditions, but could be further stimulated by treating the hsHAS3 expressing COS-7 cells with CPT-cAMP [8-(4-chlorophenylthio)-cyclic adenosine monophosphate], PMA (phorbol 12-myristate 13-acetate), and EGF (epidermal growth factor), which are activators for protein kinase A (PKA), protein kinase C (PKC), and protein tyrosine phosphorylation, respectively. It was further demonstrated that treatment of cells with PMA, which induces phosphorylation by PKC, increases the amount of HAS activity recovered in the plasma membrane fraction, but not in ER/Golgi membrane fraction (Vigetti et al., 2009). On the other hand, treatment of the plasma membrane fraction with alkaline phosphatase reduced the HAS activity. These data suggest that phosphorylation of HASs activates the HAS activity, although the possibility remains that phosphorylation may also affect other proteins that are involved in, or regulate, HAS activity (Vigetti et al., 2009).

HAS activity is also affected by N-glycosylation (Vigetti et al., 2009). Treatment of cells with tunicamycin, which inhibits N-glycosylation and induces ER stress, increased HAS activity in the plasma membrane and ER/Golgi membrane fractions. The same effect was observed when membrane fractions from untreated cells were treated with N-glycosidase F. Again, this post-translational modification may occur directly on HASs, or on other proteins that regulate HASs. Tunicamycin treatment was shown previously to stimulate the formation of intracellular HA cables, which span from perinuclear and/or ER membranes to the extracellular space, and play roles in inflammation to recruit monocytes (Hascall et al., 2004). It is thus plausible that the intracellular HASs are involved in synthesizing the HA cables.

Roles of ABC transporter and protein complexes in HA biosynthesis

HASs at the plasma membrane were also thought to simultaneously catalyze HA synthesis and extrude the growing HA chains to the extracellular space. This is proven not true by a recent finding in *Streptococcus pyogenes* that HA synthesis and capsular formation are inhibited by an insertional mutation in the gene encoding an ATP-binding cassette (ABC) transporter (Ouskova et al., 2004), which is homologous to human multidrug resistance transporter ABC-C (MRP5) (Prehm and Schumacher, 2004). Treating human fibroblast cells with compounds known to inhibit MRP5 and other multidrug resistance transporters, resulted in a reduction in HA production and in proliferation of fibroblasts (Prehm and Schumacher, 2004), which require HA for detachment during cell division and growth (Brecht et al., 1986). Several of the inhibitors also decrease HA synthase activity, suggesting the coordination between HA synthesis and transport. Comparison of the specificity of the inhibitors and the observed inhibitory effects pointed to MRP5 as the most likely candidate for an HA transporter. Further investigation demonstrated the inhibition of HA export in a dose-dependent manner upon silencing of MRP5 expression in mouse fibroblasts by RNA interference (Schulz et al., 2007). HA export by MRP5 could be competed by HA oligosaccharides that were introduced into the cells by osmotic lysis of pynocytotic vesicles, and by MRP5 substrates such as fluorescein and cGMP. While their exact organization remains to be elucidated, the coordination between the synthesis and transport of HA may suggest an interaction between the HA synthesizing enzymes and the HA transporter.

In certain cancer cell lines, *hsHAS1* has been found to generate aberrant splice variants named Va, Vb, and Vc, collectively referred to as HAS1-Vs (Ghosh et al., 2009). While normal, full-length HAS1 (HAS1-FL) is typically localized at the plasma membrane, HAS1-Vs (i.e.

HAS1-Va, Vb, and Vc) are located intracellularly in the cytoplasm and/or the Golgi apparatus, and are able to synthesize intracellular HA. Intriguingly, co-expression of HAS1-FL and HAS1-Vs in HEK293 cells resulted in the expression of these proteins intracellularly, suggesting that HAS1-Vs affect the subcellular localization of HAS1-FL. Immunoprecipitation experiments demonstrated that each HAS1-Vs variant interacts with itself, with other HAS1-Vs variants, and with HAS1-FL. Chemical cross-linking and non-reducing SDS-PAGE experiments further identified multimeric protein complexes containing HAS1-Vs and HAS1-FL, which could be released as monomers by treatment with the reducing agent β-mercaptoethanol, indicating that HAS1 and its splice variants can multimerize into heterooligomers via intermolecular disulfide bonds. At least HAS1-Vc was shown to stimulate anchorage-independent growth of transfected murine fibroblast NIH3T3 cells *in vitro* and subcutaneous tumor growth *in vivo*, confirming its role in tumorigenesis (Ghosh et al., 2009). This is in agreement with many reports showing that overexpression of HASs and exogenous HA treatment lead to enhanced cancer cell growth and malignancy (reviewed in Tammi et al., 2002; Toole, 2004; Volpi et al., 2009). It would be interesting to see if the formation of HAS1 multimeric complexes also occurs in normal cells, and if there is a connection between the complex formation and HAS activity. It is surely plausible that a protein complex between HASs and the ABC transporter is necessary for the HA synthesizing and exporting activities to function properly. Further research will be necessary to answer these questions.

HEPARAN SULFATE BIOSYNTHESIS IN VERTEBRATES

Structure and function of the glycosaminoglycan heparan sulfate

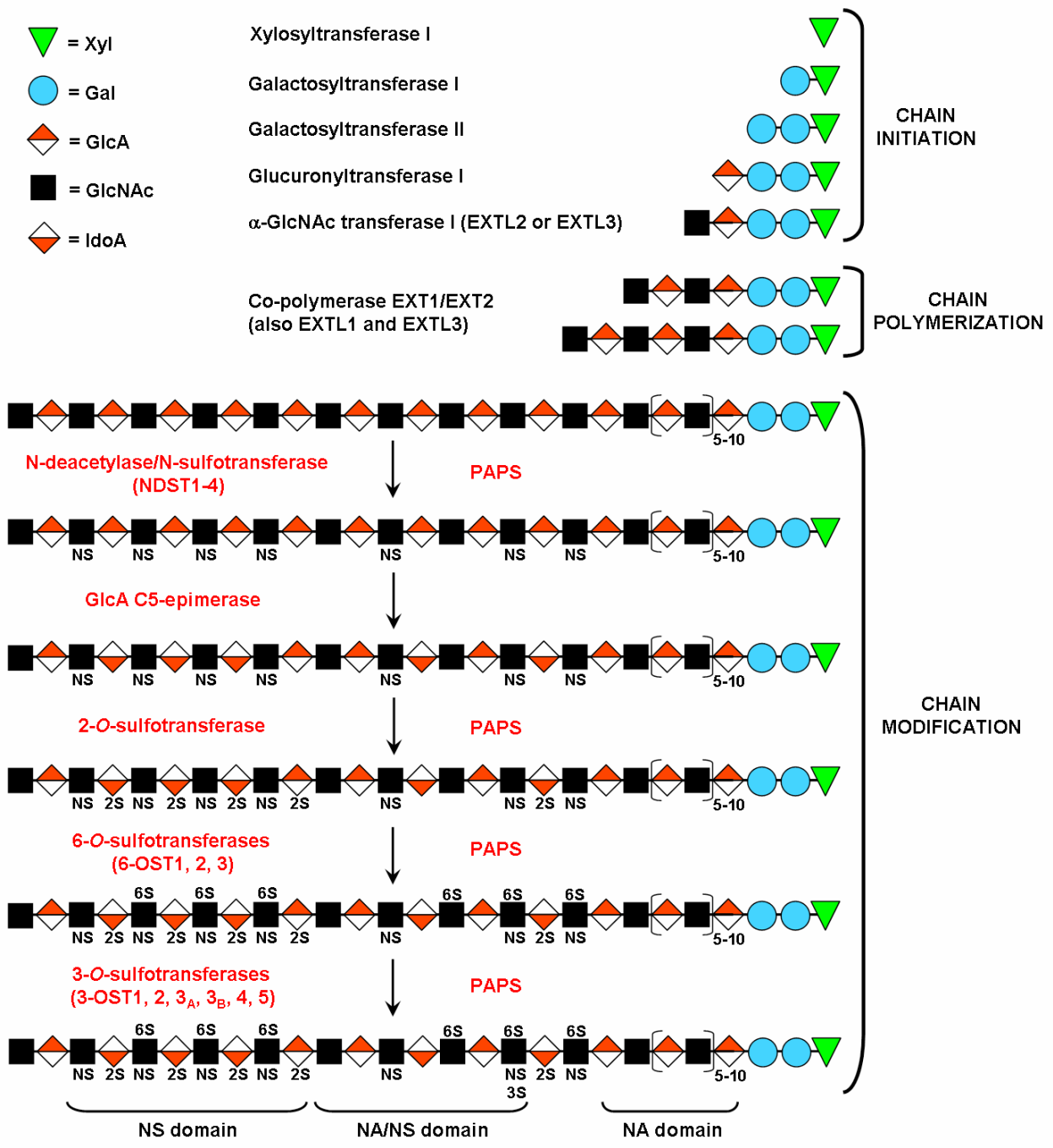
Heparan sulfate (HS) is a linear polyanionic extracellular polysaccharide found on the surface of almost all multicellular animal cells, including *Hydra*, *C. elegans*, *Drosophila*, zebrafish, mouse, and human (Bishop et al., 2007). It consists of a linkage tetrasaccharide [→4)-β-D-GlcA-(1→3)-β-D-Gal-(1→4)-β-D-Gal-(1→4)-β-D-Xyl-(1→] that is further elongated by 50-200 units of alternating disaccharide repeat [→4)-β-D-GlcA-(1→4)-α-D-GlcNAc-(1→] (Figure 2.6), resulting in a molecular mass of 25-100 kDa (Whitelock and Iozzo, 2005; Skidmore et al., 2008). The polysaccharide chain is covalently attached via an *O*-linkage to a Ser residue of a core protein, forming a heparan sulfate proteoglycan (HS-PG) comprised of both the core protein and the polysaccharide chain(s) (Whitelock and Iozzo, 2005). The common core proteins in mammals include: 1) syndecan, a family of type I transmembrane proteins with extended extracellular domains (Figure 2.7); 2) glypican, a family of globular proteins with glycosyl phosphatidyl inositol (GPI) anchors (Figure 2.7); and 3) basement membrane proteins collagen XVIII, agrin, and perlecan (Bernfield et al., 1999; Whitelock and Iozzo, 2005). The core proteins play a role in determining HS functionality by presenting the HS chains in the ECM, and as such control the level of HS expression.

The disaccharide repeat portion of HS is further modified by replacement of the *N*-acetyl group of some of the GlcNAc residues with a sulfate group, by partial epimerization of β-D-GlcA into α-L-iduronic acid (IdoA), and by partial *O*-sulfations of GlcNAc residues at C-6 and C-3, and of GlcA or IdoA at C-2 (Whitelock and Iozzo, 2005). HS was first identified as a contaminant in the preparation of heparin, another GAG that has the same biosynthetic pathway as HS. In comparison to HS, heparin is more extensively modified, such that it can be

Figure 2.6. A schematic of heparan sulfate structure and biosynthesis.

Legend at the top left corner provides the identity of each sugar residue. The heparan sulfate biosynthetic enzymes involved in each step are indicated. NS domain: highly sulfated region; NA domain: un-modified region; NA/NS domain: mixed region containing both un-modified and sulfated residues.

(Adapted with modification from Esko and Selleck, 2002).


= Xyl
= Gal
= GlcA
= GlcNAc
= IdoA
Xylosyltransferase I
Galactosyltransferase I
Galactosyltransferase II
Glucuronyltransferase I
α-GlcNAc transferase I (EXTL2 or EXTL3)
CHAIN INITIATION
Co-polymerase EXT1/EXT2
(also EXTL1 and EXTL3)
CHAIN POLYMERIZATION
N-deacetylase/N-sulfotransferase
(NDST1-4)
PAPS
GlcA C5-epimerase
2-O-sulfotransferase
6-O-sulfotransferases
(6-OST1, 2, 3)
3-O-sulfotransferases
(3-OST1, 2, 3A, 3B, 4, 5)
CHAIN MODIFICATION
NS
2S
6S
3S
5-10
NS domain
NA/NS domain
NA domain

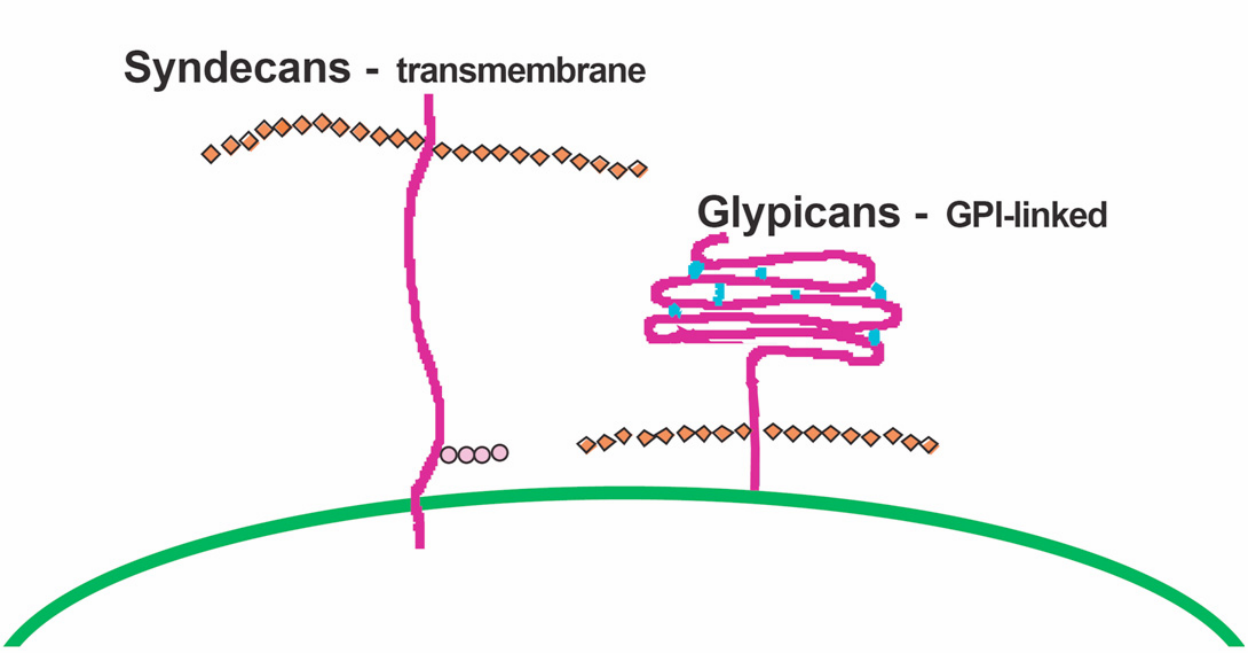


Figure 2.7. A schematic of two types of heparan sulfate-harboring proteoglycans at the cell surface, illustrating the site of heparan sulfate chain attachment in different proteoglycans. Orange diamonds represent heparan sulfate chains, while pink circles represent a chondoritin sulfate chain that may also be present in syndecans. (Adapted from Bernfield et al., 1999).

considered as a more homogenous, highly sulfated version of HS (Linker et al., 1958; Whitelock and Iozzo, 2005; Skidmore et al., 2008). The rather clustered modifications of HS result in blocks of highly sulfated regions (6-10 disaccharide units), which are negatively charged and relatively rigid (highly sulfated domains (NS domains)) (Figure 2.6) (Bernfield et al., 1999; Skidmore et al., 2008). These regions are interspersed with more flexible, unmodified regions (NA domains; 16-20 disaccharide units) and with transitional mixed regions (NA/NS domains) that lie between the other two domains (Figure 2.6). All the modifications, and their extent and distribution along the polymer chains, contribute to the high sequence diversity of HS which accommodates its function to bind a vast assortment of protein ligands (Bernfield et al., 1999).

Hundreds of protein ligands have been reported to interact with HS chains, including cytokines (e.g. chemokines, interleukins, stromal cell-derived factor 1 (SDF-1), tumor necrosis factor α (TNF-α)); growth factors (e.g. fibroblast growth factors (FGFs) and their receptors (FGFRs), transforming growth factors (TGFs), vascular endothelial growth factors (VEGFs)); lipases, proteases, and protease inhibitors; and various ECM proteins (e.g. fibronectin, laminin, collagens, endostatin) (Whitelock and Iozzo, 2005; Bishop et al., 2007). The ligand binding occurs through ionic interactions between particular sequences of basic amino acid residues (Lys, Arg) on the surface of the protein ligands and specific, negatively charged glycosyl sequences on the HS chains (Whitelock and Iozzo, 2005). Binding to HS serves to modulate the ligand activity by increasing the local concentration of the ligands, and thereby augmenting receptor activation at low ligand concentrations, or by creating and maintaining gradients of the signaling molecules (Bernfield et al., 1999; Whitelock and Iozzo, 2005). The functionality of HS is therefore very important in cell signaling, morphogenesis, and development, as well as in angiogenesis, inflammation, and tissue injury and repair (Whitelock and Iozzo, 2005; Bishop et

al., 2007). In human, a defect in HS chain elongation, due to mutations in the polymerizing enzymes encoded by the tumor suppressor genes *EXT1* and *EXT2* (see below), results in hereditary multiple exostoses (HME), a bone disorder characterized by cartilaginous tumor growth (exostoses/osteochondromas) at the growth plate of long bones that can turn malignant in 1-2% of patients. Mutations that affect HS-PGs have also been implicated in tumor progression, Alzheimer's and Parkinson's diseases, atherosclerosis, diabetes, and pathogen infections (Whitelock and Iozzo, 2005; Bishop et al., 2007; Nadanaka and Kitagawa, 2008).

Overview of heparan sulfate biosynthesis

The biosynthesis of HS is catalyzed by a series of enzymes that work sequentially. The process is highly specific and tightly regulated in that the reaction product of one enzymatic activity becomes the substrate for the next enzyme activity (Esko and Selleck, 2002; Whitelock and Iozzo, 2005). HS synthesis can be divided into three stages: chain initiation, chain elongation, and chain modification (Figure 2.7). Chain initiation comprises the synthesis of the tetrasaccharide core on the Ser residues of a proteoglycan protein (linkage region) and the transfer of an α-1,4-GlcNAc residue onto the tetrasaccharide core structure. Since HS and chondroitin/dermatan sulfate (CS/DS) share the same tetrasaccharide core structure, the transfer of this first α-1,4-GlcNAc becomes a critical, determining step to direct the biosynthetic process towards making an HS chain, instead of making a CS/DS chain that is determined by a transfer of a β-1,3-GalNAc. Chain elongation takes place by polymerization of the alternating disaccharide α-1,4-GlcNAc-β-1,4-GlcA. Finally, the HS chain modifications, as described in the previous section, occur as the HS chain is being polymerized and serve to create specific binding-sites for the various HS-binding proteins.

The enzymes involved in HS biosynthesis have all been identified and the encoding genes cloned (see below; Figure 2.7; Table 2.2). These proteins, with the one exception of the 3-*O*-sulfotransferase 1 (3-OST1) that lacks a transmembrane domain, are type II transmembrane proteins functioning in the Golgi apparatus (Shworak et al., 1999; Esko and Selleck, 2002; Schon et al., 2006). Moreover, as will be described below, the HS biosynthetic enzymes comprise a mixture of single-acting enzymes, dual-acting enzymes, and protein complexes. Some of the enzymes, especially those modifying the HS chains, are present in multiple isoforms that display different expression patterns. Esko and Selleck have suggested that the HS biosynthetic enzymes may engage in protein complexes, named GAGOSOMEs, which contain different compositions of enzymes and/or isoforms (Esko and Selleck, 2002). Such protein complexes may help explain the high efficiency and rapidity of HS synthesis, and the diversified but specific structures of HS chains. The fact that many of the HS biosynthetic enzymes do form homo- and/or hetero-oligomers (see below) supports this hypothesis.

Initiation of heparan sulfate synthesis

HS, CS, and DS share the same tetrasaccharide linkage region, and enzymes that synthesize this core structure. The first and rate-limiting step in GAG biosynthesis is the transfer of a xylose residue from UDP-Xyl to a Ser residue of the Ser-Gly/Ala dipeptides, which typically have one or more acidic residues in close proximity, in a proteoglycan core protein (Esko and Zhang, 1996; Schon et al., 2006). This step is followed by sequential additions of the two Gal in β-1,4 and β-1,3 linkages, respectively, and the GlcA residues in a β-1,3 configuration. The four reactions are catalyzed by single enzymes xylosyltransferase I (XT-I), galactosyltransferase I (GalT-I), galactosyltransferase II (GalT-II), and glucuronyltransferase I (GlcAT-I), respectively

Table 2.2. Heparan sulfate biosynthetic enzymes.

(*) Isoforms refer to multiple genes encoding the same enzyme activity.

Enzyme	Enzymatic activity	Enzyme type
Heparan sulfate initiation (synthesis of linkage region)		
Xylosyltransferase I (XT-I)	Attach Xyl to Ser residue of core protein	Single-acting enzyme
Galactosyltransferase I (GalT-I)	Transfer Gal in β-1,4 linkage onto Xyl-O-Ser	Single-acting enzyme
Galactosyltransferase II (GalT-II)	Transfer Gal in β-1,3 linkage onto Galβ1-4 Xyl-O-Ser	Single-acting enzyme
Glucuronyltransferase I (GlcAT-I)	Transfer GlcA in β-1,3 linkage onto Galβ1-4Galβ1-4Xyl-O-Ser	Single-acting enzyme
Heparan sulfate chain elongation		
EXT1	HS co-polymerase (α-1,4-GlcNAcT/β-1,4-GlcAT)	Dual-acting enzyme, protein complex
EXT2	HS co-polymerase (α-1,4-GlcNAcT/β-1,4-GlcAT)	Dual-acting enzyme, protein complex
EXTL1	Transfer GlcNAc in α-1,4 linkage onto N-acetylheparosan and GlcAβ1-3Galβ1-*O*-C_2H_4NH-benzyloxycarbonyl	Single-acting enzyme
EXTL2	Transfer GlcNAc or GalNAc in α-1,4 linkage onto the tetrasaccharide linkage region	Dual-acting enzyme
EXTL3	Transfer GlcNAc in α-1,4 linkage onto *N*-acetylheparosan	Single-acting enzyme
Heparan sulfate chain modification		
N-deacetylase/N-sulfotransferase (NDST) (4 isoforms)*	N-deacetylation and N-sulfation of GlcNAc	Dual-acting enzyme; four isoforms
Glucuronyl C5-epimerase (Epi)	Convert GlcA into IdoA	Single-acting enzyme, protein complex
Uronosyl-2-*O*-sulfotransferase (2-OST)	2-*O*-sulfation of IdoA or GlcA	Single-acting enzyme, protein complex
6-*O*-sulfotransferase (6-OST) (3 isoforms)*	6-*O*-sulfation of glucosamine	Single-acting enzyme, homooligomer
3-*O*-sulfotransferase (3-OST) (6 isoforms)*	3-*O*-sulfation of glucosamine	Single-acting enzyme

(Esko and Selleck, 2002; Sugahara and Kitagawa, 2002). The human genes encoding these proteins have been cloned and their enzymatic activities demonstrated by heterologous expression (Kitagawa et al., 1998; Almeida et al., 1999; Okajima et al., 1999; Wei et al., 1999; Gotting et al., 2000; Bai et al., 2001; Schon et al., 2006). All four enzymes are type II transmembrane proteins, shown to localize to the medial Golgi by co-localization with the medial-Golgi marker α-mannosidase II (Wei et al., 1999; Bai et al., 2001; Pinhal et al., 2001; Schon et al., 2006). They are also ubiquitously expressed in human organs and tissues, in agreement with the widespread occurrence of GAGs. The crystal structure of GlcAT-I has been resolved (Negishi et al., 2003), and the amino acids important for substrate binding and catalysis of the GalT-I have recently been described (Talhaoui et al., 2010).

Human XT-I has been further characterized with regards to protein domain organization. The globular catalytic domain of XT-I is believed to start at a.a. position 261, since recombinant protein lacking the first 260 a.a. at the N-terminus still retains 100% catalytic activity (Muller et al., 2006). A DWD motif in the catalytic domain has been shown necessary for catalysis (Gotting et al., 2004). Recombinant protein truncation studies showed that XT-I requires its N-terminal sequence of 214 a.a., which comprises a cytosolic tail, TMD, and stem region, for proper Golgi localization (Schon et al., 2006). The requirement for the stem region of XT-I for retention in the Golgi, suggests that XT-I may interact or form a protein complex/aggregate via its stem domain with other proteins within the Golgi. In this case, the TMD may function only to anchor the protein in the membrane. Interestingly, it has been reported that a great amount of human XT-I is also secreted out of the cells in a soluble form, although its function in the extracellular matrix is still unclear (Kuhn et al., 2001). Cloning of a highly homologous XT-II

gene has been described, however its catalytic activity and possible function in GAG synthesis remains to be demonstrated (Gotting et al., 2000).

The GAG tetrasaccharide core structure is also modified by phosphorylation of the Xyl at C-2 position, and by sulfation of both Gal residues at the C-2 and/or C-6 positions (Gulberti et al., 2005). While both modifications have been identified in CS and DS, only the C-2 phosphorylated Xyl has been found in HS. Recently, over-expression of a kinase responsible for the phosphorylation of Xyl was shown to result in the production of more HS chains, but the size of HS chains remains unchanged, leading to a suggestion that C-2 phosphorylation of Xyl functions to regulate the number of HS (or GAG) chains synthesized (Koike et al., 2009).

The critical addition of a α-1,4-GlcNAc at the non-reducing end of the tetrasaccharide linkage region commits the sugar chain to HS synthesis (Esko and Zhang, 1996; Esko and Selleck, 2002). In contrast, addition of a β-1,4-GalNAc commits the sugar chain to CS/DS synthesis, while the addition of an α-1,4-GalNAc prevents any further polymerization. Transfer of the α-1,4-GlcNAc is catalyzed by glucuronosyltransferase II (GlcAT-II). Since a proteoglycan core protein may bear both HS and CS/DS chains, the specificity of GlcAT-II for the protein acceptor is based on recognizing the amino acids in the vicinity of the tetrasaccharide. These amino acids include clusters of acidic residues, hydrophobic residues, and tandemly repeated Ser-Gly motif. Two proteins, EXTL2 and EXTL3, have been shown to have the GlcAT-II activity. Discussions on EXTL2 and EXTL3 are presented in the next section on HS chain elongation, since EXTL2 and EXTL3 are members of the EXT protein family, the other members of which are involved in HS chain elongation.

Heparan sulfate chain elongation

Following the attachment of the first α-1,4-GlcNAc, HS polymerase (GlcA/GlcNAc transferase) catalyzes the HS chain polymerization by the alternating addition of β-1,4-GlcA and α-1,4-GlcNAc. The five type II transmembrane proteins of the EXT (exostosin) protein family, namely EXT1 and EXT2, and three EXT-like proteins EXTL1, 2, and 3, are the enzymes responsible for the polymerization of HS chains (Nadanaka and Kitagawa, 2008). The five proteins are of 330-919 a.a. in length, and share 25-30% overall homology and 30-50% homology in the conserved C-terminal region (Esko and Selleck, 2002).

As mentioned above, mutations in the genes encoding EXT1 and EXT2 result in hereditary multiple exostoses (HME), thus these two proteins are considered as tumor suppresors (Lind et al., 1998). Data from independent research groups on individual expression of EXT1 and EXT2 in cell lines defective in HS biosynthesis (murine *sog9,* gro2C), in COS-1 and COS-7 cells, and in yeast *Pichia pastoris*, indicate that each protein displays GlcAT and GlcNAcT activities *in vitro* and is localized to the ER (Lind et al., 1998; McCormick et al., 1998; Kobayashi et al., 2000; McCormick et al., 2000; Senay et al., 2000). EXT1 appears to have significantly higher GlcAT and GlcNAcT activities compared to EXT2, which has relatively low activity. However, EXT1 and EXT2 form homo- and heterooligomers when co-expressed. The EXT1/EXT2 heterocomplex has substantially higher transferase activities than either EXT1 or EXT2 alone, and is localized to the Golgi apparatus, where the HS synthesis takes place. Mixing EXT1 and EXT2 that were separately expressed and purified, did not result in co-immunoprecipitation nor enhanced enzyme activity, indicating that *in vivo* interaction is necessary to productively form a functional GlcAT/GlcNAcT complex.

Characterization of the enzymatic activities of recombinant soluble EXT1, EXT2 and the EXT1/EXT2 complex, reveals that EXT1 and the EXT1/EXT2 complex are able to polymerize HS chains *in vitro*, as shown by the elongation of exogenous *N*-acetylheparosan oligosaccharide substrates in the presence of UDP-GlcA and UDP-GlcNAc (Busse and Kusche-Gullberg, 2003). Interestingly, EXT1 and the EXT1/EXT2 complex generate reaction products with different characteristics. EXT1 adds up to ~10 sugar units onto 10-mer and 11-mer oligosaccharide substrates with a GlcA or a GlcNAc at the non-reducing end, respectively, resulting in elongated oligomers of various sizes (both even- and odd-number DP, e.g. DP 12, 13, 14, and so on) and with either a GlcA or a GlcNAc at the non-reducing end. In contrast, the EXT1/EXT2 complex produces elongated oligomers of even-numbered DP (DP 12, 14, 16, 18) that have a GlcA at the non-reducing end, suggesting that the EXT1/EXT2 complex has more efficient GlcAT rather than GlcNAcT activity. EXT2, on the other hand, added only a single GlcA onto an 11-mer *N*-acetylheparosan substrates with a GlcNAc at the non-reducing end. A further investigation demonstrated that the GlcNAcT activity of EXT1 alone was 10-fold higher than the GlcAT activity *in vitro*, based on the lower K_m values for UDP-GlcNAc and *N*-acetylheparosan oligosaccharide substrates in the GlcNAcT reaction compared to the GlcAT reaction (Wei et al., 2000).

EXTL1, 2, and 3 are not associated with hereditary multiple exostoses (HMEs), and do not have GlcAT activity (Kitagawa et al., 1999; Kim et al., 2001). EXTL2 is an α-1,4-*N*-acetylhexosaminyltransferase (Kitagawa et al., 1999). It is capable of transferring α-GlcNAc onto GlcAβ1-3Galβ1-*O*-naphthalenemethanol, which contains the non-reducing end disaccharide of HS linkage region, suggesting that EXTL2 functions as GlcAT-II that initiates the elongation of HS chains. However, EXTL2 also transfers α-GalNAc from UDP-GalNAc

onto HS linkage region tetrasaccharide-Ser, *N*-acetylchondrosine (GlcAβ1-3GalNAc), and GlcAβ1-3Galβ1-*O*-naphthalenemethanol, suggesting that it may also be able to put an α-GalNAc cap on the HS linkage region and prevent any further elongation. Crystal structures have been resolved of murine EXTL2 alone, as well as in complex with UDP-GlcA, with UDP-GalNAc, and with both UDP and the acceptor analog GlcAβ1-3Galβ1-*O*-naphthalenemethanol (Negishi et al., 2003; Pedersen et al., 2003). These analyses reveal three highly conserved amino acid residues (Asn_{243}, Asp_{246}, Arg_{293}) at the active site, and suggest a retaining S_Ni-like mechanism for the α-1,4-*N*-acetylhexosaminyltransferase reaction catalyzed by EXTL2.

In contrast to EXTL2, EXTL3 has α-GlcNAcT activity towards both *N*-acetylheparosan oligosaccharides and GlcAβ1-3Galβ1-*O*-C_2H_4NH-benzyloxycarbonyl (Cbz), indicating that EXTL3 can function in both HS chain initiation and elongation (Kim et al., 2001). On the other hand, EXTL1 is likely to function in the HS chain elongation, since it has α-GlcNAcT activity only towards *N*-acetylheparosan oligosaccharides.

The catalytic activities of the EXT proteins are clearly overlapping to some extent, and thus the exact *in vivo* function of each protein remains to be determined. All human EXT protein family members are expressed ubiquitously, except for EXTL1, which is expressed only in certain tissues, such as skeletal muscle, brain and heart (Sugahara and Kitagawa, 2002). This suggests that EXTL1 may be involved in synthesizing unique HS structures with tissue-specific functions. It has been shown in HEK293 cells that silencing using RNA interference of human EXT1 or EXT2, results in the synthesis of shorter HS chains, while overexpression of EXT1 resulted in increased HS chain length, which was even more pronounced when both EXT1 and EXT2 were overexpressed (Busse et al., 2007). These data are in agreement with the enzymatic function of EXT1/EXT2 complex in HS chain elongation. In contrast, HEK293 cells

transformed with RNAi constructs of EXTL3 synthesize longer HS chains. In this case, it is possible that a decrease in EXTL3, which can function in both HS chain initiation and elongation, inhibits HS chain initiation and results in a decreased number of HS chains that are polymerized more extensively. Overexpression of EXT2 or EXTL3 alone, however, did not seem to have significant effect on the chain length. It has also been reported that cell lines defective in EXT1 were still able to produce small amounts of HS (Okada et al., 2010). The silencing of EXT2 or EXTL2 expression, but not of EXTL3, further reduced HS synthesis in these cell lines. It was suggested that, in the absence of EXT1, EXT2 and EXTL2 cooperate to polymerize HS chains.

Heparan sulfate chain modification

The first modification to occur on HS chains is the removal of an acetyl group from GlcNAc residues (forming $GlcNH_2$), followed by substitution of the free amino group with sulfate (forming GlcNS) obtained from the cofactor 3'-phosphosadenyl-5'-phosphosulfate (PAPS) (Esko and Selleck, 2002). This is a key event since it is a prerequisite for all other HS modifications. The two-step reaction is catalyzed by GlcNAc *N*-deacetylase/*N*-sulfotransferases (NDSTs), a bifunctional enzyme for which four isoforms, NDSTs 1 through 4, have been identified in vertebrates (Kusche-Gullberg and Kjellen, 2003). The NDSTs have the *N*-deacetylase and *N*-sulfotransferase domains located at the N- and C-termini, respectively (Kakuta et al., 1999; Duncan et al., 2006). The action of NDST appears to be processive, as indicated by the presence of stretches of *N*-sulfated GlcNAc (NS domains) and of non-modified GlcNAc (NA domains) within HS chains (Esko and Lindahl, 2001). It has been recently shown with NDST1 and NDST2, that in the absence of PAPS the enzymes remove the acetyl groups

from fewer GlcNAc residues at random positions in the HS chains (Carlsson et al., 2008). However, in the presence of PAPS, NDSTs perform both *N*-deacetylation and *N*-sulfation of consecutive GlcNAc residues, thus generating NS domains. The degree and pattern of modifications of the NS domains change with different concentrations of PAPS. The expression level of EXT1 and EXT2 has also been shown to affect NDST1 expression in HEK293 cells (Presto et al., 2008). Intriguingly, while overexpression of EXT2 increases the level of NDST1 expression, NDST1 N-glycosylation, and HS sulfation, overexpression of EXT1 gives opposite results. Analyses of the cell line overexpressing both EXT2 and NDST1 indicated that the two proteins interact with each other and can be co-immunoprecipitated, leading to a speculation that NDST1 and EXT1 compete for binding to EXT2.

Following the action of NDSTs, glucuronyl C5-epimerase (Epi) catalyzes the conversion of some of the GlcA in HS chains into IdoA, and uronosyl-2-*O*-sulfotransferase (2-OST) catalyzes the 2-*O*-sulfation of some of the GlcA and IdoA residues. Each protein is encoded by a single gene, which has been cloned and heterologously expressed (Crawford et al., 2001; Li et al., 2001; Rong et al., 2001). Epi and 2-OST co-localize with the medial-Golgi marker α-mannosidase II in normal cells (Pinhal et al., 2001). However, both proteins are localized to the ER when Epi is fused to an ER retention signal (p33). These results, supported by co-immunoprecipitation data, indicate a protein complex formation between Epi and 2-OST, which may contribute to the efficiency of the consecutive actions of both proteins. The epimerization reaction catalyzed by Epi is reversible in *in vitro* soluble systems, but appears to be irreversible *in vivo* (Hagner-McWhirter et al., 2004). Epi epimerizes GlcA residues in GlcNS-GlcA disaccharide units, and does not react with *O*-sulfated GlcA/IdoA, nor with GlcA/IdoA surrounded by *O*-sulfated glucosamine (Esko and Selleck, 2002). 2-OST, on the other hand, has

marked substrate preference for IdoA rather than GlcA residues (Smeds et al., 2010). The resolved crystal structure of chicken 2-OST in complex with PAPS showed that 2-OST forms a homotrimeric complex and revealed a.a. residues that appear critical for IdoA versus GlcA substrate preference (Bethea et al., 2008).

The last two modifications of HS chains are the *O*-sulfation at the C-6 and C-3 positions of the glucosamine residues, catalyzed by 6-*O*-sulfotransferase (6-OST) and 3-*O*-sulfotransferase (3-OST), respectively (Whitelock and Iozzo, 2005). Three 6-OST isoforms sharing 50-57% homology have been identified, with different expression patterns and substrate preferences (Habuchi et al., 2000). Transcript expression of 6-OST1 is observed predominantly in the liver, and that of 6-OST2 mainly in the brain and spleen, while 6-OST3 expression appears to be ubiquitous. With regards to substrate specificity, 6-OST1 prefers to act on IdoA-GlcNS disaccharide units, while 6-OST2 and 6-OST3 can transfer sulfate groups onto both IdoA-GlcNS and GlcA-GlcNS, although the substrate preference of 6-OST2 appears to be concentration-dependent. Human 6-OST2 has also been found to exist in two splice variant forms, designated h6-OST2 and h6-OST2S, respectively, which also differ in expression patterns and substrate specificities (Habuchi et al., 2003). All three 6-OST isoforms have been shown to localize to the *trans*-Golgi stacks, and to form homooligomers mediated, at least in part, by their stem regions (Nagai et al., 2004). Multiple 3-OST isoforms, comprising 3-OST1, 2, 3_A, 3_B, 4, and 5, have also been identified and characterized to have different expression patterns and substrate specificities (Kusche-Gullberg and Kjellen, 2003; Mochizuki et al., 2008). All 3-OST isoforms are type II transmembrane protein, except for 3-OST1 that does not have transmembrane domain but is still localized in the Golgi, suggesting that it may engage in an interaction/protein complex with other Golgi resident proteins.

PROTEIN N-GLYCOSYLATION:

ROLE OF THE OLIGOSACCHARYLTRANSFERASE COMPLEX

Overview of the eukaryotic N-glycan biosynthetic pathway

N-glycosylation is an important and common protein modification, whereby a glycan is attached co-translationally via a β-glycosylamide linkage to an Asn residue in a polypeptide chain within the sequon Asn-X-Ser/Thr (Weerapana and Imperiali, 2006). Previously thought to occur only in eukaryotes, it is now established that protein N-glycosylation also occurs in Bacteria and Archaea (Messner, 2004; Eichler and Adams, 2005; Abu-Qarn et al., 2008). Protein N-glycosylation serves multiple purposes intra- and extracellularly. It is essential for normal growth and development, as exemplified by congenital disorders of glycosylation (CDG), a group of human diseases caused by defects in N-glycosylation (Stanley et al., 2008). In plants, mutations in the N-glycan biosynthetic genes have been shown to affect cell wall composition, adaptive responses to osmotic stress, salt tolerance, fertility, and viability of embryos (reviewed in Pattison and Amtmann, 2009). Intracellularly, protein N-glycosylation plays a crucial role in protein folding, oligomerization, quality control, sorting and transport in the secretory pathway (Helenius and Aebi, 2001, 2004). It also contributes to protein solubility and protection against proteases, and extracellularly, to cell-cell interaction and adhesion.

In eukaryotes, the N-glycosylation biosynthetic pathway is compartmentalized in the ER and Golgi apparatus (Figure 2.8) (Helenius and Aebi, 2004; Weerapana and Imperiali, 2006; Pattison and Amtmann, 2009). In the cytosol and ER, the process starts with assembly of the 14-saccharide N-glycan core structure ($Glc_3Man_9GlcNAc_2$) on a dolichyl-phosphate (Dol-P) precursor, catalyzed in a step-wise manner by a series of highly specific glycosyltransferases (Burda and Aebi, 1999; Burda et al., 1999). Sequential transfers onto Dol-P of two GlcNAc and

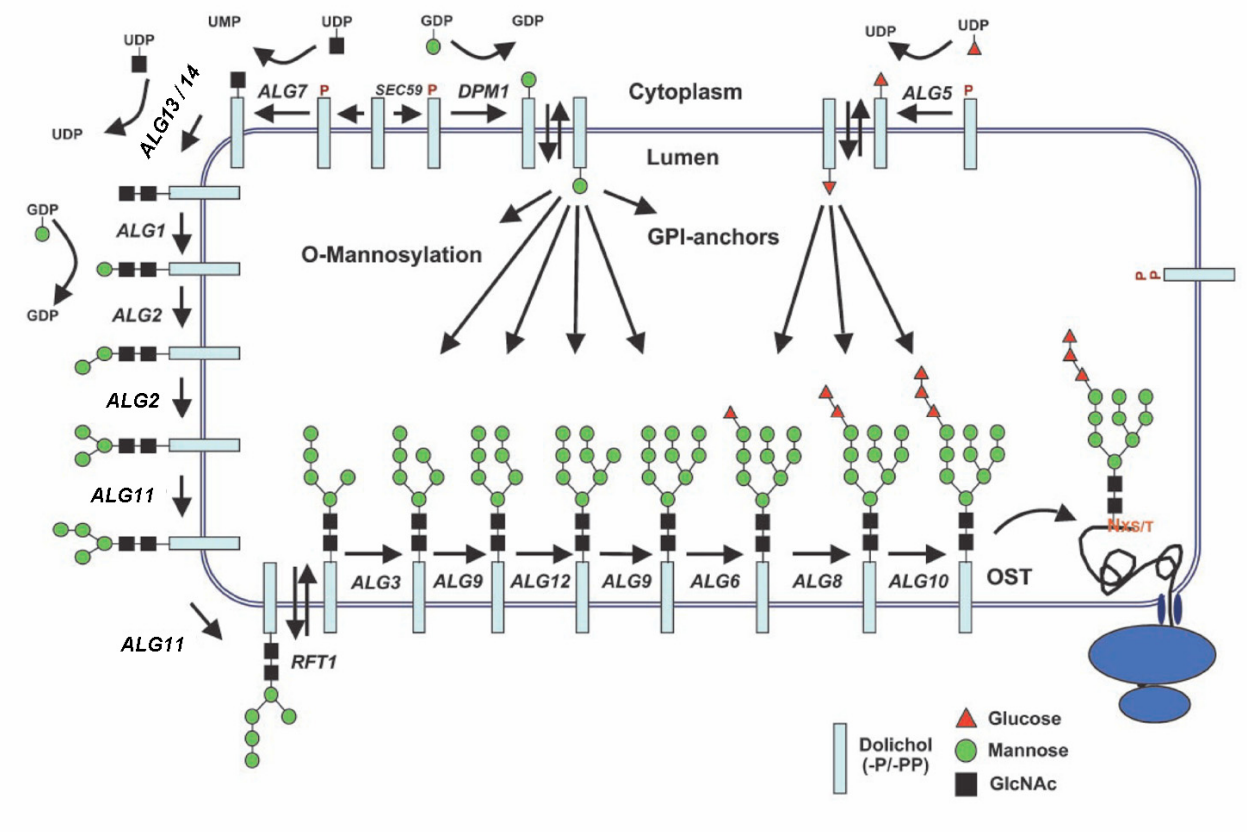


Figure 2.8. Biosynthesis of core N-glycan and its subsequent transfer to a nascent polypeptide chain in the ER.

The yeast biosynthetic enzymes involved in each step are indicated. (Adapted with modification from Helenius and Aebi, 2004).

five Man residues from the nucleotide-sugar precursors UDP-GlcNAc and GDP-Man, respectively, occur on the cytoplasmic side of the ER membrane. The $Man_5GlcNAc_2$-PP-Dol is re-oriented / flipped to the luminal side of the ER membrane, a process thought to be facilitated by RFT1, an ATP-dependent flippase. Assembly of the N-glycan core structure continues in the ER lumen by the addition of four Man and three Glc residues, using Dol-P-Man and Dol-P-Glc as sugar donors. All the enzymes involved in the N-glycan core structure assembly (the dolichol pathway) have been identified, the last ones being two bifunctional enzymes Alg2, which catalyzes sequential α-1,3- and α-1,6-mannosylations, and Alg11, which catalyzes two sequential α-1,2-mannosylations (Figure 2.8) (O'Reilly et al., 2006; Kampf et al., 2009).

The completed $Glc_3Man_9GlcNAc_2$ core oligosaccharide is transferred by oligosaccharyltransferase, a multi-subunit enzyme complex, from the lipid carrier onto a growing polypeptide chain as it is being translated (Helenius and Aebi, 2001, 2004; Weerapana and Imperiali, 2006; Pattison and Amtmann, 2009). The removal of the terminal two Glc residues by glucosidases I and II enters the nascent glycoprotein into the Calnexin-Calreticulin cycle, an ER chaperone system that facilitates protein folding. In this system, calnexin and calreticulin bind to the newly synthesized glycoprotein and promote its proper folding. Upon release from calnexin or calreticulin, the remaining Glc residue is trimmed by glucosidase II. UDP-Glc:glycoprotein glucosyltransferase (GT) recognizes glycoproteins that are mis-folded at this point, and adds the one Glc molecule back onto the $Man_9GlcNAc_2$, re-entering the glycoprotein back into the Calnexin-Calreticulin cycle. The cycle is repeated until the protein is properly folded, at which time an ER-resident mannosidase cleaves one Man residue. The resulting $Man_8GlcNAc_2$-containing glycoprotein is then transported into the Golgi, where further trimming and re-

glycosylation take place by the action of several glycosidases and glycosyltransferases, leading to the synthesis of diverse and elaborate glycan structures.

The ER-portion of the N-glycosylation pathway is well-conserved in eukaryotes, and thus, the oligomannosidic structures of N-glycans are common to yeast, animals, and plants (Weerapana and Imperiali, 2006; Pattison and Amtmann, 2009). However, the maturation steps to form the high-mannose, complex, and hybrid N-glycans in the Golgi apparatus are somewhat different between these different classes of organisms. For examples, N-glycans that are highly branched and mannosylated are present in yeast. Plant N-glycans are typically smaller than animal N-glycans and contain no sialic acid but rather contain β1,2-xylose and/or core α1,3-fucose (Dirnberger et al., 2002). Animal N-glycans can be very complex and elaborate, and may contain sugar residues such as GalNAc and sialic acid (Weerapana and Imperiali, 2006).

Oligosaccharyltransferase is a multi-subunit protein complex

As mentioned above, oligosaccharyltransferase (OST) catalyzes the critical step of transferring the N-glycan core structure *en bloc* from the lipid-linked precursor to a polypeptide. OST is an integral membrane enzyme complex containing multiple protein subunits (Knauer and Lehle, 1999; Kelleher and Gilmore, 2006). In eukaryotes, the OST complex has been isolated and most extensively characterized in yeast (*Saccharomyces cerevisiae*) and vertebrates. Nine OST subunits have been identified in yeast (Table 2.3). Five of the yeast subunits, Wbp1, Swp1, Stt3, Ost1, and Ost2, are essential for viability, and thus are considered as the main components of OST. The other four subunits, Osts 4 through 6, are not essential for viability, but are required for maximal catalytic activity of OST. Eight OST subunits have also been identified in vertebrates, and so far three subunits have been identified in plants (Table 2.3.)

Table 2.3. The subunits of eukaryotic OST complexes in yeast, vertebrates, and plants.

([a]) The yeast and vertebrate OST subunits have been identified biochemically as parts of the OST complexes in *Saccharomyces cerevisiae* and in mammals, respectively. ([b]) The plant OST subunits listed are the subunit homologs for which *Arabidopsis thaliana* mutant lines have been described in published literature. Biochemical evidence to support the association of these plant subunits in an OST complex remains to be demonstrated.

The molecular mass of each subunit, if already reported in literature, is in brackets following the subunit name. Yeast subunits that are essential for viability, and their homologs in vertebrates and plants, are highlighted in green. STT: Staurosporine and Temperature Sensitive; DAD1: Defender Against apoptotic Death 1; DGL1: Defective Glycosylation 1.

Yeast[a]	Vertebrates[a]	Plant[b]
Ost1 (64/62 kDa)	Ribophorin I (66 kDa)	
Ost2 (16 kDa)	DAD1 (12 kDa)	AtDAD1 (12.7 kDa) AtDAD2 (12.7 kDa)
Ost3 (34 kDa)	N33 (32-33 kDa)	
Ost4 (3.6 kDa)		
Ost5 (9.5 kDa)		
Ost6 (32 kDa)	IAP	
Wbp1 (47 kDa)	OST48 (48 kDa)	DGL1
Swp1 (30 kDa)	Ribophorin II (63/64 kDa)	
Stt3 (~60 kDa)	STT3A STT3B	STT3A STT3B

(Kelleher and Gilmore, 2006; Pattison and Amtmann, 2009). Two small proteins, namely DC2 and KCP2 (keratinocyte-associated protein 2), were also identified as potential subunits of mammalian OST (Shibatani et al., 2005). DC2 has a low level of identity to a portion of yeast Ost3 and Ost6. Two isoforms of STT3, designated A and B, are present in vertebrates and plants, and two isoforms of DAD, the Ost2 homolog, have also been found in plants (Danon et al., 2004). All these subunits are integral membrane proteins, with the numbers of transmembrane spanning domains varying from one in Ost1, Wbp1, and Ost4, to thirteen in Stt3 (Figure 2.9) (Kelleher and Gilmore, 2006).

Oligosaccharyltransferase isoforms have different functions and/or specificity

Much experimental data demonstrate the presence of more than one form of OST complex. Glycerol gradient and ion exchange chromatography of a digitonin extract from canine microsomes, indicates the existence of two OST isoforms in vertebrates (Kelleher et al., 2003). Both isoforms contain mammalian OST subunits ribophorin I and II, OST48, and DAD1, and either STT3A or STT3B. The STT3B-containing OST complex was found to have a significantly higher V_{max} than the STT3A-containing complex. In addition, the STT3A-containing OST complex showed greater preference for fully-synthesized N-glycan core oligosaccharides (i.e. $Glc_3Man_9GlcNAc_2$), while the STT3B-containing complex did not seem to discriminate between the complete versus assembly-intermediate oligosaccharide substrates.

Analyses of yeast OST using the split-ubiquitin system and blue native gel electrophoresis (BN-PAGE) also identified two OST isoforms; one containing Ost3, and the other containing Ost6 (Schwarz et al., 2005; Yan and Lennarz, 2005). In addition to Ost3 or Ost6, both isoforms contain the remaining six yeast OST subunits, thus forming two distinct

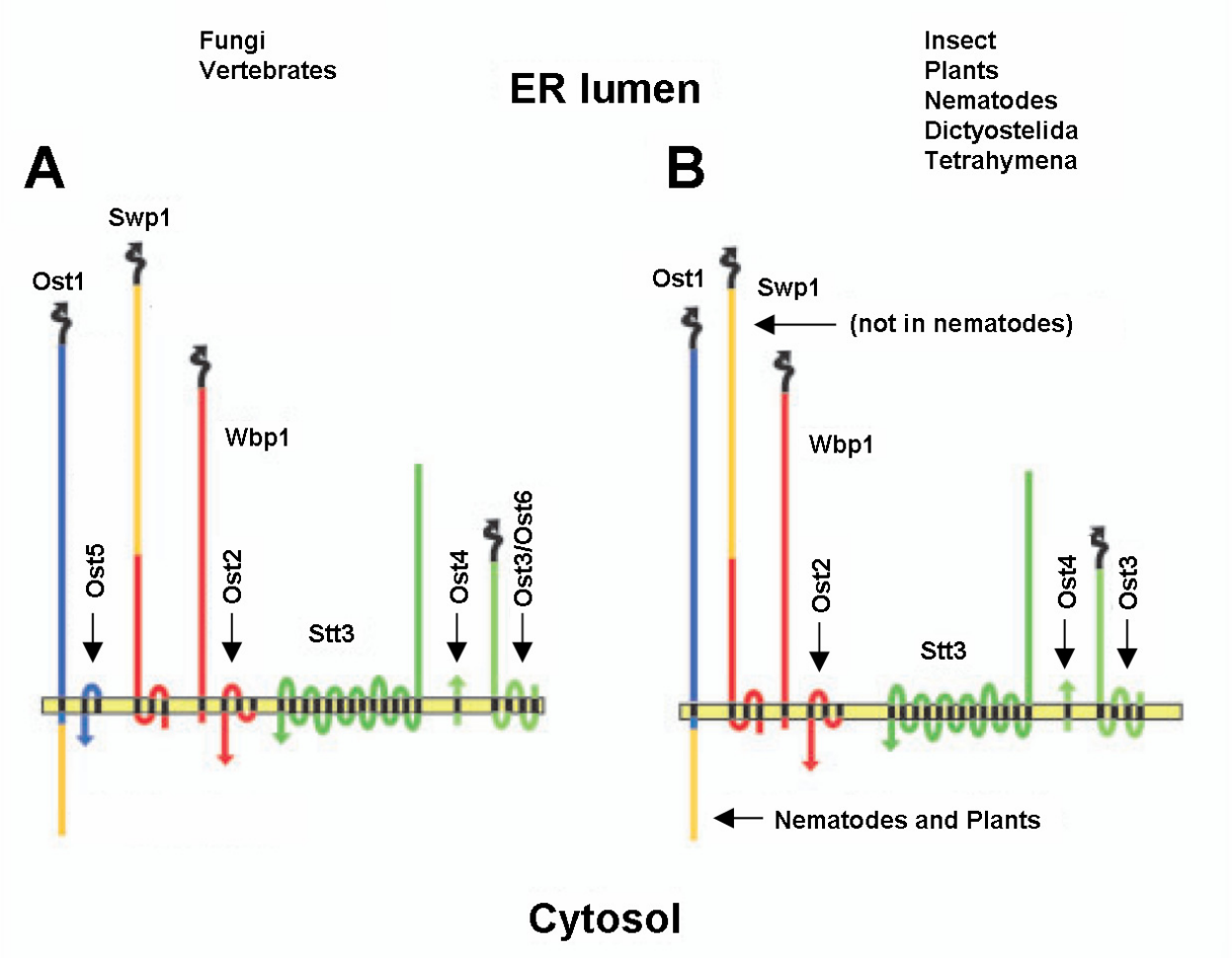


Figure 2.9. Membrane topology of the oligosaccharyltransferase (OST) subunits.

(A) OSTs in fungi and vertebrates consist (or are predicted to consist) of seven or eight subunits.

(B) OSTs in insect, plants, nematodes, dictyostelida, and tetrahymena are predicted to consist of seven subunits.

(Adapted from Kelleher and Gilmore, 2006).

eight-subunit complexes. These findings were also supported by data from pair-wise split ubiquitin experiments, in which Ost3 and Ost6 showed identical patterns of interaction with other OST subunits, and by data that Ost6 co-immunoprecipitated with all other OST subunits except Ost3 (Yan and Lennarz, 2005; Yan et al., 2005). The OST complexes resolve at ~500 kDa upon separation by BN-PAGE. Further analysis of the OST complexes from yeast mutant strains Δost4, Δost3, Δost6, and the double mutant Δost3Δost6, demonstrated that in the absence of Ost4, there is no incorporation of either Ost3 or Ost6 into the OST complexes (Spirig et al., 2005). These data suggest a role for Ost4 as an assembly factor to integrate Ost3 and Ost6 into the OST complexes. It is also in agreement with previous experiments showing that Ost3 did not co-immunoprecipitate with any OST subunits, and that Stt3 co-immunoprecipitated with all OST subunits except Ost3, in extracts from the Δost4 mutant (Karaoglu et al., 1997). Another experiment also showed that cross-linking between Stt3 and Ost3 could only occur in the presence of Ost4 (Kim et al., 2003).

The Ost3- and Ost6-containing OST complexes interact with two distinct but highly similar translocon trimeric complexes, i.e. Sec61 and Ssh1, respectively (Yan and Lennarz, 2005). More specifically, Ost3- and Ost6 containing complexes interact with Sbh1 and Sbh2, respectively, which are the β-subunits of the translocon complexes. OST (presumably Ost3-containing OST) and Sec61 have previously been shown to associate with each other via multiple protein-protein interactions (Chavan et al., 2005). Interestingly, mammalian STT3A-containing OST has been observed to resolve in several high molecular mass complexes (~500, ~600, and ~700 kDa, respectively), which correspond to the OST, OST associated with Sec61 translocon complex, and OST associated with both the Sec61 translocon complex and a translocon-related tetrameric protein complex named TRAP, respectively (Shibatani et al., 2005).

Furthermore, the yeast OST, Sec61 complex, and ribosomes were shown to co-sediment and co-purify in experiments using sucrose gradient centrifugation and affinity purification with Concanavalin A beads, respectively (Harada et al., 2009). These results indicate that yeast OST, Sec61 complex, and ribosomes, interact directly with each other and form a ternary complex, thereby coupling together three processes, i.e. protein translation, translocation, and N-glycosylation. It remains to be determined whether yeast Ost6-containing OST and the Ssh1 translocon complex also engage in a similar ternary complex.

The different OST isoforms in yeast and vertebrates apparently have specific functions with regards to protein N-glycosylation. Schulz and Aebi used mass spectrometry analysis to compare the site-specific N-glycosylation occupancy of glycoproteins extracted using endoglycosidase H from yeast cell wall of wild-type and of mutant strains defective in either Ost3 or Ost6 (Schulz and Aebi, 2009). They observed that the absence of Ost3 or Ost6 resulted in OST isoform-specific partial underglycosylation of certain N-glycosylation sites, suggesting that one OST isofom can N-glycosylate specific sequons more efficiently than the other isoform, and vice versa. By using the RNAi technique to silence the expression of STT3A or STT3B, the catalytic functions of the STT3A- and STT3B-containing mammalian OST were individually observed with regards to the N-glycosylation of several different proteins (Ruiz-Canada et al., 2009). The results indicate that STT3A-containing OST functions to co-translationally N-glycosylate a nascent polypeptide as it is being translocated into the ER lumen. The STT3B-containing OST, but not the STT3A-containing isoform, appears to efficiently N-glycosylate a sequon located near to the signal sequence cleavage site of a secreted protein (procathepsin C). STT3B-containing OST also functions to post-translationally N-glycosylate a sequon located at the C-terminal portion of a polypeptide. These observations in yeast and in vertebrates suggest

that the two OST isoforms (i.e. the Ost3- and Ost6-containing OSTs in yeast, and the STT3A- and STT3B-containing OSTs in vertebrates) may work cooperatively to ensure a highly efficient N-glycosylation process.

Both yeast Ost3 and Ost6 are predicted to have a thioredoxin-like fold, which is located in the ER luminal domains of the proteins and contain a CxxC active-site motif (Fetrow et al., 2001). Studies of Ost6 showed that this protein exhibits a high reducing capacity, is present in the more stable, oxidized form in the ER, is capable of reducing a disulfide bond in a peptide substrate, and has reductase activity towards insulin (Schulz et al., 2009). Analyses of the crystal structures of Ost6 in both the oxidized and reduced states, indicate that reduction of Ost6 causes a loop located near the CxxC active site to become flexible, and thus, induces a conformational change in the active site. It was proposed that reduced Cys residues in the nascent polypeptide might perform a nucleophilic attack on the disulfide bond(s) in Ost6. This may result in intermolecular disulfide bond formation between Ost6 and the nascent polypeptide, which may serve to secure the nascent polypeptide to the OST complex to facilitate the N-glycosylation process. The strong reductase activity of Ost6 may subsequently act on the intermolecular disulfide bond, thereby releasing the now N-glycosylated polypeptide. Since Ost3 and Ost6 are highly homologous, the same reductase activity may also be present in Ost3.

Functions of other OST subunits

Multiple lines of evidence support Stt3 and its orthologs as the active subunit of the OST complexes. The above data, showing that STT3A- and STT3B-containing mammalian OSTs have different oligosaccharide substrate preference and reaction kinetics, demonstrate that the STT3s contribute significantly to catalysis, and suggest that the STT3 subunits contain the active

sites of the OST complexes (Kelleher et al., 2003; Kelleher and Gilmore, 2006). It has also been shown that STT3A, but not the other subunits, can be photo-cross-linked with acceptor peptides bearing a photoreactive Lys derivative adjacent to a cryptic N-glycosylation site (Nilsson et al., 2003). Using this photo-cross-linking technique, the interaction between the active site of STT3 and the nascent polypeptide chains has been studied in detail (Karamyshev et al., 2005). It was recently reported that the trypanosome *Leishmania major* has only four homologs of STT3, and no homologs of any other yeast OST subunits (Nasab et al., 2008; Hese et al., 2009, and references therein). The four STT3 homologs, however, are able and sufficient to N-glycosylate polypeptides, although with a more lax oligosaccharide selectivity. Three of the four homologs could complement a deletion of the yeast Stt3 gene and the encoded proteins became incorporated into the yeast OST complex. Microsomal preparations from *L. major* were able to transfer non-glucosylated ($Man_{7\text{-}9}GlcNAc_2$) and glucosylated ($Glc_{1\text{-}3}Man_9GlcNAc_2$) glycans with the same efficiency *in vitro* (Castro et al., 2006). When the *L. major* STT3s were heterologously expressed in yeast and became incorporated into the yeast OST complex, however, the enzymes highly prefer the completely-synthesized $Glc_3Man_9GlcNAc_2$ oligosaccharides as substrates. These observations indicate that the oligosaccharide substrate specificity/preference is dictated by multiple components of the OST complex, and not solely by the catalytic subunit STT3.

The function of ribophorin I, the vertebrate homolog of yeast Ost1, has recently been investigated. A cross-linking experiment showed that ribophorin I associates with a subset of membrane proteins (Wilson et al., 2005). Depletion of ribophorin I by RNAi significantly decreases N-glycosylation of type II, and of certain type I, transmembrane proteins, but has no effect on the N-glycosylation of proteins with multiple transmembrane spanning domains or of secretory proteins, suggesting that the function of ribophorin I is substrate-specific (Wilson and

High, 2007; Wilson et al., 2008). RNAi of ribophorin I also prevented cross-linking from occurring between a ribophorin I-dependent protein and the STT3A subunit. It has therefore been suggested that ribophorin I functions to facilitate access of certain integral membrane proteins to the catalytic site of OST.

Apart from Stt3/STT3A/STT3B, Ost1/ribophorin I, Ost3, Ost6, and Ost4, which have been discussed above, little information is available on the functions of the other OST subunits (Kelleher and Gilmore, 2006; Lennarz, 2007). The yeast Wbp1 subunit and its interacting partner Ost2 are thought to be involved in recognition of the lipid-linked oligosaccharide precursors. The Swp1p subunit has been shown to interact with Ost1, and thus is suggested to assist Ost1 in its function to deliver nascent polypeptide to the active site of OST. The function of Ost5 is at present unknown. The ternary structure of yeast OST has been resolved at 12 Å resolution using cryo-electron microscopy (cryo-EM), in which the positions of several subunits relative to each other were determined and a molecular mechanism of OST was proposed (Figure 2.10) (Li et al., 2008). Analysis of the cryo-EM micrograph of yeast OST revealed a two-fold symmetry, suggesting a dimeric configuration (Chavan et al., 2006). This is in agreement with the ~500 kDa size of OST observed by blue native gel electrophoresis, since the sum of molecular masses of eight subunits forming the OST complex is ~250 kDa.

CONCLUSIONS AND RELEVANCE

Polysaccharides are secondary gene products. Polysaccharide biosynthesis is governed primarily by the coordinated actions of an array of glycosyltransferases and, where applicable, modifying enzymes. Research in the past 30 years has moved us forward significantly in

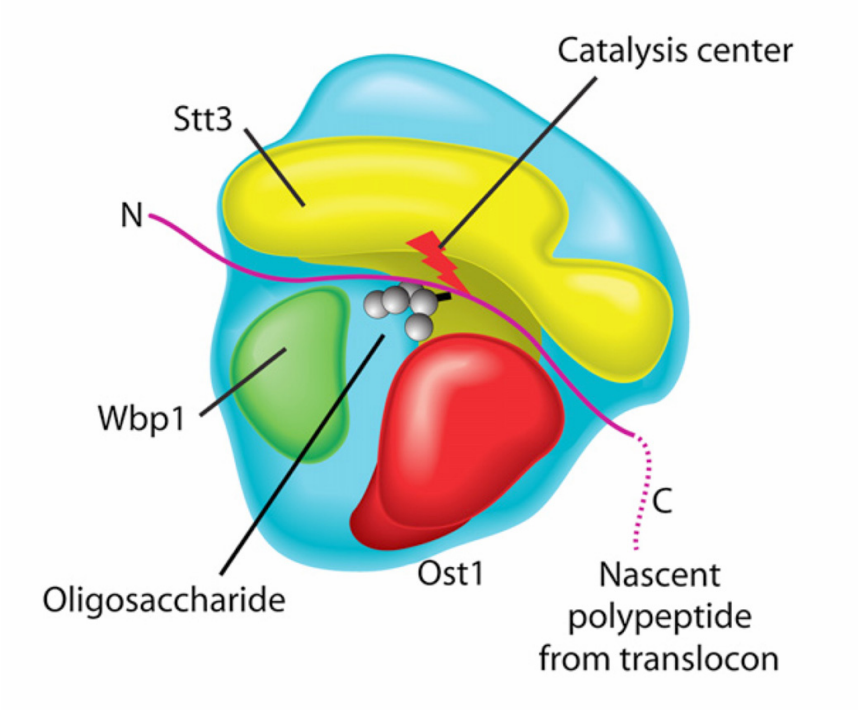


Figure 2.10. A model depicting the relative positions of three subunits of yeast OST and the molecular mechanism of N-glycan transfer ("peptide-threading mechanism") proposed by Li and co-workers (Li et al., 2008).

The model shows the top (luminal) view of the OST complex, and the positions of Stt3 (yellow), Wbp1 (green), and Ost1 (red), which are based on the structure of yeast OST resolved at 12 Å resolution by cryo-electron microscopy. Not depicted in the model is the Swp1 subunit, which is located at the bottom (cytosolic) side. In the proposed mechanism, Wbp1 subunit recognizes and binds the oligosaccharide (represented by the grey spheres). Ost1 subunit scans the nascent polypeptide (pink thread) for the Asn-X-Ser/Thr sequon as it threads through the groove between the Stt3 subunit on one side, and Wbp1 and Ost1 subunits on the other side. Upon detection of a sequon, the oligosaccharide bound to Wbp1 is transferred to the Asn residue of the polypeptide, the process of which is catalyzed by the Stt3 subunit. (Adapted from Li et al., 2008).

understanding these processes, yet it is clear that our knowledge of the biosyntheses of most, if not all, polysaccharides is far from complete.

The literature reviewed in this chapter illustrates the complexity and diversity of biosynthetic mechanisms that give rise to the numerous types of polysaccharide structures. First, in eukaryotes, polysaccharides are synthesized at different locations within a cell, including the plasma membrane, the compartments of the secretory pathway, and, in the case of starch biosynthesis in higher plants, the plastids. The different locations of synthesis influence the corresponding biosynthetic machineries. For example, polysaccharide syntheses in the secretory pathway proceeds in a "conveyor belt" manner with the synthesizing and modifying enzymes positioned strategically along the pathway, allowing for the production of structurally complex polysaccharides containing various glycosyl residues and modifications. On the other hand, the polysaccharide syntheses located at the plasma membrane, such as those of cellulose and hyaluronan, appear to produce simpler structures and to be coupled with some kind of polysaccharide extrusion mechanism. Furthermore, syntheses of different polysaccharides may occur via formation of one linkage by one glycosyltransferase, via dual-activity enzymes catalyzing two different reactions, and/or via *en bloc* transfer of pre-synthesized oligosaccharides onto growing polymer chains.

Not one of the five polysaccharide syntheses discussed in this chapter is identical to another, yet there are many similarities between them. For example, most of the endomembrane enzymes have type II transmembrane topology, while those at the plasma membrane are membrane-bound proteins with multiple membrane-spanning domains. Furthermore, a variety of polysaccharide structures with different biological functions can be found within the same class of polysaccharides, and their biosyntheses are clearly regulated very specifically to cater to the

needs of the different cell/tissue types and developmental stages. One way to achieve this is by producing enzymes with multiple isoforms, each with distinct specificity, distinct products, and/or partial redundancy, which are differentially expressed both spatially and temporally. In addition, protein complexes can be formed, which contain different combinations of enzymes and/or isoforms, and thus contribute to the structural diversity of the synthesized polysaccharides. As illustrated in this chapter, the involvement of protein complexes appears to be a common theme in polysaccharide biosynthetic pathways.

This dissertation presents the identification of the GAUT1:GAUT7 HG:GalAT complex, which is the first pectin biosynthetic enzyme complex to be described. This is a major step in our quest to understand the biosynthesis of pectin, arguably the most structurally complex polysaccharide in nature. To put our data in perspective and to make the next move in our research, it is important to consider the mechanisms and protein complex involvements in the syntheses of other polysaccharides that have been elucidated, which are summarized in this chapter. At the same time, however, we need to keep an open mind about the real possibility that pectin synthesis may involve novel mechanisms that are different from those previously described. In this regard, it is noteworthy that an in-depth phylogenetic study of the GT8 family, the family to which GAUT1 and GAUT7 belong, has indicated that the GAUT genes belong to the GT8 class I (Yin et al., 2010). Class I is largely a plant specific clade, with some metazoan members for which the functions have not been definitively identified. The GAUTs, along with plant GATLs and GATRs, were likely acquired from an unicell cyanobacteria progenitor, and likely to have acquired plant specific functions. Therefore, the precise functions, involvement in protein complexes, and post-translational modifications of the GAUT family members will need to be empirically determined.

CHAPTER 3

GAUT1 AND GAUT7 ARE THE CORE OF A PLANT CELL WALL PECTIN BIOSYNTHETIC HOMOGALACTURONAN:GALACTURONOSYLTRANSFERASE COMPLEX[1]

[1] Melani A. Atmodjo, Yumiko Sakuragi, Xiang Zhu, Amy J. Burrell, Sushree S. Mohanty, James A. Atwood III, Ron Orlando, Henrik V. Scheller, and Debra Mohnen. To be re-submitted by invitation to the Proceedings of the National Academy of Sciences.

ABSTRACT

Plant cell wall pectic polysaccharides are arguably the most complex carbohydrates in nature. Progress in understanding pectin synthesis has been slow due to its complex structure and difficulties in purifying and expressing the low abundance, Golgi-membrane bound pectin biosynthetic enzymes. *Arabidopsis* GAUT1 was previously identified as an α-1,4-galacturonosyltransferase that synthesizes homogalacturonan, the most abundant pectic polysaccharide [Sterling JD*, et al.* (2006) *Proc Natl Acad Sci U S A* 103(13):5236-5241]. We now show that GAUT1 functions in a protein complex with the homologous GAUT7. Surprisingly, although both GAUT1 and GAUT7 are type II membrane proteins with single N-terminal transmembrane spanning domains, the N-terminal region of GAUT1, including the transmembrane domain is cleaved *in vivo*. This raises the question of how the processed GAUT1 is retained in the Golgi, the site of homogalacturonan biosynthesis. We show that the anchoring of GAUT1 in the Golgi requires association with GAUT7 to form the GAUT1:GAUT7 complex. Proteomics analyses also identified 12 additional proteins that interact with the GAUT1:GAUT7 complex. This study provides conclusive evidence that the GAUT1:GAUT7 complex is the catalytic core of a homogalacturonan:galacturonosyltransferase complex and that cell wall matrix polysaccharide biosynthesis occurs via protein complexes. The processing of GAUT1 to remove its N-terminal transmembrane domain and its anchoring in the Golgi by association with GAUT7 provides a paradigm for how the large number of plant cell wall biosynthetic glycosyltransferases and polysaccharide-modifying proteins can assemble into protein complexes to drive the synthesis of the complex and developmentally and environmentally plastic plant cell wall.

INTRODUCTION

Pectin is the most structurally complex cell wall polysaccharide in plants, requiring at least 67 transferases for synthesis (Mohnen, 2008; Caffall and Mohnen, 2009). It comprises ~35% of the primary cell wall in dicots, non-graminaceous monocots, and gymnosperms, and 2-10% in grasses (Mohnen, 2008). Pectin is a family of polysaccharides including homogalacturonan (HG), rhamnogalacturonan I (RG-I), rhamnogalacturonan II (RG-II), and xylogalacturonan, which are defined by the presence of α-1,4-linked D-galactopyranosyluronic acid (GalA) residues (Mohnen, 2008; Caffall and Mohnen, 2009). Pectins have multiple functions in plant growth, development, and disease resistance including roles in cell:cell adhesion, wall porosity, cell elongation, and wall extensibility (Willats et al., 2001; O'Neill et al., 2004; Ezaki et al., 2005; Derbyshire et al., 2007). Pectins form the matrix in which the load-bearing cellulose/hemicellulose network is embedded, provide structural support in primary walls, influence secondary wall formation in fibers and woody tissues, and are a reservoir of oligosaccharide signaling molecules (Caffall and Mohnen, 2009; Singh et al., 2009; Brutus et al., 2010; Lionetti et al., 2010). The gelling and stabilizing properties of pectin are exploited for food enhancement and industrial purposes, and pectin has multiple health benefits including lowering cholesterol and serum glucose levels, inhibiting cancer growth and metastasis, and prebiotic function in the gut (Behall and Reiser, 1986; Thakur et al., 1997; Manderson et al., 2005; Jackson et al., 2007). An understanding of how pectin is synthesized is required to manipulate and exploit the diverse pectin structures for industrial, agronomic and biomedical uses (Mohnen et al., 2008).

HG is the defining and most abundant pectic domain, constituting ~55-70% of pectin. HG is a linear homopolymer of α-1,4-linked GalA that is partially methylesterified at C-6 and *O*-

acetylated at the O-2/O-3, modifications that affect pectin structure and function (Willats et al., 2001). α-1,4-Galacturonosyltransferase (GalAT, E.C. 2.4.1.43) synthesizes HG by catalyzing transfer of GalA from uridine-diphosphate-GalA (UDP-GalA) onto HG acceptors (Doong and Mohnen, 1998). HG:GAlactUronosylTransferase 1 (GAUT1) was identified by tandem mass spectrometry of *Arabidopsis* solubilized membrane preparations enriched for GalAT activity; transient expression in HEK293 cells; and immunoabsorption of GalAT activity from SP-Sepharose-purified *Arabidopsis* solubilized membrane fractions (SP fraction) using anti-GAUT1 antibodies (Sterling et al., 2006). *GAUT1* encodes a protein of 673 amino acids (a.a.) and predicted molecular mass of 77.4 kDa, pI of 9.95, and type II transmembrane topology, consistent with the Golgi localization of GAUT1 and HG:GalAT activity (Sterling et al., 2001; Dunkley et al., 2006). GAUT1 belongs to glycosyltransferase family 8 and the *Arabidopsis* GAUT1-related gene superfamily consisting of 15 GAUT and 10 GAUT-like (GATL) genes (Sterling et al., 2006; Caffall et al., 2009; Cantarel et al., 2009; Yin et al., 2010).

GAUT7, another GAUT family member, was the only predicted glycosyltransferase that co-purified with GAUT1 in GalAT-enriched detergent-solubilized *Arabidopsis* membrane proteins from suspension culture cells (Sterling et al., 2006). Despite 60% sequence similarity (36% sequence identity) to GAUT1, GAUT7 had no HG:GalAT activity when transiently expressed in HEK293 cells (Sterling et al., 2006), raising the question of whether GAUT7 functions in a biosynthetic complex with GAUT1. Here we show that *Arabidopsis* GAUT1 and GAUT7 form a GAUT1:GAUT7 HG:GalAT complex and that GAUT7 anchors a proteolytically processed form of GAUT1 in the Golgi.

RESULTS AND DISCUSSION

GAUT7 associates with GAUT1 in an HG:GalAT complex

Polyclonal antibodies against a.a. positions 82-103 of *Arabidopsis* GAUT7 predicted stem region recognized a broad doublet band of ~75 kDa, confirmed to be GAUT7 by mass spectrometry (MS) sequencing (Figure 3.S1, B and D). Immunoabsorption of GAUT7 from *Arabidopsis* SP fraction using the anti-GAUT7 antibodies caused an antibody-dependent depletion of GalAT activity from the supernatant and recovery of GalAT activity in the anti-GAUT7-immunoabsorbed pellet (Figure 3.S1). The results suggested that GAUT7 was either a GalAT or part of a GalAT complex.

To test if GAUT7 exists in a GalAT complex with GAUT1, independent anti-GAUT1 and anti-GAUT7 immunoabsorbed-proteins were separated by SDS-PAGE and analyzed by immunoblotting using anti-GAUT1 and anti-GAUT7-specific sera. *Arabidopsis* GAUT1 and GAUT7 co-immunoprecipitated from the SP fraction (Figures 3.1A and 3.S1), demonstrating biochemically that GAUT1 and GAUT7 exist in a protein complex.

Bimolecular fluorescence complementation (BiFC) (Hu et al., 2002) was used to corroborate the above finding and identify *in vivo* interactions between GAUT1 and GAUT7. Fluorescence complementation with characteristic Golgi signal morphology was observed when GAUT1 and GAUT7, fused to complementary split halves of YFP, were transiently co-expressed in tobacco (*Nicotiana benthamiana*) leaves (Figure 3.1, B and C), verifying that GAUT1 and GAUT7 interact in the Golgi *in vivo*. Negative controls (ARAD1-Yn/GAUT1-Yc, ARAD1-Yn/GAUT7-Yc) did not fluoresce while positive ARAD1 (ARABINAN DEFICIENT 1) control (ARAD1-Yn/ARAD1-Yc) did (Figure 3.1, F-H), as expected for this putative arabinosyltransferase (Harholt et al., 2006) that forms homodimers and localizes to Golgi (J.

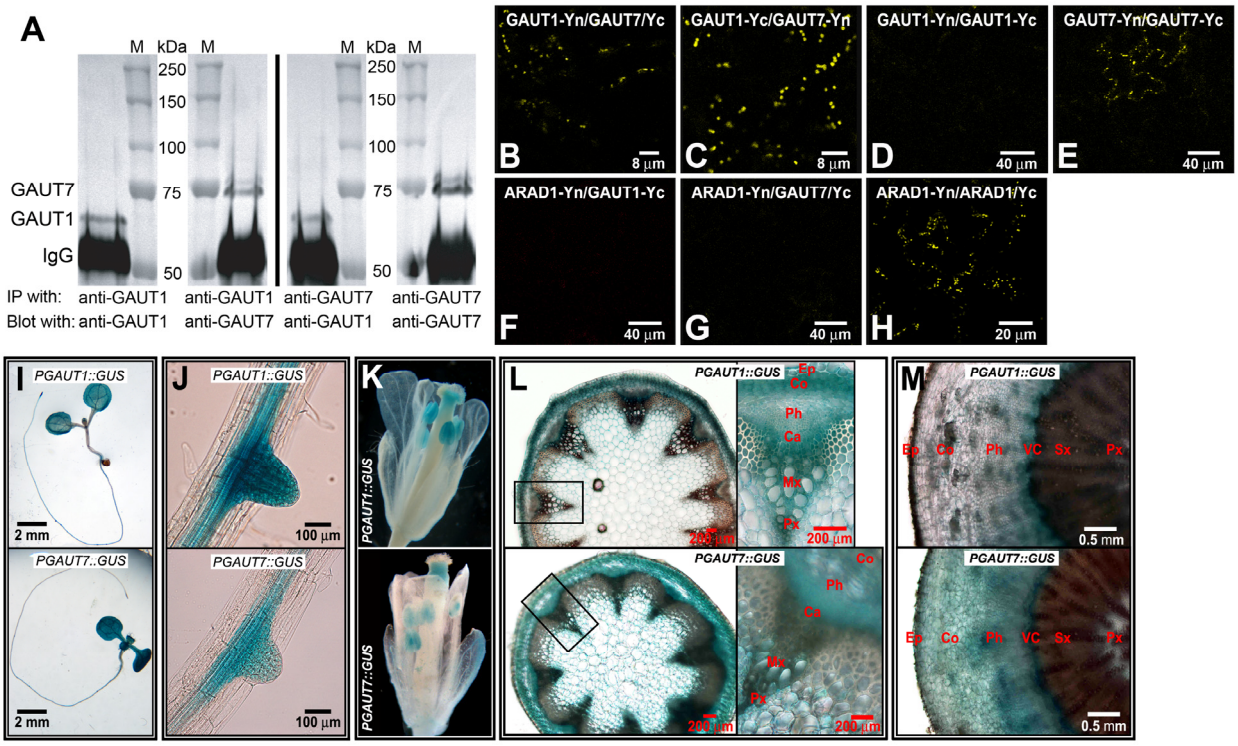


Figure 3.1. GAUT1 and GAUT7 interact in a protein complex.

(A) Co-immunoprecipitation of GAUT1 and GAUT7. Anti-GAUT1 and anti-GAUT7 antibody-immunoabsorbed proteins from *Arabidopsis* SP fraction were separated by SDS-PAGE and immunoblotted with anti-GAUT1 or anti-GAUT7 sera. The experiment was done thrice with similar results. M: molecular mass protein marker. IP: immunoprecipitation. IgG: IgG heavy chain detected by secondary antibody.

(B to H) BiFC analysis of GAUT1 and GAUT7. Constructs were transiently co-expressed in *N. benthamiana* leaves, with BiFC of ARAD1 as control. YFP signals are in yellow. Individual expression of each construct gave no YFP signals. Results were verified by three independent experiments (except ARAD1 negative control which was done twice).

(I to M) *Arabidopsis GAUT1* and *GAUT7* promoter:*GUS* construct expression in whole (I) and developing lateral roots (J) of 7-day-old seedlings, and in flowers (K), stems (L), and tap roots (M) of mature 6-8 week plants. Similar results were observed in multiple T2 plants from at least

5 independent T1 lines. Ep-epidermis, Co-cortex; Ph-phloem, Ca-cambium, mx-metaxylem, px-protoxylem (panel L) or primary xylem (panel M), VC-vascular cambium, Sx-secondary xylem.

Harholt, C.N. Søgaard, Y. Sakuragi, H.V. Scheller, manuscript submitted). No signal was observed upon co-expression of GAUT1-Yn and GAUT1-Yc (Figure 3.1D). Co-expression of GAUT7-Yn and GAUT7-Yc gave variable fluorescence (Figure 3.1E), leaving the possibility that GAUT7 may dimerize *in vivo*. The biochemical and BiFC results provide *in vitro* and *in vivo* confirmation that GAUT1 and GAUT7 exist in a protein complex. Protein complexes have been reported in the biosynthesis of glycoconjugates (de Graffenried and Bertozzi, 2004) (e.g. N-linked glycoproteins, glycolipids and proteoglycans) and in plant starch and cellulose syntheses (Hannah and James, 2008; Taylor, 2008). However, confirmed function of a protein complex in plant cell wall matrix polysaccharide synthesis has not been previously reported.

To determine if *GAUT1* and *GAUT7* are co-expressed in *Arabidopsis*, we analyzed *Arabidopsis* gene expression databases and *GAUT1* and *GAUT7* promoter:*GUS* construct expression in transgenic *Arabidopsis*. Microarray data (https://www.genevestigator.com/) show similar expression patterns of *GAUT1* and *GAUT7* in all plant tissues, with Pearson correlation coefficients (*r*-values) of 0.600 and 0.684 from Gene Co-expression Analysis Toolbox (http://genecat.mpg.de) and Arabidopsis Co-expression Analysis Tool (http://www.cressexpress.org/), respectively (Table 3.S3). Promoter:*GUS* fusions (Figure 3.1, I-M) indicate high expression of both genes in meristematic regions, vascular tissues and reproductive organs, and support a role for both proteins in primary and secondary wall syntheses. Strong *GAUT1* and *GAUT7* co-expression was observed in seedling cotyledon tips and vasculature, leaf primordia and outer edges of young leaves, and in 6-8 week-old plant cambium, phloem, epidermis, cortex, stem metaxylem and protoxylem, anthers, pollen, floral distal stigma, and root vascular cambium and phloem. The extensive co-expression of *GAUT1*

and *GAUT7* in dividing cells and vascular tissues is consistent with their function in a protein complex.

The immunoprecipitated GAUT1:GAUT7 complex (Figures 3.1 and 3.S1) transfers GalA from UDP-GalA onto HG oligosaccharide (oligogalacturonide (OGA)) acceptors, similar to the reported GalAT activity in solubilized membrane preparations (Doong and Mohnen, 1998). To test if the GAUT1:GAUT7 complex transfers GalA onto RG-I and/or RG-II acceptors, acceptor substrate specificity was examined by comparing OGAs of degrees of polymerization (DP) 7-23; RG-I backbone oligomers of DP 6-26 with either a rhamnosyl residue (RG-I-R) or a GalA residue (RG-I-G) at the non-reducing end; and RG-II monomer. Figure 3.2A shows that the immunoprecipitated GAUT1:GAUT7 complex, as well as GalAT activity in the *Arabidopsis* SP fraction, is highly selective for OGAs, verifying the GAUT1:GAUT7 complex as an HG:GalAT.

GAUT1:GAUT7 complex is held together by disulfide bond(s)

Western analysis of GAUT1 and GAUT7 separated by non-reducing SDS-PAGE confirmed the GAUT1:GAUT7 protein complex and indicated disulfide bonding between GAUT1 and GAUT7. GAUT1 and GAUT7 from *Arabidopsis* SP fraction migrate as monomeric forms when separated by reducing SDS-PAGE (~60 kDa for GAUT1; ~75 kDa for GAUT7) (Figure 3.2B). However, both proteins electrophoresed at ~185 kDa when separated by non-reducing SDS-PAGE (Figure 3.2B). The similar size of GAUT1 and GAUT7 immunoreactive bands upon non-reducing SDS-PAGE suggests that GAUT1 and GAUT7 are held together by intermolecular disulfide bond(s). GAUT1 and GAUT7 each contain 8 cysteine residues and one or more of these is proposed to function in disulfide bond formation (see below).

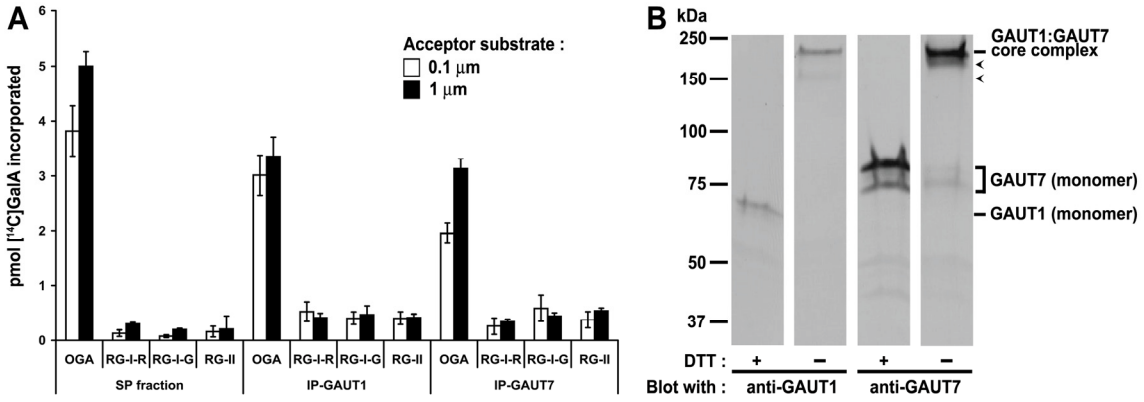


Figure 3.2. GAUT1:GAUT7 GalAT complex is selective for HG substrate and contains non-reducing SDS-PAGE resistant bonds.

(A) GalAT activities of *Arabidopsis* SP fraction and of anti-GAUT1- and anti-GAUT7-immunoprecipitated GAUT1:GAUT7 complex were tested at 0.1 and 1 μM of the pectic acceptors: OGA DP 7-23; RG-I oligomers DP 6-26 with either rhamnose (RG-I-R) or GalA (RG-I-G) at the non-reducing ends; and RG-II monomer. Data are mean ± SD (n = 3).

(B) GAUT1 and GAUT7 resolve at higher masses in non-reducing SDS-PAGE. *Arabidopsis* SP fraction separated by SDS-PAGE in the presence or absence of 25 mM DTT, and analyzed by immunoblotting with anti-GAUT1 or anti-GAUT7 sera. Protein bands common to GAUT1 and GAUT7 on the immunoblots are estimated to be ~185 kDa (noted as GAUT1:GAUT7 core complex). Arrowheads indicate additional GAUT1 or GAUT7 high molecular weight protein bands.

Repetitive high stringency proteomics was used to characterize the non-reducing SDS-PAGE-stable GAUT1:GAUT7 complex (outlined in Figure 3.S4). The complex was immunoprecipitated four times; twice with antigen-purified anti-GAUT1 specific antibodies and twice with antigen-purified anti-GAUT7 specific antibodies, both covalently attached to magnetic beads. After stringent washing (see Methods), the immunoprecipitants were eluted from the beads and resolved by non-reducing SDS-PAGE. The ~185 kDa GAUT1:GAUT7 complex was excised from the gel (Figure 3.S5-A), in-gel trypsin digested, and subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS). GAUT1 and GAUT7 were the only two proteins identified by LC-MS/MS in each independent anti-GAUT1- and anti-GAUT7-specific-IgG immunoprecipitant (IP-GAUT1 and IP-GAUT7, respectively; Table 3.S1). Neither GAUT1 nor GAUT7 were detected in the preimmune IgG immunoprecipitation control (IP-Control). The results establish that the non-reducing SDS-PAGE-stable GAUT1:GAUT7 complex consists only of GAUT1 and GAUT7.

GAUT1 is post-translationally cleaved

GAUT1 is predicted to encode a 77.4 kDa protein, yet always resolved at ~60 kDa ((Sterling et al., 2006); Figures 3.1A, 3.2B, and 3.S1). We proposed that the discrepancy between the predicted and observed size of GAUT1 was due to post-translational proteolytic processing of GAUT1 *in planta*. To test this we used three independent anti-GAUT1 antibodies, reactive against GAUT1 a.a. positions 132-154, 341-365, and 448-472, respectively ((Sterling et al., 2006) and this paper). While all three antibodies detected recombinant GAUT1 expressed in HEK293 cells (Sterling et al., 2006), only the antibodies generated against a.a. 341-365 and 448-

472 recognized GAUT1 in immunoblots of *Arabidopsis* SP fraction (Figure 3.S2), leading us to hypothesize that *Arabidopsis* GAUT1 is proteolytically-cleaved *in vivo* in the N-terminal region.

N-terminal sequencing of GAUT1 excised from a reducing SDS-PAGE-blotted membrane yielded the peptide sequence RANELVQ, which indicated a cleavage between Met_{167} and Arg_{168} and would yield a processed *Arabidopsis* GAUT1 of 58.6 kDa and pI of 9.3, consistent with the observed characteristics of GAUT1 in the SP fraction. The cleavage site was also supported by LC-MS/MS analyses of immunoprecipitated GAUT1:GAUT7 complex which gave no GAUT1 peptide sequence N-terminal to Ala_{169} (Figure 3.S5, C and D). Cleavage of GAUT1 *in planta* could activate enzyme activity, as has been described for *Arabidopsis* and tobacco type-I pectin methylesterases (Wolf et al., 2009) or may facilitate specific association of GAUT1 with GAUT7 in the GAUT1:GAUT7 complex. Further studies are required to establish the function(s) of GAUT1 processing.

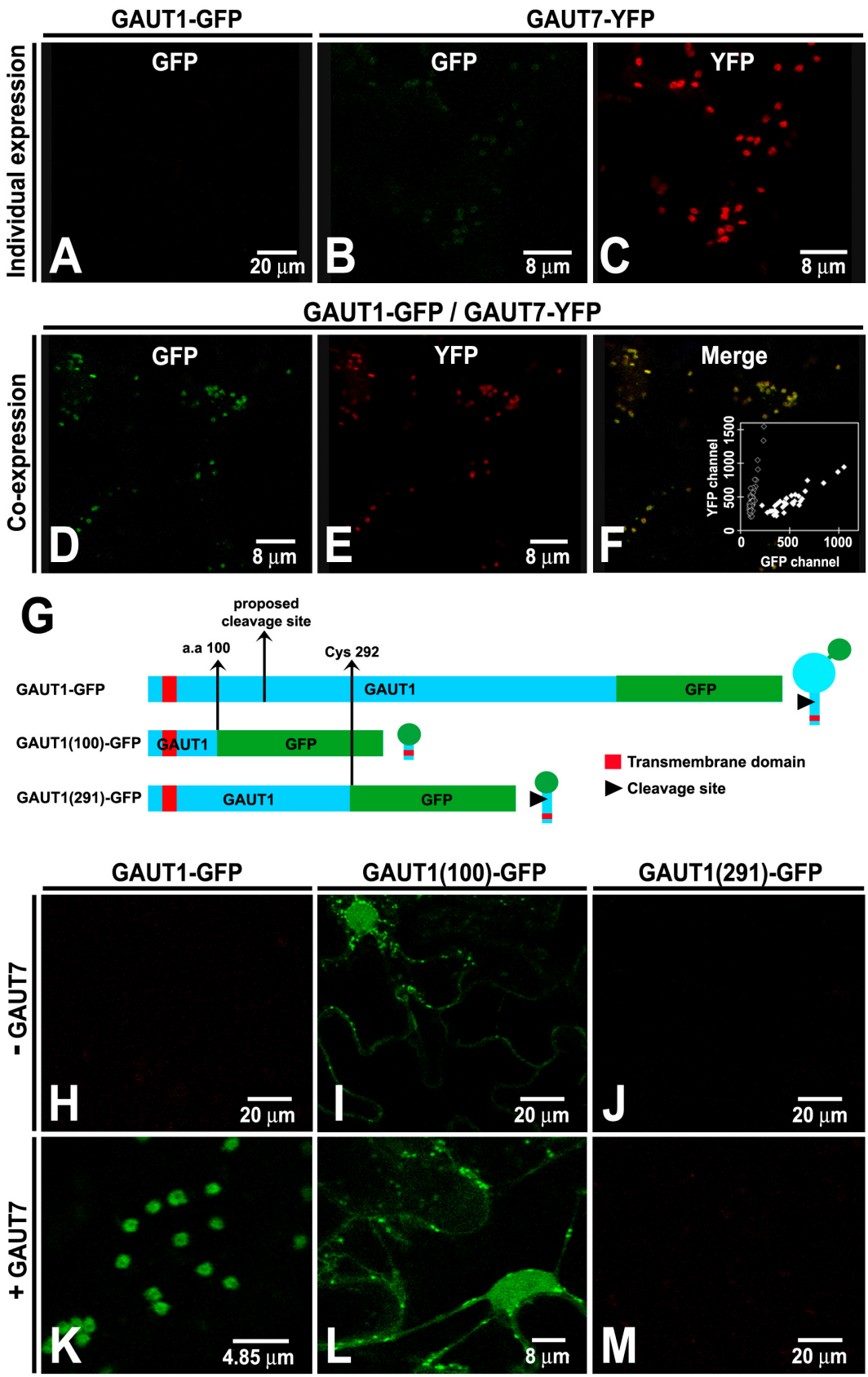
GAUT1 is dependent on GAUT7 for retention in the Golgi

As a consequence of a cleavage between Met_{167} and Arg_{168}, GAUT1 would be devoid of its transmembrane domain (TMD) and secreted out of the cell unless a tethering mechanism retained it in the Golgi. We hypothesized that GAUT1 was held in the Golgi by protein-protein interactions with GAUT7. To test this, GAUT1-GFP and GAUT7-YFP constructs were individually- and co-expressed in tobacco leaves. When individually expressed, GAUT1-GFP did not yield a Golgi signal (Figure 3.3A) while GAUT7-YFP did (Figure 3.3C). However, when co-expressed, GAUT1-GFP accumulation in Golgi was observed which overlapped with GAUT7-YFP Golgi signal (Figure 3.3, D-F). Similar co-localization of GAUT1 and GAUT7 in Golgi was observed when GAUT1-GFP was co-expressed with non-tagged GAUT7 (Figure 3.3,

Figure 3.3. Golgi retention of the cleaved GAUT1 relies on presence of GAUT7.

(A to F) Transient individual expression (A-C) and co-expression (B-F) of GAUT1-GFP and GAUT7-YFP in *N. benthamiana* leaves. GFP and YFP signals were detected by sequential scanning. Insert (panel F) plots mean pixel intensity from each Golgi (detected in the GFP versus the YFP channels), showing Golgi signals in GAUT7-YFP individual expression experiments (◇) and in the GAUT1-GFP/GAUT7-YFP co-expression experiments (◆). The insert demonstrates that GFP signals detected upon GAUT1-GFP/GAUT7-YFP co-expression (D) are due to the Golgi accumulation of GAUT1-GFP in the presence of GAUT7-YFP, and not to background signal from GAUT7-YFP (B). Results were verified in at least three independent experiments.

(G to M) Transient expression of C-terminally truncated GAUT1-GFP fusion constructs (G) in absence (H-J) or presence (K-M) of GAUT7. GAUT1-GFP: full length GAUT1 fused to GFP; GAUT1(100)-GFP: first 100 a.a. of GAUT1 fused to GFP; GAUT1(291)-GFP: first 291 a.a. of GAUT1 fused to GFP.


GAUT1-GFP
GAUT7-YFP
Individual expression
GFP
GFP
YFP
A
B
C
20 μm
8 μm
8 μm
GAUT1-GFP / GAUT7-YFP
Co-expression
GFP
YFP
Merge
D
E
F
8 μm
8 μm
YFP channel
GFP channel
0
500
1000
1500
G
proposed cleavage site
a.a 100
Cys 292
GAUT1-GFP
GAUT1(100)-GFP
GAUT1(291)-GFP
GAUT1
GFP
Transmembrane domain
Cleavage site
GAUT1-GFP
GAUT1(100)-GFP
GAUT1(291)-GFP
- GAUT7
+ GAUT7
H
I
J
K
L
M
20 μm
20 μm
20 μm
4.85 μm
8 μm
20 μm

H and K). These results demonstrate that retention of GAUT1 in Golgi requires the presence of GAUT7 and suggest that GAUT7 acts as a membrane-anchor for GAUT1.

To further explore the mechanism of GAUT1 tethering, three GAUT1-GFP fusion constructs (Figure 3.3G) were transiently expressed in tobacco, with and without co-expressed non-tagged GAUT7. The constructs were: 1) GAUT1-GFP (GAUT1 full length), 2) GAUT1(100)-GFP containing only the first 100 a.a. of GAUT1 including the predicted TMD but not the proposed cleavage site, and 3) GAUT1(291)-GFP containing the first 291 a.a. of GAUT1 including both the predicted TMD and the cleavage site but not the first cysteine residue beyond the predicted TMD (Cys_{292}). Transient expression of GAUT1-GFP yielded GAUT7-dependent fluorescence accumulation in Golgi (Figure 3.3, H and K; Figure 3.S3). In contrast, despite the absence or presence of GAUT7, GAUT1(100)-GFP yielded a broad localization pattern, including endoplasmic reticulum (ER) and Golgi (Figure 3.3, I and L), while GAUT1(291)-GFP accumulated no observable signal (Figure 3.3, J and M; Figure 3.S3). GAUT1(100)-GFP seems sufficient to target the protein to the secretory pathway but does not provide Golgi-specific localization. GAUT1(291)-GFP, however, likely underwent proteolytic cleavage, releasing the C-terminally-located GFP from the membrane and resulting in secretion into the apoplast where GFP generates weak or no fluorescence due to the low pH environment (Zheng et al., 2004). The results indicate a cleavage site between a.a. 100-291, consistent with the proposed cleavage site between Met_{167} and Arg_{168}, and underscore the importance of disulfide bonding between GAUT1 and GAUT7 to retain GAUT1 in the Golgi. The results also indicate a region downstream of a.a. 291 in GAUT1 as necessary for protein-protein interactions between GAUT1 and GAUT7.

Proteomics analyses identify GAUT1:GAUT7 core complex putative interacting proteins

We reasoned that the ~185 kDa disulfide-bonded GAUT1:GAUT7 core complex may associate non-covalently with additional proteins to form a larger pectin synthesis complex. The estimated size of ~590 kDa for the detergent-solubilized polygalacturonic acid (PGA) synthase (synonymous with HG:GalAT) from azuki bean is consistent with a large pectin synthesis complex (Ohashi et al., 2007). The GAUT1:GAUT7 HG:GalAT core complex could, for example, interact stably or transiently with methyltransferases to form methylesterified HG, a common form of HG in plant walls.

To test for possible GAUT1:GAUT7 core complex-associating proteins, we immunoprecipitated the GAUT1:GAUT7 complex a total of six times; three times using anti-GAUT1 (IP-GAUT1) and three times using anti-GAUT7 (IP-GAUT7) antigen-purified antibodies covalently bound to magnetic beads. The immunoprecipitants were stringently washed prior to elution from the beads to reduce non-specific ionic, hydrophobic, and detergent micellar interactions. Eluted immunoprecipitants were separated by reducing SDS-PAGE and analyzed by LC-MS/MS (Figure 3.S4). Since preliminary data showed immunoprecipitated proteins from ~50-125 kDa, we focused on these for the proteomic analyses (Figure 3.S5-B).

Ten proteins, including GAUT1 and GAUT7, were identified consistently in each IP-GAUT1 and IP-GAUT7 immunoprecipitant from three independently-prepared *Arabidopsis* SP fractions, but not in the IP-Control (using pre-immune IgG-beads), as determined using two high stringency proteomics data filtering methods (i.e. False Discovery Rate and Probability analyses; Table 3.S2). Four additional proteins were >4-fold more abundant in IP-GAUT1 and IP-GAUT7 than in the IP-Control (Table 3.S2). The fourteen proteins, including GAUT1 and GAUT7, represent the GAUT1:GAUT7 core complex and its putative associating proteins (Table 3.1).

While it remains possible that all or some of the fourteen proteins form a large holocomplex, the relatively low Normalized Spectral Abundance Factor (NSAF) values of the twelve additional proteins, compared to those of GAUT1 and GAUT7, suggest that these proteins may transiently interact with the GAUT1:GAUT7 core complex. Several GAUT1:GAUT7 core complex-associating proteins are noteworthy. For example, the two dehydration-responsive proteins (At4g18030, At4g00740) contain a putative methyltransferase domain (domain of unknown function 248, InterPro:IPR004159) and have homology (30-37% a.a. sequence identity, 47-55% similarity) to QUASIMODO2 (At1g78240), a putative homogalacturonan:methyltransferase (HG-MT) (Mouille et al., 2007). HG-MT activity and the two putative HG-MT proteins identified here localize to the Golgi (Goubet and Mohnen, 1999; Dunkley et al., 2006) and are co-expressed with GAUT1 and GAUT7 (Table 3.S3). These results suggest a close relationship between HG synthesis and modification, and support the proposition that methylation of HG occurs as it is synthesized, or immediately thereafter (Mohnen, 2008; Caffall and Mohnen, 2009). The presence of KORRIGAN1 (KOR1/At5g49720), a membrane-bound endo-1,4-β-glucanase implicated in cellulose biosynthesis (Taylor, 2008), as a putative associating protein warrants further investigation. *KOR1* mutants are dwarf, defective in cell elongation, and have reduced wall crystalline cellulose that is compensated by increased pectin (Molhoj et al., 2002). KOR1 localizes to a heterogeneous population of intracellular compartments including Golgi (Taylor, 2008). Perhaps most intriguingly, two homologs of mammalian ribophorins I and II (At1g76400 and At4g21150, respectively) are among the putative interacting proteins. These ribophorins are subunits of oligosaccharyltransferase (OST), an enzyme complex that transfers oligosaccharides *en bloc* from dolichol pyrophosphates onto proteins as they translocate into the ER lumen (Kelleher and Gilmore, 2006). The identification of the ribophorins raises the

Table 3.1. Members of the GAUT1:GAUT7 GalAT core complex and putative associating proteins as revealed by proteomics analyses of the immunoprecipitated complex. Proteomics analyses were performed on three independent sets of immunoprecipitations performed using anti-GAUT1-IgG (IP-GAUT1), anti-GAUT7-IgG (IP-GAUT7), and preimmune IgG (IP-Control). GAUT1 and GAUT7 are the only identified components of the non-reducing SDS-PAGE stable core complex, while fourteen proteins (including GAUT1 and GAUT7) are considered as GAUT1:GAUT7 GalAT complex and its putative associating proteins. Data shown are proteins reproducibly identified in both IP-GAUT1 and IP-GAUT7, and not in IP-Control, based on Probability analysis. The associating proteins also includes those that were >4-fold more abundant in the IP-GAUT1 and IP-GAUT7 compared to IP-Control. Proteins are listed based on their total protein scores in the IP-GAUT1 data sets.

Protein ID	Gene ID	Sequence Name	Protein Length (a.a.)	Protein Weight (kDa)	IP-GAUT1					IP-GAUT7				
					Total Protein Probability	Total Score	Total Spectra	NSAF* Value	Corrected NSAF* Value	Total Protein Probability	Total Score	Total Spectra	NSAF* Value	Corrected NSAF* Value
GAUT1:GAUT7 non-reducing agent stable core complex														
NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1)	673	77.31	1	530.98	19	0.555	0.625	1	256.18	8	0.434	0.583
NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7)	619	69.69	1	362.67	14	0.445	0.375	1	196.85	7	0.417	0.417
GAUT1:GAUT7 complex and putative complex-associating proteins														
NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1)	673	77.31	1	1573.58	122	0.073	0.095	1	2013.38	228	0.047	0.061
NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7)	619	69.69	1	1372.61	167	0.111	0.107	1	1668.63	327	0.072	0.071
NP_176170.1	At1g59610	ADL3 (ARABIDOPSIS DYNAMIN-LIKE 3)	920	100.15	1	973.23	48	0.020	0.021	1	1083.70	64	0.009	0.009
NP_177766.1	At1g76400	ribophorin I family protein	614	68.58	1	780.90	41	0.026	0.026	1	980.38	143	0.032	0.031
NP_172500.1	At1g10290	ADL6 (DYNAMIN-LIKE PROTEIN 6)	914	99.09	1	736.39	36	0.015	0.016	1	1000.20	68	0.010	0.010
NP_193537.2	At4g18030	dehydration-responsive family protein	621	70.27	1	721.96	25	0.016	0.016	1	612.05	42	0.009	0.009
NP_193847.2	At4g21150	ribophorin II (RPN2) family protein	691	74.60	1	641.43	19	0.011	0.011	1	1098.01	65	0.013	0.013
NP_190258.1	At3g46740	TOC75-III (translocon outer membrane complex 75-III)	818	89.12	1	608.83	28	0.014	0.013	1	747.03	49	0.008	0.008
NP_186956.2	At3g03060	ATPase	628	69.51	1	541.41	30	0.019	0.019	1	513.82	58	0.013	0.012
NP_179673.1	At2g20800	NDB4 (NAD(P)H DEHYDROGENASE B4)	582	65.31	1	487.12	17	0.011	0.012	1	1093.47	117	0.027	0.027

NP_565435.1	At2g18330	AAA-type ATPase family protein	636	71.13	1	423.99	36	0.023	0.022	1	765.71	105	0.023	0.022
NP_188616.1	At3g19820	DWF1 (DIMINUTO 1)	561	65.33	1	409.62	15	0.010	0.011	1	901.14	112	0.027	0.027
NP_567184.1	At4g00740	dehydration-responsive protein-related	600	67.49	1	385.40	13	0.009	0.009	1	549.74	36	0.008	0.008
NP_199783.1	At5g49720	KOR1 (KORRIGAN)	621	69.13	1	229.31	9	0.006	0.006	1	243.22	15	0.003	0.003

Total probability: the estimate of the likelihood that the protein assignment is correct based on the peptide probabilities for the identified peptides mapping to the protein.
Total score: the sum of the peptide scores (Mascot Ion Scores) for all peptides matching to a protein.
Total spectra: the sum of all spectra generating peptide identifications to a given protein, i.e. the number of the redundant protein spectra.
***NSAF values**: Normalized Spectral Abundance Factor values are protein normalized spectral counts, calculated as the number of spectral counts for protein X (SpC) divided by the number of amino acid in protein X (L) divided by the sum of SpC/L for all proteins in the experimental data set.
****Corrected NSAF values** were calculated based on the size of the cleaved GAUT1, i.e. 506 a.a. long.

question of whether the biosynthesis of HG, or of pectin in general, occurs not by the classic polymer elongation by sequential action of glycosyltransferases, but rather via *en bloc* transfer of oligosaccharide domains to a growing polysaccharide chain.

Models of GAUT1:GAUT7 HG:GalAT core complex

We propose dimeric and heterotrimeric models for the GAUT1:GAUT7 core complex that is stable in non-reducing SDS-PAGE (Figure 3.4). LC-MS/MS analyses revealed a relative molar abundance of GAUT1 to GAUT7 in the core complex between 1.1 to 1.7, as calculated from the NSAF values (Table 3.S1). The ratios suggest a GAUT1:GAUT7 stoichiometry of 1:1, 2:1, or a mixture of both, in the core complex. The fact that the GAUT1:GAUT7 core complex reproducibly resolved at ~185 kDa in non-reducing SDS-PAGE is more consistent with the 2:1 stoichiometry, which would give a predicted size of 186.9 kDa for a heterotrimeric disulfide-bonded complex of two mature GAUT1 (58.6 kDa) and one GAUT7 (69.7 kDa) subunits. However, the possibility remains that the GAUT1:GAUT7 core complex exists as a mixture of 1:1 and 2:1 stoichiometry *in vivo*. It is noteworthy that in the LC-MS/MS analyses identifying the GAUT1:GAUT7 core complex-associating proteins, the obtained NSAF values suggest a GAUT1:GAUT7 ratio of 1:1, with GAUT7 slightly more abundant than GAUT1. One possible explanation for the increased relative molar abundance of GAUT7 in the larger complex with the associating proteins, is that GAUT7 may dimerize transiently and non-covalently *in vivo*, as suggested by BiFC (Figure 3.1). Such dimerization could facilitate interactions between the GAUT1:GAUT7 core complex and the associating proteins. GAUT1 is the catalytic subunit (Sterling et al., 2006) of the GAUT1:GAUT7 HG:GalAT complex. GAUT7 functions, at least in part, to anchor GAUT1 to the Golgi membrane. Interestingly, it was recently reported that

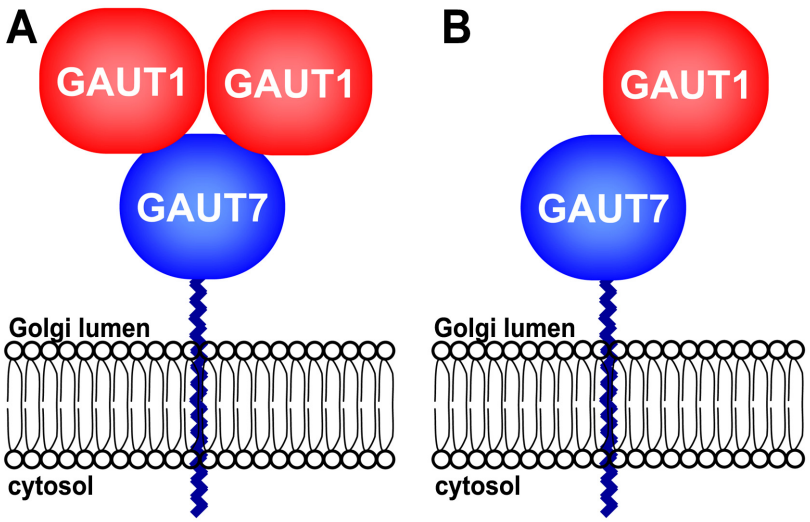


Figure 3.4. Models of the *Arabidopsis* GAUT1:GAUT7 core complex. NSAF values from LC-MS/MS indicate a GAUT1:GAUT7 stoichiometry of 2:1 (A) or 1:1 (B).

GAUT7 carries an amino acid substitution in a proposed catalytic residue that would render it non-catalytic (Yin et al., 2010). This proposition needs experimental verification, but is consistent with a GAUT1-anchoring, non-catalytic role for GAUT7. We cannot exclude, however, that GAUT7 may have an enzymatic function other than the HG:GalAT activity tested previously (Sterling et al., 2006), or that interaction of GAUT7 with GAUT1 may be required for optimal HG:GalAT activity in the GAUT1:GAUT7 complex.

The work reported here indicates that pectin HG synthesis occurs via tethering of a GalAT catalytic subunit (GAUT1) to a Golgi membrane-bound protein anchor (GAUT7). How widespread this phenomenon is in the synthesis of cell wall polysaccharides remains to be determined. It is interesting to speculate, however, that Golgi-tethering proteins, like GAUT7, may play a broader role in promoting the association of glycosyltransferases and polysaccharide-modifying enzymes (such as methyltransferases) into protein complexes to achieve the synthesis of specific wall polysaccharide domains. In this regard, it is noteworthy that GAUT7 homologs are expanded in the grass family (Yin et al., 2010), suggesting a unique role for this class of proteins in grasses and possibly in grass cell wall synthesis. Further studies of the GAUT1:GAUT7 core complex and its associating proteins are expected to increase our understanding of the mechanism of HG and pectin synthesis, and expand our view of plant cell wall synthesis and structure.

METHODS

Immunoprecipitation

Anti-GAUT1 polyclonal antibodies were previously generated (Sterling et al., 2006). Anti-GAUT7 antibodies were generated similarly using multiple antigenic peptides corresponding to GAUT7 a.a. 82-103 (DEVLQKINPVLPKKSDINVGSR). For co-immunoprecipitation of GAUT1 and GAUT7, 30 μl anti-GAUT1 or anti-GAUT7 antibody-conjugated Dynabeads M-280 Sheep Anti-rabbit IgG (as described (Sterling et al., 2006)) were incubated with 500 μl SP fraction (prepared as detailed in *SI Materials and Methods*). Immunoabsorbed materials were washed twice in PBS, three times in Buffer A (50 mM Hepes pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 25% (v/v) glycerol, 1% (v/v) Triton X-100), and once in PBS. PBS washes were done to remove non-specific ionically interacting proteins, and to lower the Triton X-100 concentration below the critical micelle concentration to remove non-specific proteins that may have associated via detergent micelles. Buffer A washes removed non-specific, hydrophobically interacting proteins. Immunoprecipitants were analyzed by SDS-PAGE followed by immunoblotting with anti-GAUT1 and anti-GAUT7 sera (serum dilutions of 1:3,000 and 1:10,000, respectively). Immunoprecipitations for acceptor substrate specificity studies were as above, except that antibody-beads and SP fraction were incubated at 1:1 ratio (v/v).

GalAT activity assay for substrate specificity studies

GalAT activity was measured in 30 μl reactions containing 10 μl enzyme, 50 mM Hepes pH 7.3, 0.2 M sucrose, 0.05% (w/v) BSA, 25 mM KCl, 1 μM or 0.1 μM of oligosaccharide substrate, 6.9 μM UDP-[^{14}C]GalA (specific activity 180.3 mCi/mmol; 1 Ci = 37 GBq) and 1.25 mM $MnCl_2$. Reactions were incubated 30 min at 30°C and products measured using a

precipitation method (Doong and Mohnen, 1998). Acceptor substrates tested were OGA mixture (DP 7-23) (Doong and Mohnen, 1998), RG-I-R and RG-I-G oligomers (DP 6-26) prepared as described (*SI Materials and Methods*), and RG-II monomer from wine (gift from Malcolm O'Neill, CCRC).

BiFC and co-expression of recombinant GAUT1 and GAUT7 fused to fluorescent proteins

Bacterial strains, growth conditions, plasmid constructions, and settings for confocal microscopy are detailed in *SI Materials and Methods*. Transfection of *N. benthamiana* was as described (Vainonen et al., 2008). Overnight cultures of *A. tumefaciens* bearing plasmid constructs were harvested, resuspended in buffer (100 μM acetosyringon, 10 mM $MgCl_2$, 10 mM MES pH 5.6), and incubated at RT for 2 hrs. The optical densities of *Agrobacterium* strains containing ARAD1, GAUT1, GAUT7, and P19 (Voinnet et al., 2003) constructs were 0.05, 0.2, 0.1, and 0.1, respectively. *Agrobacterium* cell suspensions were infiltrated into leaves of greenhouse-grown, 4-week old *N. benthamiana*. Sections of infiltrated leaves were observed 4-6 days after infiltration by confocal scanning laser microscopy (TCS SP2, Leica Microsystems, Wetzlar, Germany).

Proteomics analyses of GAUT1:GAUT7 complex

The experimental scheme is presented in Figure 3.S4. Three independent experiments were done for studies of the non-reducing SDS-PAGE-stable core complex and the complex-associating proteins. In each independent proteomics experiment, three immunoprecipitation treatments were carried out: IP-GAUT1, IP-GAUT7, and IP-Control (using anti-GAUT1-specific IgG, anti-GAUT7-specific IgG, and total pre-immune IgG, respectively). To ensure high

stringency results, the antigen-purified specific antibodies (see *SI Materials and Methods*) were covalently attached to magnetic beads (Dynabeads M-270 Epoxy (Invitrogen, Carlsbad, CA)) at 450 μg IgG per 22.5 mg beads. The GAUT1:GAUT7 protein complex was immunoprecipitated by incubating antibody-beads with 1.5 ml *Arabidopsis* SP fraction for 2 hrs at 4°C with end-to-end rotation. Beads were stringently washed (see Methods section on Immunoprecipitation above), and immunoabsorbed proteins eluted by adding 675 μl 100 mM glycine pH 2.5. After brief mixing, supernatant containing the eluted proteins was moved to a new tube and neutralized by adding 75 μl 1 M Tris pH 9.0. The immunoprecipitation procedure was performed repeatedly with the same antigen-beads using fresh SP fraction (total of four and three times for the core complex and complex-associating proteins, respectively).

Pooled eluted protein from each treatment was concentrated, and separated by 7.5% SDS-PAGE using non-reducing and reducing conditions to study the core complex and the complex-associating proteins, respectively. One fifth of each sample was separated and analyzed by Sypro Ruby-stained SDS-PAGE and the remainder separated by un-stained SDS-PAGE. To isolate GAUT1:GAUT7 core complex, excision was made on the non-reducing gel as illustrated in Figure 3.S5-A, using immunoblots of the same sample for comparison. In the study of associating proteins, the ~50-125 kDa region of the un-stained reducing gel was excised into 8 sections (illustrated in Figure 3.S5-B). This range was selected since preliminary LC-MS/MS analyses indicated that identified proteins were from this region of the gel.

In-gel trypsin digestion and LC-MS/MS analyses are described in *SI Materials and Methods*. ProteoIQ version 1.5.02 software (BioInquire, Athens, GA) was used to analyze the protein and peptide identifications by employing separately two statistical algorithms on-board the software, i.e. the False Discovery Rate (FDR) and the Probability methods, with procedures

and parameters as detailed in *SI Materials and Methods*. Resulting protein identifications for each analysis (i.e. IP-Control, IP-GAUT1, and IP-GAUT7 data sets) were manually compared to identify proteins present in both IP-GAUT1 and IP-GAUT7, but not in IP-Control. Proteins from respective studies were considered either as components of the core complex or as the complex-associating proteins. Comparison of final protein lists obtained using the two statistical methods was done to select proteins reproducibly identified with both algorithms (for further details see *SI Text*).

Further information is provided in *SI Materials and Methods*.

SUPPORTING INFORMATION

SI Text – Proteomics Data Analysis

In the study of the non-reducing SDS-PAGE stable GAUT1:GAUT7 core complex, FDR analysis identified 2 protein groups in IP-Control, 3 groups in IP-GAUT1, and 13 groups in IP-GAUT7, while Probability analysis identified 8, 2, and 3 proteins groups in IP-Control, IP-GAUT1, and IP-GAUT7, respectively. In both analyses, GAUT1 and GAUT7 were consistently the only two proteins identified exclusively in the IP-GAUT1 and IP-GAUT7 (Table 3.S1).

In the study of the core complex-associating proteins, FDR analysis identified 28, 49, and 172 proteins groups in the IP-Control, IP-GAUT1, and IP-GAUT7 data sets, respectively. All proteins in the IP-Control data set that were also found in the IP-GAUT1 and IP-GAUT7 data sets were subtracted from the IP-GAUT1 and IP-GAUT7 data sets. Eleven of these proteins, however, were present in ≥4-fold abundance (based on normalized spectral count (N-SC) values) in both IP-GAUT1 and IP-GAUT7, compared to the IP-Control, and were "TOP PROTEIN" (the best fit protein whose sequence can account for all the identified peptides within a given protein

group) in the IP-GAUT1 and IP-GAUT7 data sets. Comparison of the remaining proteins in the IP-GAUT1 and IP-GAUT7 data sets resulted in 16 proteins, in addition to GAUT1 and GAUT7, that were co-identified in both. These 16 proteins and the 11 proteins present in >4-fold abundance in the IP-GAUT1 and IP-GAUT7 data sets, a total of 27 proteins, are considered as candidate proteins associating with the GAUT1:GAUT7 complex based on the FDR analysis (Table 3.S2).

Application of Probability analysis in studying the GAUT1:GAUT7 complex-associating proteins resulted in 34, 33, and 102 protein groups in the IP-Control, IP-GAUT1, and IP-GAUT7 data sets, respectively. Manual comparison of the data sets revealed 9 proteins, in addition to GAUT1 and GAUT7, that were identified exclusively in both IP-GAUT1 and IP-GAUT7, and were not present in the IP-Control. Additionally, four proteins were found >4-fold more abundant in both IP-GAUT1 and IP-GAUT7 compared to in IP-Control. These 13 proteins constitute candidate GAUT1:GAUT7 complex-associating proteins based on the Probability analysis (Table 3.S2). This number is approximately half of that based on the FDR analysis described above, suggesting that Probability analysis is more stringent than FDR analysis in this study. Final comparison of the protein lists from both analyses revealed an overlap of 14 proteins, including GAUT1 and GAUT7, that are consistently identified using both algorithms (Table 3.1) and are thus considered as putative GAUT1:GAUT7 complex-associating proteins to be studied in the future.

SI Materials & Methods

Materials

Dowex 50W X8 resin and all chemicals, unless otherwise noted, were from Sigma (St. Louis, MO). SP-Sepharose, PD-10 desalting column, HiTrap Protein A HP column, and NHS-

activated Sepharose 4 Fast Flow were from GE Healthcare (Piscataway, NJ). Dynabeads M-280 Sheep anti-Rabbit IgG, Dynabeads M-270 Epoxy, and Sypro-Ruby protein gel stain were from Invitrogen (Carlsbad, CA). Uridine diphosphate α-D-[^{14}C]glucuronic acid (UDP-D-[^{14}C]Glc*p*A, 6.67 GBq/mmol) was purchased from Perkin Elmer (Waltham, MA). PVDF membranes and Amicon Ultra 4 centrifugal devices were from Millipore (Billerica, MA). The 3,3',5,5'-tetramethylbenzidine (TMB) substrate kits were from Vector Labs (Burlingame, CA). Multiple Antigenic Peptides (MAPs) of GAUT1 and GAUT7 were purchased from Sigma Genosys (St. Louis, MO). Spectra/Por dialysis tubing was from Spectrum Laboratories, Inc. (Rancho Dominguez, CA).

Preparation of SP-Sepharose-purified Arabidopsis membrane proteins (SP fraction)

SP-Sepharose-purified membrane proteins from *Arabidopsis* were prepared essentially as described in Sterling et al. (2006) with slight modifications. Briefly, 15 ml of solubilized *Arabidopsis* membrane proteins (prepared as described in Guillaumie et al., 2003) were loaded onto a 10 ml SP-Sepharose column equilibrated with Buffer A (50 mM Hepes, pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 25% (v/v) glycerol, 1% (v/v) Triton X-100). After washing with 20 ml of Buffer A, bound proteins were eluted with a NaCl step gradient in Buffer A (10 ml each of 100 mM, 200 mM, 300 mM, 400 mM, 500 mM, and 600 mM NaCl in Buffer A). The majority of GalAT activity, as well as GAUT1 and GAUT7, were eluted in the 300 mM NaCl and the first 2.5 ml of the 400 mM NaCl fractions. These fractions were pooled, desalted over a PD-10 column, and stored at -80°C until use.

Synthesis of UDP-D-[^{14}C]GalpA

UDP-D-[^{14}C]Gal*p*A was synthesized enzymatically from UDP-D-[^{14}C]Glc*p*A as described (Liljebjelke et al., 1995) with slight modification. UDP-[^{14}C]GlcA (50 μCi) was lyophilized and resuspended in 50 μl reaction mixture containing 50 mM ammonium formate pH 7.4, 20% glycerol, and UDP-D-glucuronic acid 4-epimerase (kindly provided by Sarah Inwood, CCRC-UGA) at an enzyme:substrate ratio of 2:5 (w/w). The reaction was incubated at 37°C for 3 hrs and terminated by addition of 300 μl of 5:1 (v/v) chloroform/methanol. Reaction product UDP-[^{14}C]GalA was extracted and purified as described (Liljebjelke et al., 1995).

GAUT7 immunodepletion

For the GAUT7 immunodepletion curve shown in Figure 3.S1, the immunoprecipitations, GalAT activity assays, and Western blotting were performed as previously described for GAUT1 (Sterling et al., 2006).

Production of RG-I oligomers of degrees of polymerization (DP) 6-26 from Arabidopsis seed mucilage

Arabidopsis seed mucilage was extracted by incubating one ml (~0.75 g) of wild-type seeds with 100 ml water at 50°C with stirring for 1 hr. This step was repeated six times and the mucilage-containing supernatant from each step was pooled. The total extract was filtered through 2 layers of miracloth (Calbiochem, San Diego, CA), concentrated on a rotary evaporator, and lyophilized. The dried mucilage was reconstituted in 10 mM ammonium formate, pH 4.5 and treated overnight with 5 U (1 U = 1 μmol of reducing sugar produced per minute) of *Aspergillus niger* endo-polygalacturonase-II (EPG-II; gift of Carl Bergmann, CCRC) per ~700

mg mucilage in 40 ml volume. Following digestion, the reaction was dialyzed (Spectra/Por 6; 1000 MWCO dialysis tubing) against water, filtered through a 0.45 μm Target nylon filter (Thermo Fisher Sci., Rockwood, TN), and lyophilized. The resulting EPG-treated mucilage is referred to as RG-I crude polymer.

Two types of linear rhamnogalacturonan oligomers were generated, as described (Mutter et al., 1998) with modification, from the RG-I crude polymer using the following approach. (1) RG oligomers containing non-reducing end rhamnosyl (Rha) residues ($(R\text{-}G)_n$, named RG-I-R) were generated by digestion of RG-I crude polymer with RG-hydrolase (Schols et al., 1990; Mutter et al., 1998) (from *Aspergillus aculeatus*, 2.05 mg protein/ml, a kind gift from Henk Schols, Wageningen University). (2) RG oligomers with non-reducing end galacturonosyl (GalA) residues ($(G\text{-}R)_n$, named RG-I-G) were generated by controlled acid hydrolysis of RG-I crude polymer (Renard et al., 1997). RG-I-R oligomers were made in a 200 μl reaction containing 50 mM sodium acetate buffer pH 5.4, 3 mg RG-I crude polymer, and 0.256 μg RG-hydrolase (0.085 μg enzyme/mg substrate). The reaction was incubated at 40°C for 48 hrs, terminated by boiling for 5 min, and desalted by passage through a 2 ml Dowex 50W X8 column using 3 column volumes of sterile water. The desalted digestion product was lyophilized and dissolved in water. RG-I-G oligomers were made in a 500 μl reaction containing 5 mg RG-I crude polymer and 0.1 M HCl. The reaction was incubated at 80°C for 4 hrs and terminated by addition of 50 μl of 1 M NaOH. The reaction product was desalted by dialysis (Spectra/Por CE; 500 MWCO dialysis tubing) against water, lyophilized, and dissolved in water. These basic reactions were scaled-up as necessary.

RG-I-R and RG-I-G oligomers were separated by High pH Anion Exchange Chromatography (HPAEC) over a Dionex (Sunnyvale, CA) CarboPac PA-1 column with

carbohydrate detection using a Dionex electrochemical detector (ECD) in the pulsed amperometric detection (PAD) mode. Analytical separation was done over a 4x250 mm column with a CarboPac PA Guard column (3x25 mm) at a flow rate of 1 ml/min and addition of post-column 400 mM NaOH to obtain a final pH of 12.5-13.0 in the eluent for optimal electrochemical detection. Semi-preparative separation was done over a 9x25 mm column at a 2 ml/min flow rate and with post-column addition of 1 M NaOH to obtain a final pH of 12.5-13.0 in the eluent for electrochemical detection. Both analytical and semi-preparative separations were done using the following sodium acetate buffer gradient: 0-2 min, 0.05 M; 2-15 min, 0.05-0.4 M; 15-47 min, 0.4-0.8 M; 47-50 min, 0.8-1 M; 50-55 min, 1 M. The identities of the RG oligosaccharide peaks of DP 4-18 were verified by MALDI-TOF analysis. Briefly, the individual peaks were collected from analytical separations (with post-column base and ECD on) into tubes containing acid (HCl) to immediately neutralize the solutions. These peak fractions were desalted, lyophilized, re-dissolved in sterile water, and analyzed for mass by MALDI-TOF on an Applied Biosystem 4700 Proteomics Analyzer (Applied Biosystem, Foster City, CA) using both negative and positive reflector modes. A mixture of oligomers of DP 6-26 (DP refers to the total number of monosaccharides in the chain, e.g. a DP 6 oligomer, regardless of the type, consists of three Rha and three GalA residues) was collected for each type of the RG-I oligomer from semi-preparative runs with the post-column base turned off (collection was made based on retention time of previous trial runs with post-column base on), dialyzed to remove salt, and lyophilized. The DP 6-26 RG-I oligomers of each type (i.e. RG-I-G and RG-IR) were reconstituted in water and stored at -20°C until use.

Bimolecular Fluorescence Complementation and co-expression of GAUT1 and GAUT7 fused to fluorescent proteins

Bacterial strains and bacterial and plant growth conditions. *Escherichia coli* DH10B was used for cloning and *Agrobacterium tumefaciens* PGV3850 C58C1 was used for transient transfection of *Nicotiana benthamiana*. The bacteria were grown in LB medium at 37°C and 28°C, respectively. The concentrations of antibiotics used during selection were as follows: kanamycin 50 µg/ml for *E. coli*, and kanamycin 50 µg/ml, rifampicin 100 µg/ml, and ampicilin 50 µg/ml for *A. tumefaciens*. *N. benthamiana* was grown in the greenhouse under 28°C-16 hr day and 22°C-8 hr night conditions.

Plasmid constructions. The construction of recombinant proteins was carried out as previously described (Vainonen et al., 2008) to generate fluorescent proteins fused to the C-termini of GAUT and ARAD1 proteins. The GAUT1, GAUT7, and ARAD1 coding sequences were amplified by gene specific synthetic oligonucleotides containing USER-cloning tails and USER cloning sites as underlined and italicized, respectively: GAUT1-F (nt96, GGCTTAAUATGGCGCTAAAGCGAGGGCTGGCTTAAUATGGCTATGGCGGAAATGGCAACGA), GAUT1-R (nt97, GGTTTAAU*TAAGGATCCTTAATTAAGCCTCAGCGG*TTCATGAAGGTTGCAAC), GAUT7-F (nt98, GGCTTAAUATGAAAGGCGGAGGCGGTGG), GAUT7-R (nt99, GGTTTAAU*TAAGGATCCTTAATTAAGCCTCAGCGG*AGGATTCACGTTACAGTCAC), ARAD1-F (nt122, GGCTTAAUATGGCGCGTAAATCTTCCCT), ARAD1-R (nt123, GGTTTAAU*TAAGGATCCTTAATTAAGCCTCAGCGG*AAATGGAAGTGATAAGACCGGT). The PCR product was mixed with the PacI/Nt.BbvCI-digested plasmid pCAMBIA330035Su and treated with USER enzyme mix (New England Biolabs, Beverly, MA) for 35 min at 37°C and 25

min at 25°C. The reaction mix was directly used to transform *E. coli* DH10B competent cells. The resulting plasmids, pGAUT1.1 and pGAUT7.2, were digested with PacI and Nt.BbvCI and annealed with PCR products of the full-length, Yn (residues 1-155), and Yc (residues 156-238) fragments of Venus-YFP amplified with the synthetic oligonucleotides nt59 and nt34, nt59 and nt35, and nt61 and nt34, respectively, as described above. The plasmid L13.1 containing ARAD1 coding sequence was treated similarly except that the primers nt59 and nt61 were replaced by nt100 and nt101, respectively. Sequences of the oligonucleotides were as follows: nt59 (GGCTTAAUCTGGGTAGCGGTGGAATGGTGAGCAAGGGCGAGGAG), nt34 (GGTTTAAUTTACTTGTACAGCTCGTCCAT), nt61 (GGCTTAAUCTGGGTAGCGGTGGAATGGACAAGCAGAAGAACGGCATC), nt35 (GGCTTAAUCATGTGAGCAAGGGCGAGGAG), nt100 (GGCTTAAUTGTTGTCCAGGATGTTGTCTGGGTAGCGGTGGAATGGTGAGCAAGGGC GAGGAG), nt101 (GGCTTAAUTGTTGTCCAGGATGTTGTCTGGGTAGCGGTGGAATGGACAAGCAGAAG AACGGCATC). For generation of C-terminally truncated GAUT1 containing the first N-terminal 100 and 291 a.a. residues fused C-terminally to GFP (GAUT1(100)-GFP and GAUT1(291)-GFP, respectively), the 300 nt and 873 nt of 5'-end of GAUT1 were amplified by PCR using primer pairs nt97/nt149 and nt96/nt150, respectively, and were ligated with PacI/Nt.BbvCI-digested plasmid USER-GFP (Sakuragi et al., In print) as described above. The sequences of the nucleotides were as follows: nt149 (GGTTTAAUCCCTTAAAAGAATCAAGGCTAAG) and nt150 (GGTTTAAUCCGTCATATAACTGCATCTTAGC). The plasmids were directly sequenced for

verification of the inserts. The plasmids were introduced into *A. tumefaciens,* and transformed cells were subsequently selected in the presence of kanamycin, rifampicin, and ampicilin.

Confocal microscopy. Sections of the infiltrated tobacco leaves were observed 4-6 days after infiltration using a confocal scanning laser microscope (TCS SP2, Leica Microsystems, Wetzlar, Germany). BiFC analysis was done using an excitation wavelength of 514 nm with constant energy setting of 40%, and emission wavelengths of 535-560 nm with constant gain setting between 833 and 950 that varied between experiments but were kept at a constant value within each experiment. For co-expression of GFP and YFP, sequential scanning was with excitation at 488 nm and 514 nm with constant energy settings of 40% and 43%, respectively, and emission at 495-510 nm and 545-560 nm with constant gain setting of 748 and 747, respectively. For co-expression of the GFP-fused GAUT1 truncated constructs with non-tagged GAUT7, excitation was at 488 nm with the energy setting of 40%, and emission at 495-560 nm with gain setting of 941. Chloroplast autofluorescence was detected at 650-707 nm. Scan speed was 800 Hz with a line average of 8. For quantification of GFP and YFP signals upon GAUT1-GFP and GAUT7-YFP co-expression, areas with characteristic Golgi morphology were manually selected and the pixel mean value within each area in each of the GFP and YFP detection channels was extracted using Leica LCS Lite software (Leica Microsystems, Wetzlar, Germany). The values were subsequently plotted using Microsoft Excel (Microsoft Denmark, Hellerup, Denmark). For quantification of GAUT1-GFP, GAUT1(100)-GFP, and GAUT1(292)-GFP, image processing was applied to each image using ImageJ software (http://rsbweb.nih.gov/ij/). The minimum threshold was set at 50 to filter out the background noise. Average pixel area value of GAUT1-GFP signals when co-expressed with GAUT7 was 8.06 ± 7.32, therefore a size filter ranging from 2 to 23 pixel area values was used to segment the

signals above the threshold value in all images to minimize the contribution of noise and non-Golgi signals derived from ER, nuclei, and cytosol. The quantity of GFP signals was expressed as i) the number of objects with GFP signal in each image and ii) the sum of the pixel area with the GFP signal in each image. At least three independent images were analyzed for each recombinant protein and values were plotted using Microsoft Excel.

Analysis of the GAUT1 and GAUT7 promoter:GUS fusions

Promoter regions for *GAUT1* and *GAUT7* (1932 and 868 bp upstream of the start codons, respectively) were PCR-amplified and cloned into pGEM-TE vector (Promega, Madison, WI). Upon sequence verification, the promoter fragments were fused to the bacterial *uidA* (*gusA*) gene encoding β-glucuronidase (GUS) in pBI101 vector (Arabidopsis Biological Resource Center, Columbus, OH) (Jefferson et al., 1987). The resulting plasmids were transformed into *Agrobacterium tumefaciens* GV3101::pMP90 by electroporation. Stable transformation of wild-type *Arabidopsis thaliana* (Col-O) was performed using the floral dip method (Clough and Bent, 1998). Kanamycin-resistant lines were selected, allowed to seed set, and the T2 transgenic plants were grown for GUS expression analyses. Histochemical staining of plant tissues for GUS activity was performed as described by Jefferson et al. (Jefferson et al., 1987). Plant materials harvested from T2 plants were submerged in staining solution (50 mM sodium phosphate pH 7.0, 15% (v/v) methanol, 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 0.05% (v/v) Tween-20, and 2 mM 5-bromo-4-chloro-3-indolyl-β-glucuronide), vacuum-infiltrated for 10 min at 79.8 kPa, and incubated at 37°C for 0.5 to 72 hrs. Tissues were subsequently washed repeatedly with 70% (v/v) ethanol to remove pigments. Samples were analyzed using a dissecting microscope (Olympus SZH) or a light microscope (Nikon Eclipse

80i) and photographed (Nikon DS-Ri1 digital camera). Similar results were observed using multiple T2 plants from at least 5 independent T1 lines.

Purification of anti-GAUT1 and anti-GAUT7 specific antibodies

In order to minimize non-specific binding, the immunoprecipitations of the protein complex for the proteomics analyses were done using anti-GAUT1 and anti-GAUT7 specific antibodies purified from the respective antiserum using GAUT1 and GAUT7 peptide antigen columns, respectively. To make the antigen columns, Multiple Antigenic Peptides (MAPs) of GAUT1 (Sterling et al., 2006) and GAUT7 (see Methods section in the main text) were coupled at 20 mg antigen/ml beads to NHS-activated Sepharose 4 Fast Flow following the manufacturer's instructions. Total IgG was purified from anti-GAUT1 and anti-GAUT7 antiserum and from pre-immune serum using HiTrap Protein A HP 1 ml columns following the manufacturer's instructions, and the IgG concentrated and buffer-exchanged into PBS (pH ≈7.3, 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4) by ultrafiltration using Amicon Ultra-4 (30 kDa MWCO) centrifugation devices.

To purify the specific IgG, a column (~1 ml) of the GAUT1 or GAUT7 antigen-beads was prepared and equilibrated with 10 column volumes of PBS. The previously purified total IgG from anti-GAUT1 or anti-GAUT7 antiserum was applied to the respective column, and the flow-through was collected and re-applied 3-4 times to ensure maximum binding to the affinity column. After washing the column with 10 column volumes of PBS, the specific IgG was eluted by applying 5 column volumes of 100 mM citric acid, pH 2.5. Eluent was collected in 500 µl fractions in tubes containing 100 µl 1 M Tris, pH 9.0 to neutralize the eluent. Fractions containing IgG, as judged by ELISA, were pooled, concentrated and buffer-exchanged into PBS

by ultrafiltration as described above. The resulting purified specific IgG was assayed by ELISA and Western blotting to verify its function.

Mass spectrometry analyses

Proteolytic Digestion. Each gel section was cut into smaller pieces (1 x 1 mm) and washed twice with 50% acetonitrile (ACN) in 100 mM ammonium bicarbonate (Ambic). Proteins were reduced by incubating the gel bands in 10 mM dithiothreitol (DTT)/100 mM Ambic solution at 56°C for 1 hr and then alkylated in 55 mM iodoacetamide/100 mM Ambic for 1 hr at room temperature in the dark. The solution was adjusted to a pH of 8 with 100 mM Ambic. Enzymatic digestion was then performed using sequencing grade porcine trypsin (Promega, Madison, WI) with a ratio of 1:50 (w/w) enzyme to substrate and overnight incubation at 37°C. The tryptic peptides were extracted once with 100 mM Ambic/ACN (1:1) and twice with 3% formic acid/ACN (1:1). Combined extracts were dried to completion by vacuum centrifugation, resuspended in 50 μl of 0.1% formic acid and stored at -20°C before analysis by MS.

LC-MS/MS Analysis. The peptides obtained from proteolytic digestion were analyzed on an Agilent 1100 capillary LC (Palo Alto, CA) interfaced directly to a LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA). Mobile phases A and B were H_2O-0.1% formic acid and ACN-0.1% formic acid, respectively. The digested peptides were pressure-loaded for 1 hr onto a PicoFrit 11 cm x 50 μm column (New Objective, Woburn, MA) packed with 8 cm length, 5 μm diameter C18 beads. Peptides were eluted from the column into the mass spectrometer during a 90-min linear gradient from 5 to 60% of total solution composed of mobile phase B at a flow rate of 200 nl/min. The instrument was set to acquire MS/MS spectra

on the nine most abundant precursor ions from each MS scan with a repeat count of 1 and repeat duration of 5 s. Dynamic exclusion was enabled for 200 s. Raw tandem mass spectra were converted into the mzXML format and then into peak list using ReAdW software followed by mzMXL2Other software (Pedrioli et al., 2004). The peak lists were searched using Mascot 1.9 software (Matrix Science, Boston, MA).

Data analysis. Protein and peptide identifications were analyzed using ProteoIQ version 1.5.02 software (BioInquire, Athens, GA). Peptide identifications were assembled into protein groups and statistically validated using the ProValT algorithm (Weatherly et al., 2005) implemented in the software. Two statistical algorithms in the ProteoIQ software, namely the False Discovery Rate (FDR) (Weatherly et al., 2005) and the Probability (Nesvizhskii et al., 2003) methods, were used separately to analyze the data.

For database searching, a target database of annotated proteins was created from *Arabidopsis* genes obtained from NCBI (www.ncbi.nlm.nih.gov), and a decoy database was constructed by reversing the sequences in the target database. Database searches were performed against both the target and decoy databases for FDR analysis, or against only the target database for Probability analysis. In both cases the following parameters were used: full tryptic enzymatic cleavage with two possible missed cleavages, peptide tolerance of 1000 ppm, and fragment ion tolerance of 0.6 Da. Fixed modification was set as carbamidomethyl due to carboxyamidomethylation of Cys residues (+57 Da) and variable modification was chosen as oxidation of Met residues (+16 Da) and deamidation of Asp residues (+1 Da).

In FDR analyses, only proteins with protein FDR <1% were considered to be statistically significant. For the GAUT1:GAUT7 non-reducing SDS-PAGE stable core complex, at 1% FDR the mascot ion score thresholds for identification of proteins with 3 or more peptides were 23,

26, and 23; with 2 peptides, 32, 32, and 25; and with a single peptide, 41, 40, and 42 for the IP-Control, IP-GAUT1, and IP-GAUT7, respectively. For the study of the GAUT1:GAUT7 complex-associating proteins, at 1% FDR the minimum mascot ion thresholds for protein identifications with 3 or more peptides were 44, 44, and 35; with 2 peptides, 47, 47, and 40; and with a single peptide, 57, 64, and 60 for the IP-Control, IP-GAUT1, and IP-GAUT7, respectively. IP-GAUT1 and IP-GAUT7 data sets were additionally filtered for proteins present in 100% of the replicates (i.e. 3 replicates, except for the IP-GAUT1 and IP-GAUT7 core complex for which two robust replicates were used).

For Probability analysis, the resulting protein groups were validated using the onboard ProteinProphet algorithm (Nesvizhskii et al., 2003) at a probability of 0.99 for IP-GAUT1 and IP-GAUT7, and of 0.8 and 0.95 for the IP-control data sets in the studies of the GAUT1:GAUT7 core complex and of the complex-associating proteins, respectively. To ensure that only the highest quality assignments were used, prior to quantification, additional filters were applied to IP-GAUT1 and IP-GAUT7 data sets with the following criteria: unique peptides identifications ≥2, spectral count ≥5, and presence of the protein in 100% of the replicates.

Immunoprecipitation of GAUT1 for N-terminal sequencing

To accumulate enough GAUT1 protein for N-terminal sequencing, immunoprecipitation (by incubating 750 µl SP fraction per 22.5 mg IgG bead) and elution of the immunoabsorbed proteins were performed repeatedly using the anti-GAUT7-specific IgG beads (IgG beads of anti-GAUT7 antibody, instead of anti-GAUT1, were chosen since the former were about three times stronger than the latter, as judged by ELISA titer). Pooled eluted proteins were concentrated by ultrafiltration on Amicon Ultra 4 (10 kDa MWCO) centrifugation devices and

separated via reducing SDS-PAGE alongside pre-stained protein markers and a 20 μl aliquot of un-concentrated eluted proteins. Electrophoresed proteins were transferred onto a PVDF-PSQ (Millipore) membrane in CAPS buffer containing 10 mM 3(cyclohexylamino)-1-propane-sulfonic acid (CAPS) and 10% MeOH, pH 11.0. The lane containing un-concentrated proteins was immunoblotted with anti-GAUT1 antiserum to detect GAUT1 protein band, while lanes containing concentrated proteins were stained with Coomassie R-250. The position of the GAUT1 protein band, faintly stained by Coomassie, was estimated by side-by-side comparison to the immunoblotted portion of the membrane, and that section of the membrane was cut out. Membrane pieces containing the Coomassie-stained GAUT1 protein band were sent to the W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University (New Haven, CT) for N-terminal sequence analysis.

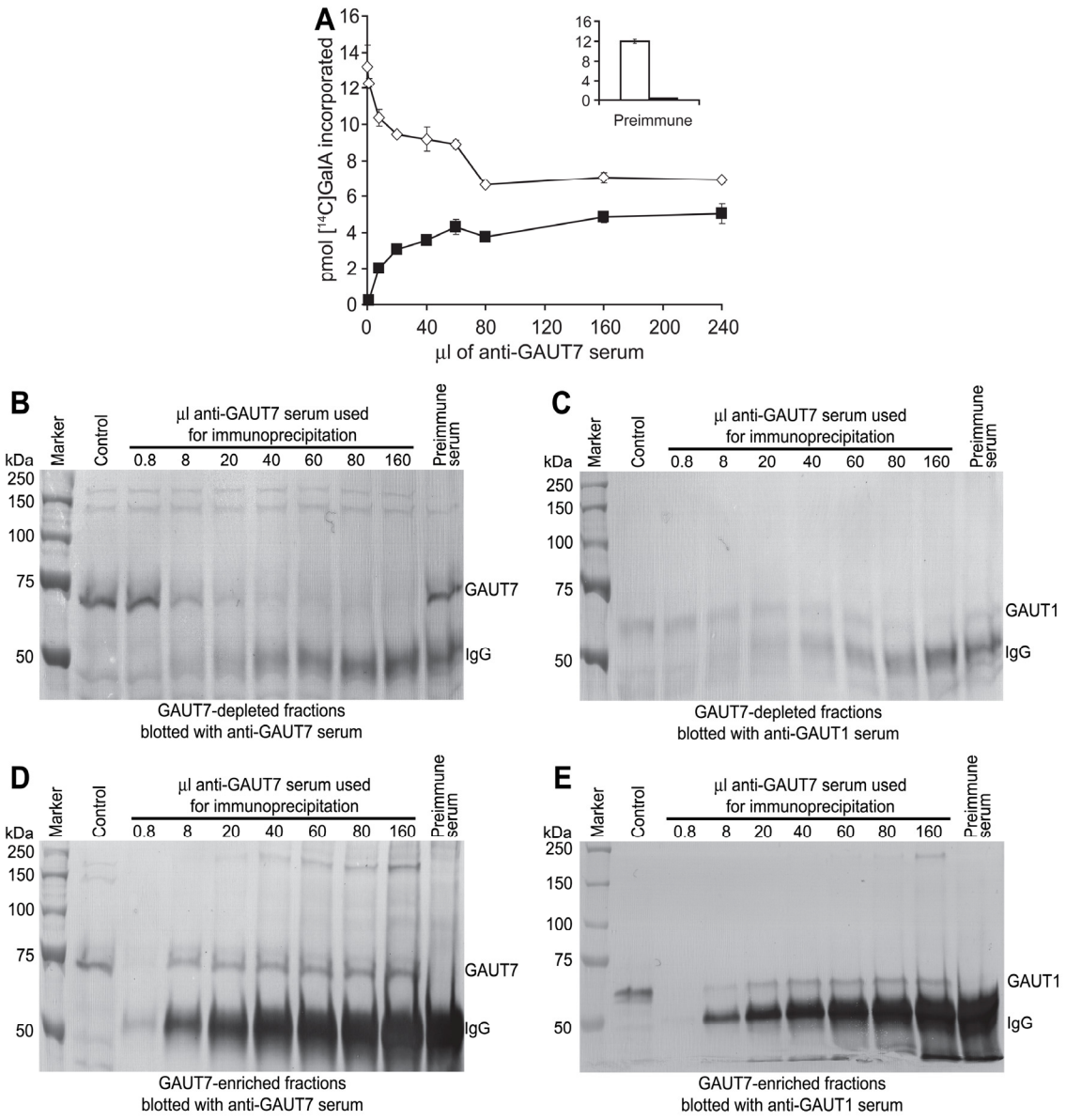


Figure 3.S1. Anti-GAUT7 serum immunoabsorbs GalAT activity from SP-Sepharose purified *Arabidopsis* solubilized enzyme (SP fraction).

(A) GalAT activity of immunodepleted (◇) and immunoabsorbed (■) fractions from immunoprecipitations using increasing amounts of anti-GAUT7 serum. Inset: Amount of GalAT activity in similar fractions treated with 160 µl preimmune serum (open bar = immunodepleted fraction; closed bar = immunoabsorbed fraction). The experiment was carried out twice with similar results.

(B and C) Western blots of the anti-GAUT7 serum-immunodepleted (supernatant) fractions probed with anti-GAUT7 and anti-GAUT1 sera, respectively.

(D and E) Western blots of the anti-GAUT7 serum-immunoabsorbed (GAUT7-enriched pellet) fractions, probed with anti-GAUT7 and anti-GAUT1 sera, respectively.

IgG: IgG heavy chain detected by the secondary antibody used in Western blotting. Control: SP fraction incubated for the same amount of time but without addition of serum. Preimmune treatment was with 160 μl of preimmune serum.

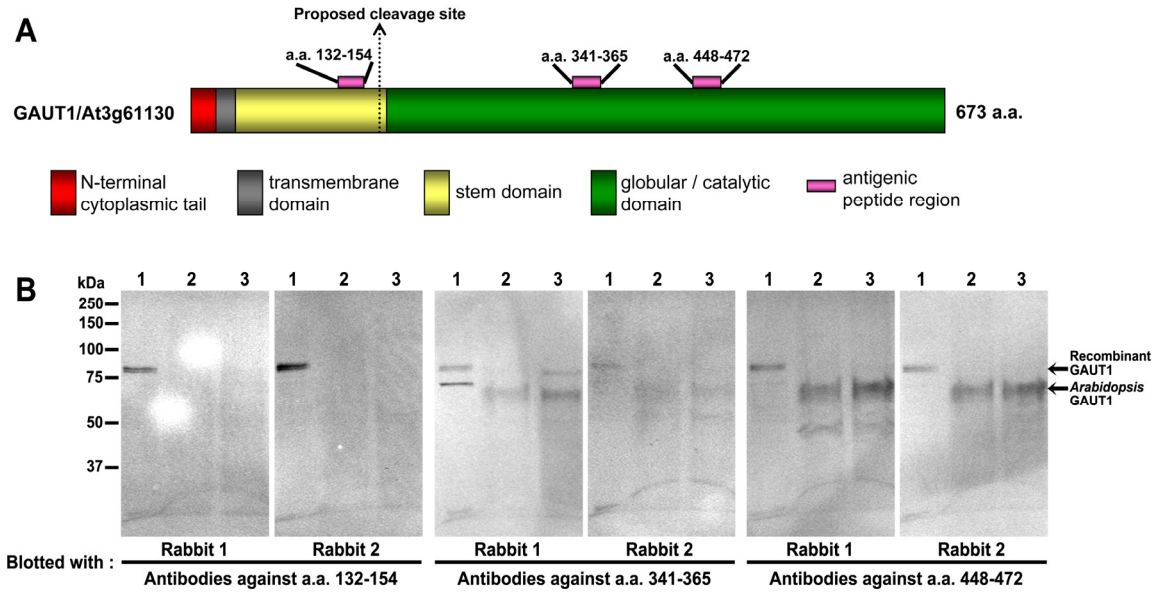


Figure 3.S2. Two out of three sets of anti-GAUT1 antibodies recognize mature *Arabidopsis* GAUT1.

(A) A schematic of GAUT1 type II transmembrane protein topology showing the positions, relative to the predicted full-length protein sequence, of the three peptide regions against which the anti-GAUT1 rabbit polyclonal antibodies were raised. N-terminal cytoplasmic tails and transmembrane domains were predicted using the HMMTOP v.2 program (http://www.enzim.hu/hmmtop/index.html); stem domains are estimated as approximately 25% of the total protein length. Arrow indicates the proposed cleavage site of *Arabidopsis* GAUT1. The positions of the different domains and the antigenic peptides are drawn to scale based on the protein's amino acid length.

(B) Western blots of recombinant and *Arabidopsis* GAUT1 probed using the three sets of anti-GAUT1 antibodies. Note that the antibodies against a.a. 132-154 recognize the recombinant GAUT1(△41) but not the *Arabidopsis* GAUT1. *Lane 1*: 10 µl of immunoprecipitated recombinant GAUT1(△41) prepared as previously described (Sterling, 2004). The amount loaded is the amount of recombinant protein immunoprecipitated using 40 µl conjugate anti-

hemagluttinin:protein A agarose beads from 1 ml lysate of human embryonic kidney (HEK) 293 cells (grown in one 15 cm^2 culture dish and harvested at 80% confluency) stably-transformed with a construct of truncated GAUT1 (missing the first 41 a.a.) fused to poly-histidine and hemagluttinin tags (Sterling et al., 2006). *Lane 2 and 3*: Fractions (5.1 μg protein/lane) from SP-Sepharose column-purified *Arabidopsis* solubilized membrane proteins eluted from the column with buffer containing 300 and 400 mM NaCl, respectively. Both of the SP-Sepharose purified fractions contain the majority of the purified GalAT activity. Western blotting was performed using the following dilutions of the six anti-GAUT1 antibodies: (from left to right) 1:10,000, 1:10,000, 1:5,000, 1:3,000, 1:3,000, and 1:1,000.

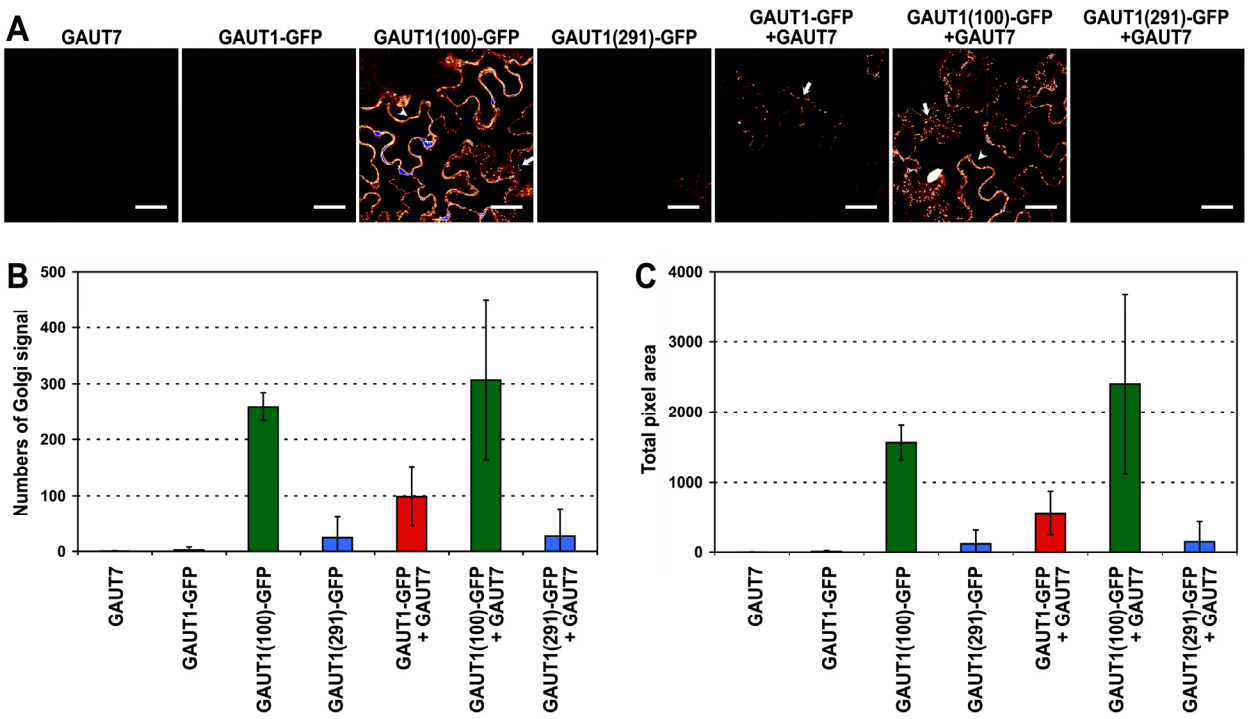


Figure 3.S3. Quantification of fluorescent signals obtained upon transient expression of GAUT1 (C-terminally truncated)-GFP fusion constructs in *N. benthamiana* leaves in the absence and presence of (non-tagged) GAUT7.

(A) Representative confocal microscopy imaging of the GFP fluorescence of the different combinations of constructs (scale bars are 40 μm). Note that this figure has a different magnification compared to that in Figure 3.3 in the main text. The average diameter of Golgi signals in both figures is between 1.4 and 1.6 μm, based on quantification of at least eighteen randomly selected Golgi signals in each figure. Examples of Golgi and ER signals are marked by white arrows and arrow heads, respectively.

(B and C) Quantification of GFP fluorescence, measured as numbers of Golgi signal and total pixel area, respectively. GAUT1-GFP signal is clearly observed in the presence of GAUT7, while not detectable in the absence of GAUT7, corroborating the results from co-expression of GAUT1-GFP and GAUT7-YFP (see Figure 3.3 in the main text). GAUT1(100)-GFP gives a

high amount of a broadly localized signal that is virtually identical both in the presence or absence of GAUT7, indicating that GAUT1(100)-GFP accumulation in the Golgi is not dependent on GAUT7. The high expression is presumably largely due to the high expression level. In contrast, GAUT1(291)-GFP does not accumulate an appreciable amount of signal regardless of the presence of GAUT7, presumably because the cleaved portion of the GAUT1-GFP fusion is secreted out of the cells. It is known that GFP fluorescence is low in the apoplast due to rapid protein degradation and to sub-optimal conditions for GFP fluorescence in the apoplast (Zheng et al., 2004).

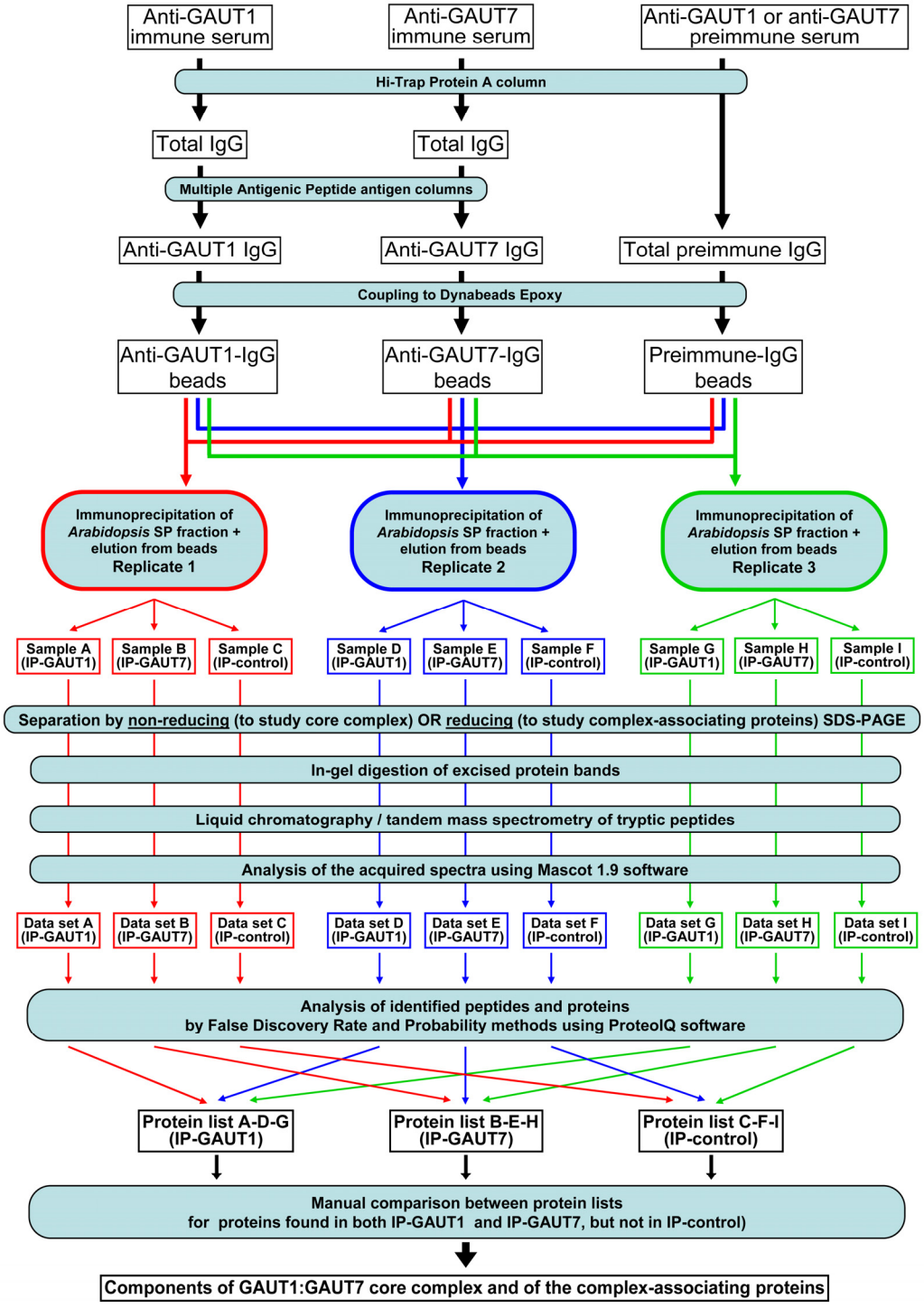
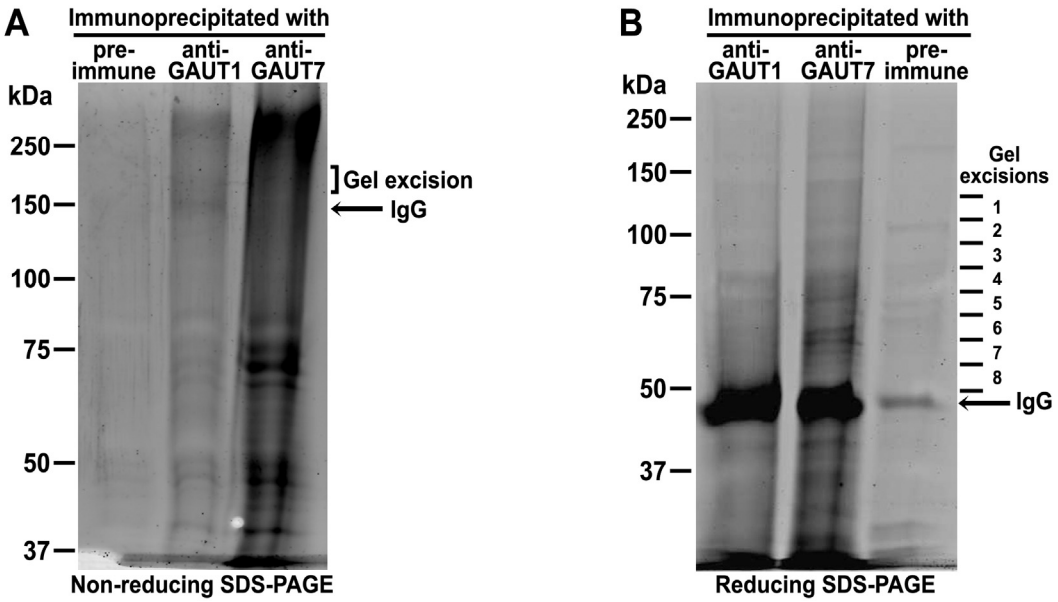


Figure 3.S4. Flow chart of sample preparation and proteomics analyses of the GAUT1:GAUT7 non-reducing SDS-PAGE stable core complex and of the GAUT1:GAUT7 complex-associating proteins.



C

```
   1  MALKRGLSGVNRIRGSGGGSRSVLVLLIFFCVFAPLCFFVGRGVYIDSSNDYSIVSVKQN   60
  61  LDWRERLAMQSVRSLFSKEILDVIATSTADLGPLSLDSFKKNNLSASWRGTGVDPSFRHS  120
 121  ENPATPDVKSNNLNEKRDSISKDSIHQKVETPTKIHRRQLREKRREMRANELVQHNDDTI  180
 181  LKLENAAIERSKSVDSAVLGKYSIWRRENENDNSDSNIRLMRDQVIMARVYSGIAKLKNK  240
 241  NDLLQELQARLKDSQRVLGEATSDADLPRSAHEKLRAMGQVLAKAKMQLYDCKLVTGKLR  300
 301  AMLQTADEQVRSLKKQSTFLAQLAAKTIPNPIHCLSMRLTIDYYLLSPEKRKFPRSENLE  360
 361  NPNLYHYALFSDNVLAASVVVNSTIMNAKDPSKHVFHLVTDKLNFGAMNMWFLLNPPGKA  420
 421  TIHVENVDEFKWLNSSYCPVLRQLESAAMREYYFKADHPTSGSSNLKYRNPKYLSMLNHL  480
 481  RFYLPEVYPKLNKILFLDDDIIVQKDLTPLWEVNLNGKVNGAVETCGESFHRFDKYLNFS  540
 541  NPHIARNFNPNACGWAYGMNMFDLKEWKKRDITGIYHKWQNMNENRTLWKLGTLPPGLIT  600
 601  FYGLTHPLNKAWHVLGLGYNPSIDKKDIENAAVVHYNGNMKPWLELAMSKYRPYWTKYIK  660
 661  FDHPYLRRCNLHE                                                720
```

D

```
   1  MALKRGLSGVNRIRGSGGGSRSVLVLLIFFCVFAPLCFFVGRGVYIDSSNDYSIVSVKQN   60
  61  LDWRERLAMQSVRSLFSKEILDVIATSTADLGPLSLDSFKKNNLSASWRGTGVDPSFRHS  120
 121  ENPATPDVKSNNLNEKRDSISKDSIHQKVETPTKIHRRQLREKRREMRANELVQHNDDTI  180
 181  LKLENAAIERSKSVDSAVLGKYSIWRRENENDNSDSNIRLMRDQVIMARVYSGIAKLKNK  240
 241  NDLLQELQARLKDSQRVLGEATSDADLPRSAHEKLRAMGQVLAKAKMQLYDCKLVTGKLR  300
 301  AMLQTADEQVRSLKKQSTFLAQLAAKTIPNPIHCLSMRLTIDYYLLSPEKRKFPRSENLE  360
 361  NPNLYHYALFSDNVLAASVVVNSTIMNAKDPSKHVFHLVTDKLNFGAMNMWFLLNPPGKA  420
 421  TIHVENVDEFKWLNSSYCPVLRQLESAAMREYYFKADHPTSGSSNLKYRNPKYLSMLNHL  480
 481  RFYLPEVYPKLNKILFLDDDIIVQKDLTPLWEVNLNGKVNGAVETCGESFHRFDKYLNFS  540
 541  NPHIARNFNPNACGWAYGMNMFDLKEWKKRDITGIYHKWQNMNENRTLWKLGTLPPGLIT  600
 601  FYGLTHPLNKAWHVLGLGYNPSIDKKDIENAAVVHYNGNMKPWLELAMSKYRPYWTKYIK  660
 661  FDHPYLRRCNLHE                                                720
```

Figure 3.S5. Proteomics analyses of the GAUT1:GAUT7 core and of the complex-associating proteins.

(A and B) Representative Sypro-Ruby-stained non-reducing (A) and reducing (B) SDS-PAGE analyses of aliquots of immunoprecipitated GAUT1:GAUT7 complexes for proteomics analyses of the core complex and of the complex-associating proteins, respectively. The same

immunoprecipitated materials were resolved by un-stained SDS-PAGE, from which gel slices were taken at the positions illustrated and subjected to in-gel trypsin digestion and subsequent LC-MS/MS. IgG: IgG heavy chain used in immunoprecipitations.

(C and D) Mass spectrometry sequence coverage of mature *Arabidopsis* GAUT1 obtained from the LC-MS/MS analyses of the GAUT1:GAUT7 complex and of the complex with associating proteins, immunoprecipitated by anti-GAUT1 (C) and anti-GAUT7 (D) specific IgGs, respectively. The identified tryptic peptide sequences are shown in red.

Table 3.S1. Analysis of non-reducing SDS-PAGE-stable core complex: components of the GAUT1:GAUT7 core complex based on False Discovery Rate and Probability analyses.

Only TOP PROTEINS, of which the protein sequence can account for all peptides within a protein group, are considered as components of the complex. * NSAF = Normalized Spectral Abundance Factor. ** Corrected NSAF values are calculated using the size of the proteolytically-processed GAUT1, i.e. 506 a.a. long (based on the proposed cleavage site between Met_{167} and Arg_{168}).

Protein Group	Source	Protein ID	Gene ID	Sequence Name	Protein Length (a.a.)	Protein Weight (kDa)	IP-GAUT1					IP-GAUT7				
							Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**	Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**
False Discovery Rate analysis																
1	TOP PROTEIN	NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	673	77.31	1	655.8	38	0.422	**0.519**	1	342.0	23	0.073	**0.623**
2	TOP PROTEIN	NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	619	69.69	1	492.3	43	0.520	**0.481**	1	287.8	17	0.059	**0.377**
					Ratio of NSAF values of GAUT1:GAUT7:					0.812	**1.081**				1.221	**1.655**
					Average ratio:					1.016						
					Average ratio (corrected):**					**1.371**						
Probability analysis																
1	TOP PROTEIN	NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	673	77.31	1	531.0	19	0.555	**0.624**	1	256.2	8	0.434	**0.583**
2	TOP PROTEIN	NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	619	69.69	1	362.7	14	0.445	**0.376**	1	196.9	7	0.417	**0.417**
					Ratio of NSAF values of GAUT1:GAUT7:					1.249	**1.660**				1.041	**1.398**
					Average ratio:					1.145						
					Average ratio (corrected):**					**1.529**						

Table 3.S2. Analysis of the GAUT1:GAUT7 complex-associating proteins: 29 and 15 proteins, including GAUT1 and GAUT7, identified using False Discovery Rate (FDR) and Probability analyses, respectively.

FDR analysis: Eighteen proteins were identified exclusively in both anti-GAUT1-IgG and anti-GAUT7-IgG immunoprecipitations (IP-GAUT1 and IP-GAUT7), but not in control (IP-Control, using preimmune IgG). Eleven additional proteins (highlighted in pink) were found to be >4-fold more abundant in IP-GAUT1 and IP-GAUT7, compared to IP-Control.

Probability analysis: Eleven proteins were identified exclusively in both anti-GAUT1-IgG and anti-GAUT7-IgG immunoprecipitations (IP-GAUT1 and IP-GAUT7), and not in control (using preimmune IgG) immunoprecipitation (IP-Control). Four additional proteins (highlighted in green) were found to be >4-fold more abundant in IP-GAUT1 and IP-GAUT7, compared to IP-Control.

Only TOP PROTEINS, of which the protein sequence can account for all peptides within a protein group, are considered as components of the complex. Proteins are listed based on their total score in IP-GAUT1. * NSAF = Normalized Spectral Abundance Factor. ** Corrected NSAF values are calculated using the size of the proteolytically-processed GAUT1, i.e. 506 a.a. long (based on the proposed cleavage site between Met_{167} and Arg_{168}).

Protein Group	Source	Protein ID	Gene ID	Sequence Name	Protein Length (a.a.)	Protein Weight (kDa)	IP-GAUT1					IP-GAUT7				
							Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**	Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**
False Discovery Rate analysis																
1	TOP PROTEIN	NP_851120.1	At5g42080	ADL1 (ARABIDOPSIS DYNAMIN-LIKE PROTEIN); GTP binding	610	68.11	1	1679.8	214	0.069	0.069	1	2271.6	234	0.028	0.028
2	TOP PROTEIN	NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	673	77.31	1	1492.0	211	0.063	0.082	1	2199.1	315	0.035	0.045

(Table 3.S2. … continued)

Protein Group	Source	Protein ID	Gene ID	Sequence Name	Protein Length (a.a.)	Protein Weight (kDa)	IP-GAUT1					IP-GAUT7				
							Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**	Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**
3	TOP PROTEIN	NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	619	69.69	1	1216.6	362	0.119	0.115	1	2133.9	496	0.059	0.059
4	TOP PROTEIN	NP_176170.1	At1g59610	ADL3 (ARABIDOPSIS DYNAMIN-LIKE 3)	920	100.15	1	1177.3	86	0.018	0.018	1	1273.1	104	0.008	0.008
5	TOP PROTEIN	NP_187337.2	At3g06810	acyl-CoA dehydrogenase-related	824	91.64	1	1124.2	116	0.027	0.028	1	370.5	32	0.003	0.003
6	TOP PROTEIN	NP_172936.1	At1g14830	ADL1C (DYNAMIN-LIKE PROTEIN 5); GTP binding / GTPase	614	68.66	1	1033.6	71	0.022	0.023	1	746.2	87	0.010	0.010
7	TOP PROTEIN	NP_177766.1	At1g76400	ribophorin I family protein	614	68.58	1	897.6	87	0.028	0.028	1	1352.6	208	0.025	0.025
8	TOP PROTEIN	NP_172500.1	At1g10290	ADL6 (DYNAMIN-LIKE PROTEIN 6)	914	99.09	1	838.9	70	0.015	0.015	1	1031.1	101	0.008	0.008
9	TOP PROTEIN	NP_193537.2	At4g18030	dehydration-responsive family protein	621	70.27	1	697.6	32	0.010	0.010	1	610.7	55	0.007	0.006
10	TOP PROTEIN	NP_197195.2	At5g16930	AAA-type ATPase family protein	644	70.93	1	660.2	72	0.022	0.022	1	840.5	144	0.017	0.016
11	TOP PROTEIN	NP_190258.1	At3g46740	TOC75-III (translocon outer membrane complex 75-III); protein translocase	818	89.12	1	660.0	46	0.011	0.011	1	739.5	54	0.005	0.005
12	TOP PROTEIN	NP_177234.1	At1g70770	unknown protein	610	66.80	1	564.8	68	0.022	0.022	1	1121.6	193	0.024	0.023
13	TOP PROTEIN	NP_565435.1	At2g18330	AAA-type ATPase family protein	636	71.13	1	559.1	55	0.017	0.017	1	955.5	133	0.015	0.015
14	TOP PROTEIN	NP_179673.1	At2g20800	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	582	65.31	1	527.9	30	0.010	0.010	1	1207.8	154	0.020	0.019
15	TOP PROTEIN	NP_193847.2	At4g21150	ribophorin II (RPN2) family protein	691	74.60	1	520.4	28	0.008	0.008	1	1164.1	96	0.010	0.010
16	TOP PROTEIN	NP_186956.2	At3g03060	ATPase	628	69.51	1	503.6	41	0.013	0.013	1	604.8	73	0.009	0.008
17	TOP PROTEIN	NP_567094.1	At3g60190	ADL4/ADLP2/DRP1E/EDR3 (DYNAMIN-LIKE PROTEIN 4); GTP binding / GTPase	624	69.74	1	448.2	24	0.008	0.008	1	578.9	34	0.004	0.004
18	TOP PROTEIN	NP_193656.2	At4g19210	ATRLI2 (Arabidopsis thaliana RNase L inhibitor protein 2)	605	68.33	1	346.5	30	0.010	0.010	1	273.3	22	0.003	0.003
19	TOP PROTEIN	NP_567184.1	At4g00740	dehydration-responsive protein-related	600	67.49	1	319.2	17	0.006	0.006	1	574.6	51	0.006	0.006
20	TOP PROTEIN	NP_188616.1	At3g19820	DWF1 (DIMINUTO 1); catalytic	561	65.33	1	298.6	24	0.008	0.008	1	868.5	142	0.019	0.018
21	TOP PROTEIN	NP_191735.2	At3g61760	dynamin-like protein B (DL1B)	610	68.02	0	248.5	13	0.004	0.004	1	559.4	37	0.004	0.004
22	TOP PROTEIN	NP_176007.1	At1g56110	NOP56 (ARABIDOPSIS HOMOLOG OF NUCLEOLAR PROTEIN NOP56)	522	58.62	1	238.1	8	0.003	0.003	1	917.4	71	0.010	0.010
23	TOP PROTEIN	NP_195540.2	At4g38270	GAUT3 (Galacturonosyltransferase 3); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	680	77.70	1	236.9	29	0.009	0.008	0.981	159.9	30	0.003	0.003
24	TOP PROTEIN	NP_850940.1	At1g08660	glycosyl transferase family 29 protein / sialyltransferase family protein	474	54.16	1	206.9	7	0.003	0.003	1	611.4	56	0.009	0.009

(Table 3.S2. … continued)

Protein Group	Source	Protein ID	Gene ID	Sequence Name	Protein Length (a.a.)	Protein Weight (kDa)	IP-GAUT1					IP-GAUT7				
							Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**	Total Protein Probability	Total Score	Total Spectra	NSAF* value	Corrected NSAF value**
25	TOP PROTEIN	NP_199783.1	At5g49720	KOR1 (KORRIGAN); hydrolase, hydrolyzing O-glycosyl compounds	621	69.13	1	191.4	10	0.003	0.003	1	214.2	18	0.002	0.002
26	TOP PROTEIN	NP_191353.2	At3g57940	unknown protein	1028	115.38	1	159.6	7	0.001	0.001	1	532.9	32	0.002	0.002
27	TOP PROTEIN	NP_181042.1	At2g34970	eIF4-gamma/eIF5/eIF2-epsilon domain-containing protein	730	81.79	1	158.7	9	0.002	0.002	1	224.3	20	0.002	0.002
28	TOP PROTEIN	NP_196947.2	At5g14430	dehydration-responsive protein-related	612	69.80	1	157.7	12	0.004	0.004	1	414.7	41	0.005	0.005
29	TOP PROTEIN	NP_564083.1	At1g19370	unknown protein	605	67.47	1	149.1	6	0.002	0.002	0.999	145.2	14	0.002	0.002
Probability analysis																
1	TOP PROTEIN	NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	673	77.31	1	1573.6	122	0.073	0.095	1	2013.4	228	0.047	0.061
2	TOP PROTEIN	NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	619	69.69	1	1372.6	167	0.111	0.106	1	1668.6	327	0.072	0.071
3	TOP PROTEIN	NP_176170.1	At1g59610	ADL3 (ARABIDOPSIS DYNAMIN-LIKE 3)	920	100.15	1	973.2	48	0.020	0.021	1	1083.7	64	0.009	0.009
4	TOP PROTEIN	NP_177766.1	At1g76400	ribophorin I family protein	614	68.58	1	780.9	41	0.026	0.026	1	980.4	143	0.032	0.031
5	TOP PROTEIN	NP_172500.1	At1g10290	ADL6 (DYNAMIN-LIKE PROTEIN 6)	914	99.09	1	736.4	36	0.015	0.016	1	1000.2	68	0.010	0.010
6	TOP PROTEIN	NP_193537.2	At4g18030	dehydration-responsive family protein	621	70.27	1	722.0	25	0.016	0.016	1	612.1	42	0.009	0.009
7	TOP PROTEIN	NP_193847.2	At4g21150	ribophorin II (RPN2) family protein	691	74.60	1	641.4	19	0.011	0.011	1	1098.0	65	0.013	0.013
8	TOP PROTEIN	NP_190258.1	At3g46740	TOC75-III (translocon outer membrane complex 75-III); protein translocase	818	89.12	1	608.8	28	0.014	0.013	1	747.0	49	0.008	0.008
9	TOP PROTEIN	NP_186956.2	At3g03060	ATPase	628	69.51	1	541.4	30	0.019	0.019	1	513.8	58	0.013	0.012
10	TOP PROTEIN	NP_179673.1	At2g20800	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	582	65.31	1	487.1	17	0.011	0.012	1	1093.5	117	0.027	0.027
11	TOP PROTEIN	NP_565435.1	At2g18330	AAA-type ATPase family protein	636	71.13	1	424.0	36	0.023	0.022	1	765.7	105	0.023	0.022
12	TOP PROTEIN	NP_188616.1	At3g19820	DWF1 (DIMINUTO 1); catalytic	561	65.33	1	409.6	15	0.010	0.011	1	901.1	112	0.027	0.027
13	TOP PROTEIN	NP_567184.1	At4g00740	dehydration-responsive protein-related	600	67.49	1	385.4	13	0.009	0.009	1	549.7	36	0.008	0.008
14	TOP PROTEIN	NP_173055.1	At1g16030	HSP70B (heat shock protein 70B); ATP binding	646	70.85	0.999	347.1	20	0.012	0.012	0.999	475.3	52	0.011	0.011
15	TOP PROTEIN	NP_199783.1	At5g49720	KOR1 (KORRIGAN); hydrolase, hydrolyzing O-glycosyl compounds	621	69.13	1	229.3	9	0.006	0.006	1	243.2	15	0.003	0.003

Table 3.S3. Analysis of the GAUT1:GAUT7 complex-associating proteins: transcriptional co-expression with GAUT1 and GAUT7 and subcellular localization (predicted and proven) of the 12 putative GAUT1:GAUT7 complex-associating proteins identified by both False Discovery Rate (FDR) and Probability analyses.

Transcriptional co-expression data were compiled from several *Arabidopsis* gene co-expression databases available on-line. Either GAUT1 or GAUT7 was used as "bait" to retrieve the Pearson Correlation Coefficients (r-values) of the GAUT1 or GAUT7 co-expression with the other proteins. Highlighted in blue are negative r-values (negative correlations), in yellow are r-values ≥0.5, and in orange are r-values ≥0.7. Subcellular localization information was from the Subcellular Proteomic Database (http://suba.plantenergy.uwa.edu.au/).

Protein Group	Protein ID	Gene ID	Sequence Name	Transcriptional Co-expression with GAUT1 and GAUT7 (Pearson Correlation Coefficient)										Subcellular localization[f]	
				Expression Angler [a]		GeneCAT[b]		AthCoR[c]		CressExpress[d]		ATTED-II[e]			
				GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GFP-fusion	Mass Spectromery
1	NP_191672.1	At3g61130	GAUT1/LGT1 (Galacturonosyltransferase 1); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	Self	0.485	Self	0.600	Self	0.590	Self	0.684	Self	0.56	N/A	Golgi
2	NP_565893.1	At2g38650	GAUT7/LGT7 (Galacturonosyltransferase 7); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	0.485	Self	0.600	Self	0.590	Self	0.684	Self	0.56	Self	N/A	Golgi
3	NP_176170.1	At1g59610	ADL3 (ARABIDOPSIS DYNAMIN-LIKE 3)	-0.062	0.334	0.162	0.319			0.364	0.478			cytosol	plasma membrane, vacuole
4	NP_177766.1	At1g76400	ribophorin I family protein	0.629	0.790	0.230	0.412	0.297	0.740	0.368	0.568			N/A	ER
5	NP_172500.1	At1g10290	ADL6 (DYNAMIN-LIKE PROTEIN 6)	-0.326	-0.231	0.144	0.261			0.405	0.487			cell plate, Golgi	Plasma membrane, vacuole
6	NP_193537.2	At4g18030	dehydration-responsive family protein	0.57	0.68	0.205	0.354	0.36	0.586	0.658	0.696	0.44	0.57	N/A	extracellular, Golgi, vacuole
7	NP_193847.2	At4g21150	ribophorin II (RPN2) family protein	0.244	0.56	0.128	0.429			0.362	0.561			N/A	ER, extracellular, mitochondrion, plasma membrane, vacuole

(Table 3.S3. … continued)

Protein Group	Protein ID	Gene ID	Sequence Name	Transcriptional Co-expression with GAUT1 and GAUT7 (Pearson Correlation Coefficient)										Subcellular localization[f]	
				Expression Angler [a]		GeneCAT[b]		AthCoR[c]		CressExpress[d]		ATTED-II[e]			
				GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GAUT1 bait	GAUT7 bait	GFP-fusion	Mass Spectromery
8	NP_190258.1	At3g46740	TOC75-III (translocon outer membrane complex 75-III); protein translocase	0.295	0.310	0.204	0.308	0.346	0.340	0.576	0.593			N/A	plastid, vacuole
9	NP_186956.2	At3g03060	ATPase	-0.435	-0.137	0.226	0.279			0.217	0.495			N/A	extracellular
10	NP_179673.1	At2g20800	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	-0.558	-0.491	0.176	0.107			0.301	0.223			mitochondria	mitochondria, peroxisome
11	NP_565435.1	At2g18330	AAA-type ATPase family protein	0.149	0.691	0.234	0.353	0.264	0.485	0.248	0.520			N/A	N/A
12	NP_188616.1	At3g19820	DWF1 (DIMINUTO 1); catalytic	0.377	0.578	0.046	0.161	0.475	0.644	0.489	0.565			ER	plasma membrane, vacuole
13	NP_567184.1	At4g00740	dehydration-responsive protein-related	0.498	0.823	0.351	0.390	0.376	0.538	0.551	0.686	0.43	0.53	N/A	Golgi
14	NP_199783.1	At5g49720	KOR1 (KORRIGAN); hydrolase, hydrolyzing O-glycosyl compounds	0.436	0.304	-0.049	0.073			0.402	0.275			cell plate, ER, plasma membrane, vacuole	plasma membrane

[a]Expression Angler (http://www.bar.utoronto.ca/ntools/cgi-bin/ntools_expression_angler.cgi) - co-expression data query was performed on Botany Array Resource Data Set

[b]GeneCAT (Gene Co-expression Analysis Toolbox - http://genecat.mpg.de/)

[c]AthCoR (The *A. thaliana* Co-Response Database - http://csbdb.mpimp-golm.mpg.de/csbdb/dbcor/ath.html). Except for co-expression between GAUT1 and GAUT7 (analyzed using Multiple Gene Query (mGQ)), analysis was done using Intersection Gene Query (isGQ) using both GAUT1 and GAUT7 on matrix atge100 (developmental series (only WT); Ath1 chip; AtGenExpress; 12200 genes); coefficient was set for parametric Pearson's linear product-moment correlation; output was set for positive, co-responding genes with probability <0.05 (95%) and equal output for both genes.

[d]CressExpress (http://www.cressexpress.org/). A default linear regression value *r2* cut-off of 0.36 is suggested by this database, which is equivalent to an *r* value of 0.6 (Srinivasasainagendra et al. (2008) Plant Physiol 147(3):1004-1016).

[e]ATTED-II version 5.5 (*Arabidopsis thaliana* trans-factor and cis-element prediction database - http://atted.jp/). Only the top 300 co-expressed genes for each query are reported by this database.

DISCUSSION AND CONCLUSIONS

The plant cell wall is a complex and dynamic extracellular matrix constructed predominantly of carbohydrate, a major component of which is the structurally complex pectic polysaccharides. A comprehensive knowledge of how pectin is synthesized is essential to understand the overall organization of the plant cell wall, and to fully exploit pectin and plant cell wall for plant improvement and human use, which is especially relevant today for biofuel production from plant biomass. Progress in the field of plant cell walls and pectin biosyntheses has been slow, largely due to the difficulty in working with membrane-bound wall biosynthetic enzymes that are low in abundance, delicate to isolate in active forms, and notoriously hard to express as recombinant enzymes in heterologous systems. It is therefore not surprising that our understanding of pectin biosynthesis is still far from complete, despite all the research efforts in this field of study.

It has been suggested that the biosynthesis of plant cell wall polysaccharides, and of pectin in particular, may involve protein complexes of glycosyltransferases (Doong and Mohnen, 1998; Sterling et al., 2001; Sterling et al., 2006; York and O'Neill, 2008). The first reported plant cell wall polysaccharide biosynthetic complex was the plasma membrane-located rosette structures known to synthesize cellulose microfibrils. Cellulose synthase (CesA) proteins, thought to be the catalytic subunits of the cellulose synthase complex, have been shown to associate with the rosettes and to form homo- and/or heterocomplexes *in vivo* (see the section on cellulose biosynthesis in Chapter 2 of this dissertation). Nevertheless, the catalytic activity of the

CesAs has not been definitively demonstrated to date. A recent report described a putative glucuronoarabinoxylan (GAX) synthase complex in wheat (Zeng et al., 2010), suggesting that protein complexes may also be involved in hemicellulose biosynthesis. The proposed GAX synthase complex is thought to include putative glycosyltranferases from GT43, GT47, and GT75, which co-immunoprecipitate from detergent-solubilized membrane preparations in a complex that cooperatively exhibits xylosyl-, arabinosyl-, and glucuronyltransferase activities. Yet again, a definitive description of the complex is lacking and the catalytic function for each of the glycosyltransferases was not assigned.

The research presented in this dissertation is aimed at forwarding our understanding of pectin homogalacturonan biosynthesis through a study focused on two Arabidopsis glycosyltransferases, GAUT1 and GAUT7, which were previously shown to be present in a partially purified GalAT enzyme preparation (Sterling et al., 2006). This dissertation provides *in vitro* and *in vivo* biochemical evidence demonstrating conclusively that GAUT1 and GAUT7 engage in an HG:GalAT complex. This is the first definitive description of a protein complex involved in the biosynthesis of plant cell wall matrix polysaccharides. At least one of the complex components is known to have HG:GalAT catalytic activity (i.e. GAUT1) and another is assigned with a membrane-anchoring function (i.e. GAUT7). Transcript co-expression of *GAUT1* and *GAUT7* in most tissues and developmental stages of Arabidopsis plants is in support of the presence and importance of the GAUT1:GAUT7 complex. Further characterization of the complex implicates disulfide bonds in the complex formation, identifies mature Arabidopsis GAUT1 as a post-translationally processed protein that depends on GAUT7 for retention in the Golgi, and identifies other proteins that may interact non-covalently with the GAUT1:GAUT7 complex.

Interaction between GAUT1 and GAUT7

As described in Chapter 3 of this dissertation, co-immunoprecipitation and BiFC data demonstrate that Arabidopsis GAUT1 and GAUT7 engage in protein-protein interactions in a functional HG:GalAT enzyme complex. GAUT1 and GAUT7 resolve at ~60 and ~75 kDa in reducing SDS-PAGE, respectively, but are observed at a higher molecular weight upon non-reducing SDS-PAGE, clearly indicating that both GAUT1 and GAUT7 engage in protein complexes held together by disulfide bonds. Each protein contains 8 cysteine residues in its amino acid sequence. Intriguingly, the high molecular weight protein band of GAUT1 is of the same size as that of GAUT7 (i.e. ~185 kDa), suggesting that both protein bands correspond to a protein complex containing both GAUT1 and GAUT7 held together by intermolecular disulfide bond(s). Although one could argue that the ~185 kDa band might correspond to a mixture of homo-oligomers of GAUT1 and of GAUT7, this seems unlikely since the probability would be very low that GAUT1 and GAUT7, having different monomeric sizes, could each form homooligomers of exactly the same size. It is noted, however, that while homodimer/oligomer formation mediated by intermolecular disulfide bonds has been documented for many Golgi-localized glycosyltransferases (El-Battari et al., 2003; Young, 2004; Hassinen et al., 2010), there are very few reports of disulfide-bonded heterocomplexes. One such heterocomplex is the Arabidopsis secondary cell wall cellulose synthase complex (CSC) containing AtCesA4, AtCesA7, and AtCesA8 (Atanassov et al., 2009). It was shown that precipitation of AtCesA7 under non-reducing condition pulled down all the three CSC subunits, while the same experiment performed under reducing condition pulled down only AtCesA7. It is thus necessary to perform similar experiments in the future on the GAUT1:GAUT7 complex, to critically test the proposed presence of disulfide bond(s) between GAUT1 and GAUT7.

It is interesting to note that in the BiFC and fluorescent protein experiments using transient expression of the constructs in *N. benthamiana* leaves, there seemed to be no interactions between the introduced Arabidopsis GAUT1 and GAUT7 with the endogenous tobacco orthologs. For example, Arabidopsis GAUT1, fused to split YFP or to full-length GFP, could have been retained in the Golgi by a tobacco ortholog of GAUT7, and resulted in observable fluorescence signals. This is clearly not the case in our experiments, where no Golgi accumulation of Arabidopsis GAUT1-GFP was detected in the absence of Arabidopsis GAUT7. It is possible that the tobacco ortholog of GAUT7 does not have sufficient structural similarity with the Arabidopsis GAUT7 to enable the former to interact with the Arabidopsis GAUT1 in the transient expression experiment. Unfortunately, protein sequence comparison of the Arabidopsis GAUT1 and GAUT7 to the tobacco orthologs is not feasible at this point due to the lack of availability of the tobacco genome and protein sequences. BLAST analyses (as of December 3, 2010) of Arabidopsis GAUT1 and GAUT7 reveal no homologous protein reported from *N. benthamiana*, and only one protein from *Nicotiana tabacum*. This tobacco Avr9/Cf-9 Rapidly Elicited Protein 231 (accession no. AAG43554.1) contains 353 a.a., with sequence identity and similarity of 25% and 46%, respectively, with Arabidopsis GAUT1, and of 26% and 44%, respectively, with Arabidopsis GAUT7. Another possibility is that the GAUT1 and GAUT7 tobacco orthologs are expressed in very low levels, in comparison to the transiently over-expressed Arabidopsis GAUT1 and GAUT7. In this case, even if the tobacco orthologs could substitute for the Arabidopsis GAUTs, the resulting fluorescence might be too low to be considered as a positive signal. Considering that the Arabidopsis GAUT1:GAUT7 complex, and perhaps also the tobacco complex of the orthologous GAUTs, is presumably formed early in the secretory pathway (i.e. in the ER) and held covalently by disulfide bond(s), the chance of "mix

and match" between the Arabidopsis and the tobacco GAUTs is relatively slim, which may also help explain the observed experimental results.

This project is another clear illustration of the difficulty of working with wall biosynthetic enzymes. The very low abundance of the GAUT1:GAUT7 complex made it necessary to perform affinity purification of the complex multiple times to obtain enough material for N-terminal sequencing of mature Arabidopsis GAUT1 and for proteomics analyses of the complex. These analyses also showed that affinity purification, despite using anti-GAUT1 and anti-GAUT7-specific IgGs purified from the polyclonal sera using peptide-antigen columns, still yielded an unexpectedly complex mixture of proteins, which obliged us to be rigorous in the experimental approach and in interpreting the data. For example, we propose RANELVQ as the N-terminal sequence of GAUT1 and the cleavage site between Met_{167} and Arg_{168}, based on the N-terminal sequencing results (see Appendix C for more details), on the proteomics sequence coverage of GAUT1 (Chapter 3, Fig. 3.S5), and on the expression of GFP fusion constructs of C-terminally truncated GAUT1 in the presence and absence of GAUT7 (Chapter 3, Figs. 3.3 and 3.S3).

The data did not allow us to determine in which precise subcellular location Arabidopsis GAUT1 is post-translationally cleaved or why it is cleaved. One possibility is that the cleavage serves to activate the GalAT enzymatic activity of GAUT1, a mechanism that has been observed in the activation and extracellular targeting of the type-I pectin methylesterases in Arabidopsis and tobacco (Wolf et al., 2009). In this event, GAUT1 would be synthesized as a full-length, inactive pre-protein. Proteolytic cleavage, which could occur upon arrival of the GAUT1 pre-protein in the Golgi, would convert GAUT1 into a mature, truncated protein that is catalytically active. Another possibility is that the proteolytic cleavage may be required for the interaction of

GAUT1 with GAUT7. This scenario, however, may not be compatible with the above scenario of the activation of GAUT1 catalytic activity, since the formation of GAUT1:GAUT7 complex is likely to occur in the ER, especially if GAUT1 and GAUT7 are indeed connected by disulfide bond(s). While it remains possible, enzymatic activation of GAUT1 by cleavage in the ER would appear premature, since HG synthesis occurs in the Golgi (Zhang and Staehelin, 1992; Staehelin and Moore, 1995; Sterling et al., 2001).

The post-translational processing of GAUT1 into the mature Arabidopsis GAUT1 lacking its transmembrane domain poses an issue for GAUT1 retention in the Golgi apparatus, where GAUT1 is expected to function in homogalacturonan synthesis. However, many proteins that do not have transmembrane domains, or are predicted not to have one, are targeted to the secretory pathway but found to reside in intracellular compartments such as the ER or Golgi apparatus, instead of being secreted out of the cells. For example, all the GATL proteins are predicted to have only signal peptides, however Arabidopsis and poplar GATL1/PARVUZ proteins have been shown to localize to the ER and Golgi apparatus (Lee et al., 2007; Kong et al., 2009), indicating that a mechanism exists to retain the supposedly soluble proteins in the intracellular compartments. A similar mechanism of interaction with anchor proteins occurs when the cleaved GAUT1 interacts with GAUT7 and is thereby retained in the Golgi.

Homo- or hetero-oligomerization as a means for Golgi retention has been described for many other type II transmembrane glycosyltransferases, some of which have been summarized in several recent reviews (El-Battari et al., 2003; de Graffenried and Bertozzi, 2004; Young, 2004). Different domains, or combinations thereof, mediate the oligomer formation of different glycosyltransferases. For instance, human N-glycosylation enzymes N-acetylglucosaminyltransferase I (GlcNAcTI) and mannosidase II (MannII) associate via their

stem regions mediated by charged residues in the stem domain of GlcNAcTI, which localizes both proteins in the medial Golgi (Nilsson et al., 1994; Nilsson et al., 1996). Targeting of one of these proteins to the ER, by grafting the cytoplasmic domain of an ER protein p33, relocated both proteins to the ER and thus proved the interaction between them (Nilsson et al., 1994). Additionally, GlcNAcTI can also form a complex, via its luminal domain, with N-acetylglucosaminyltransferase II (GlcNAcTII) (Opat et al., 2000). In the case of the GAUT1:GAUT7 complex, protein-protein interaction is likely to occur also via the luminal domains of GAUT1 and GAUT7. The transient expression of GFP fused to the first 100 and 291 a.a. of GAUT1 in the presence or absence of GAUT7 (see Chapter 3 of this dissertation), established that the region beyond a.a. position 291, which presumably corresponds to the globular catalytic domain, is necessary for complex formation. More detailed GAUT1 truncation studies will be required to determine if the GAUT1 protein region between the N-terminal Met_{167} and Arg_{291} may also be involved in the GAUT1:GAUT7 protein interaction, and to establish the exact nature of complex formation (e.g. possible involvement of certain charged residues in ionic interactions and/or Cys residues in disulfide bonding).

Comparison of GAUT1:GAUT7 complex with other functionally significant glycosyltransferase complexes

In addition to retention in the Golgi, protein complex formation may also serve other functions with regards to regulation and efficiency of the enzymatic activities catalyzed by the protein complex components. Below, several reported protein-protein interactions involving ER- and Golgi-localized glycosyltransferases are summarized, and the possible extrapolation of the

unique functions of the complexes are considered in view of HG:GalAT activity of the GAUT1:GAUT7 complex.

The mammalian core 1 β-1,3-galactosyltransferase (C1β1,3GalT, also known as T-synthase) is the enzyme that catalyzes transfer of Gal from UDP-Gal onto the GalNAc that is attached to the Ser or Thr residues of O-glycosylated proteins (Ju et al., 2002). This enzyme is known to form a disulfide-bonded homodimer, as observed by non-reducing SDS-PAGE (Ju et al., 2002), and localized to the Golgi apparatus, as shown by immunofluorescence microscopy and membrane fractionation (Ju et al., 2008). To be correctly folded and catalytically active, C1β1,3GalT requires interaction with the ER-localized Cosmc (core 1 β-1,3-GalT-specific molecular chaperone) (Ju et al., 2002; Ju et al., 2008). Cosmc is a type II transmembrane protein with 22% amino acid sequence identity to C1β1,3GalT, however Cosmc does not display catalytic activity as C1β1,3GalT does (Kudo et al., 2002; Ju et al., 2008). C1β1,3GalT and Cosmc were co-purified from a rat liver preparation and co-precipitated from Jurkat cells co-expressing epitope-tagged human Cosmc and C1β1,3GalT, establishing the protein-protein interaction between the two proteins (Ju et al., 2002; Ju et al., 2002). In the absence of functional Cosmc, the inactive, misfolded C1β1,3GalT is retained within the ER lumen, where it forms non-productive disulfide-bonded oligomers that are subsequently targeted for proteosomal degradation (Ju et al., 2002; Ju et al., 2008).

The chaperone function of Cosmc was clearly evidenced when *in vitro* incubation of heat- or chemically denatured C1β1,3GalT with Cosmc resulted in refolding of C1β1,3GalT and restoration of its activity (up to 75% of initial activity for heat-denatured protein) (Aryal et al., 2010). Cosmc appears to act as a chaperone specifically for C1β1,3GalT, since a non-specific ER chaperone BiP could not refold denatured C1β1,3GalT and Cosmc could not refold

denatured β-1,4-galactosyltransferase (β4-GalT), another Golgi-localized enzyme (Aryal et al., 2010). It is highly unlikely, however, that the subunits of the GAUT1:GAUT7 complex have the same relative roles in the complex as C1β1,3GalT and Cosmc do in the C1β1,3GalT/Cosmc complex. Cosmc is localized in the ER (Ju et al., 2008), and seems to associate only with the unfolded form of C1β1,3GalT (Ju et al., 2002). GAUT7 does not appear to function as a chaperone for GAUT1, since GAUT7 itself is localized to the Golgi, and its interaction with GAUT1 is clearly not transient as would be expected for a chaperone.

Another unique protein-protein interaction is that of the mammalian protein *O*-mannosyltransferase 1 (POMT1) and its homolog POMT2, which together form a heteromeric complex capable of catalyzing the first step in protein *O*-mannosylation by transferring a Man residue from dolichyl phosphate mannose (Dol-P-Man) onto Ser/Thr residues of certain proteins (e.g. α-dystroglycan) (Akasaka-Manya et al., 2006). Human POMT1 and POMT2 share 36% a.a. sequence identity, and both are predicted to be integral membrane proteins containing seven transmembrane spanning domains and are co-localized in the ER, where protein *O*-mannosylation is initiated (Willer et al., 2002; Akasaka-Manya et al., 2006). Co-expression of both proteins is required for enzymatic activity, since each protein alone does not have enzymatic activity (Manya et al., 2004; Manya et al., 2006). POMT1 and POMT2, as well as POMT activity, co-immunoprecipitate from solubilized membrane fractions prepared from HEK293 cells co-expressing both proteins (Akasaka-Manya et al., 2006). However, expression of the proteins separately followed by extraction, solubilization, and subsequent mixing *in vitro*, does not result in co-immunoprecipitation or detection of POMT activity. These results clearly indicate that the *in vivo* interaction between POMT1 and POMT2 has to occur, not only to form a functional enzyme, but also to physically form the heterocomplex in the first place. While it

remains unclear why co-expression of POMT1 and POMT2 is necessary for POMT activity, it has been suggested that the POMT1/POMT2 complex formation may create a novel catalytic domain, or that one protein may act as the catalytic subunit while the other may take on a regulatory role (Akasaka-Manya et al., 2006; Manya et al., 2006). With regards to the GAUT1:GAUT7 complex, the POMT1/POMT2 scenario does not seem applicable, since GAUT1 by itself has been shown to have GalAT activity upon transient expression in HEK293 cells (Sterling et al., 2006). However, as in the case of the POMT1/POMT2 complex, the possibility remains that formation of the GAUT1:GAUT7 complex may create a new catalytic site or may increase GalAT activity, or that GAUT7, in addition to its membrane-anchoring function, may also act as a regulatory molecule in HG synthesis.

The protein complex between the tumor suppressors EXT1 and EXT2 illustrates yet another functional interaction between protein complex subunits. EXT1 and EXT2 are both implicated in hereditary multiple exostoses, a genetic disorder characterized by cartilaginous tumor formation at the growth plate of long bones (osteochondromas/exostoses) (Nadanaka and Kitagawa, 2008). The two proteins are involved in the polymerization of heparan sulfate chains, with each protein exhibiting a dual enzymatic activity capable of incorporating both α-1,4-GlcNAc and β-1,4-GlcA residues from UDP-GlcNAc and UDP-GlcA, respectively, in an alternating manner onto the growing polysaccharide chains (Senay et al., 2000). It was noted, however, that EXT2 is less active in comparison to EXT1 and, in a certain experimental settings, may have only a low level of GlcAT activity and virtually no GlcNAcT activity (McCormick et al., 2000; Senay et al., 2000). The proteins encoded by *EXT1* and *EXT2* are predicted to be type II transmembrane proteins with ~35% homology to each other. Individual expression of each gene localizes the encoded protein to the ER (McCormick et al., 2000; Wei et al., 2000).

Interestingly, co-expression of EXT1 and EXT2 shifts the subcellular localization of both proteins to the Golgi apparatus, and dramatically improves the GlcNAcT/GlcAT activity to significantly higher levels than the activities of the individual components combined (Kobayashi et al., 2000; McCormick et al., 2000; Senay et al., 2000). Using dual-tagging expression systems followed by immunoprecipitation, it was further shown that EXT1 and EXT2 each can form homodimers *in vivo*, and upon co-expression both proteins form heterodimers and co-immunoprecipitate. No co-immunoprecipitation, and thus no heterodimer formation, nor augmentation of the GlcNAcT/GlcAT enzymatic activity, were observed when lysates from cells separately transfected with EXT1 and EXT2 constructs were mixed *in vitro* and subjected to the same analyses.

The same scenario as EXT1 and EXT2 is also documented for the complex between the 23% identical chondroitin sulfate synthase-1/chondroitin synthase-1 (CSS1/ChSy-1) and chondroitin sulfate synthase-2/chondroitin polymerizing factor (CSS2) (Ogawa et al., 2010). Both are type II transmembrane proteins capable of forming homo- and heterodimers *in vivo*. In this case, expression of recombinant soluble CSS1/ChSy-1 alone exhibited a dual GlcAT-II/GalNAcT-II activity, but yielded no chondroitin sulfate chain elongation *in vitro* (Kitagawa et al., 2003). Only upon co-expression of CSS1/ChSy-1 and CSS2 together did the level of GlcAT-II/GalNAcT-II activity improve significantly and result in *in vitro* polymerization of the glycosaminoglycan chain.

Both the EXT1/EXT2 and CSS1/CSS2 complexes clearly underscore the importance of protein-protein interactions *in vivo*, and indicate that other, as yet to be determined, factors may transiently facilitate the productive formation of functional complexes. This latter point is supported by the fact that the high activity present in the native biological systems cannot yet be

replicated in the heterologous expression systems used in the above-mentioned studies. The GAUT1:GAUT7 complex could function similarly to the EXT1/EXT2 and CSS1/CSS2 complexes, i.e. the GAUT1:GAUT7 enzyme complex has substantially higher catalytic and/or polymerizing activities in comparison to the individual proteins. Previous research in our laboratory has shown that transient expression in HEK293 cells of the recombinant soluble GAUT1, but not of the recombinant GAUT7, yielded low, but detectable HG:GalAT activity (Sterling et al., 2006). These results could be explained in several possible ways. (1) GAUT7 indeed does not have HG:GalAT activity. (2) GAUT7 does have HG:GalAT activity, however the transient expression system employed in the prior study was not suitable, or not sufficiently optimal, to produce an active recombinant GAUT7. It is well known that a heterologous expression system that works successfully for one protein may not necessarily work for another protein. (3) GAUT7 does have HG:GalAT activity, but is considerably less active compared to GAUT1. In this case, the HG:GalAT assay method employed in the reported study may not have been sensitive enough to detect the very low level of HG:GalAT activity of the recombinant GAUT7. This is a reasonable possibility, considering that the reported transient expression of recombinant GAUT1 resulted in only a low level of HG:GalAT activity (Sterling, 2004), leaving a relatively narrow window to assess the perhaps negligible HG:GalAT activity of the recombinant GAUT7. In all three cases, the interaction of GAUT1 and GAUT7 in the Arabidopsis GAUT1:GAUT7 complex may significantly increase the HG:GalAT activity of the enzyme complex in comparison to that of GAUT1 alone. Co-expression of recombinant GAUT1 and GAUT7 in a suitable heterologous expression system may allow functional *in vivo* complex formation, and result in an augmentation of the HG:GalAT activity of the GAUT1:GAUT7 complex.

GAUT7 may not be the only membrane anchor for GAUT1

Promoter:*GUS* reporter gene experiments clearly indicate that *GAUT1* and *GAUT7* co-express in most tissues across different developmental stages, suggesting that the GAUT1:GAUT7 complex occurs ubiquitously in Arabidopsis plants. However, it was noticed that, during a particular developmental stage of the inflorescence, only *GAUT1* transcript expression occurs in the papillae of the stigma (Fig. F-1 in Appendix F). It appears that the expression of *GAUT1* in papillae occurs when the pistil is elongating to reach the anthers to allow pollination to occur. *GAUT1* expression was not present when the pistil had already elongated past the height of the anthers, presumably after pollination and fertilization have taken place, which may correspond to the stage of silique development. This exclusive tissue- and developmental-specific expression of *GAUT1*, in the absence of *GAUT7* expression, suggests that GAUT1 may not always exist in a complex with, and be anchored to the Golgi membrane by, GAUT7. If we assume that GAUT1 is always post-translationally processed into the mature protein without its transmembrane domain, then there should be other protein partner(s) that GAUT1 may interact with to ensure its Golgi retention.

In support of the above notion, preliminary data from BiFC experiments indicate that GAUT5, another member of the GAUT gene family, can also function to anchor GAUT1 to the Golgi membrane (Y. Sakuragi, Pers. Comm.). It is intriguing that the bioinformatics analysis by Yin and co-workers suggests that not only GAUT7, but also GAUT5 and 15, are likely non-catalytic due to amino acid substitutions in a proposed catalytic residue (Yin et al., 2010). GAUT5, therefore, is a good candidate for further study as another GAUT1 interacting partner in a protein complex. A possible reason for the multi-partnering of GAUT1 could be similar to what has been shown for the primary wall cellulose synthase complex. In that case, CESA1 and

CESA3 are the essential subunits, while CESA2, 5, 6, and 9 are partially functionally, but not completely redundant, and may associate with cellulose deposition at different developmental stages (Desprez et al., 2007; Persson et al., 2007). Interaction between GAUT1 and GAUT7 or GAUT5 could therefore be a way to target GAUT1 expression to a certain place within the plant where HG synthesis is needed. It is also possible that the interaction of GAUT1 with different partners may affect the level or type of GalAT activity and/or the way HG is synthesized, so as to give rise to different pectin and/or wall structures. It will be interesting to see if GAUT5 and GAUT7 are functionally or only partially redundant in regards to when and where they are expressed in tissues and developmental stages. Unfortunately, this could not easily be assessed from publicly available databases since *GAUT5* is not yet represented in the Affymetrix ATH1 GeneChip.

Possible roles of the GAUT1:GAUT7 complex-putative interacting proteins

Proteomics analyses of the proteins pulled down by anti-GAUT1 and anti-GAUT7 antibodies, identified 12 proteins that may interact with the GAUT1:GAUT7 complex. Several of these proteins and their possible roles with regards to HG biosynthesis have been discussed in Chapter 3 of this dissertation. This section offers additional discussion on this topic.

As mentioned in Chapter 3, the two dehydration-responsive proteins are putative methyltransferases in the 29-member Arabidopsis QUA2-related gene family (Krupkova et al., 2007; Mouille et al., 2007). Mouille and co-workers reported co-expression pairing between members of the GAUT and QUA2-related gene families. Specifically, *GAUT1*, *GAUT8*, and *GAUT9* are each co-expressed with two *QUA2* homologs. Interestingly, the putative methyltransferases identified in our proteomics study (i.e. At4g18030 and At4g00740) are not

the QUA2 homologs proposed to be co-expressed with GAUT1 (i.e. At1g26850 and At4g14360), although At4g18030 is the closest homolog of At1g26850 (69% identity). *At4g18030* was reported to co-express with primary wall cellulose synthases (*AtCesA1, 3*, and *6*) and *KOR1* (see below) (Persson et al., 2005), suggesting a role in primary wall synthesis.

It is intriguing that the ribophorin I and II family proteins were identified as putative interacting proteins of the GAUT1:GAUT7 complex. Little information is available on these Arabidopsis proteins. Both have been localized to the ER by a proteomics approach (Dunkley et al., 2006), and *haples6*, a T-DNA insertion mutant of the ribophorin II family protein (At1g76400), has been reported to have defective pollen tube growth (Johnson et al., 2004). Mammalian ribophorin I and II, as well as their yeast homologs (Ost1 and Swp1, respectively), are the subunits of the oligosaccharyltransferase (OST) complex which functions in the protein N-glycosylation pathway, where they are thought to have non-catalytic functions (Kelleher and Gilmore, 2006). Nevertheless, the identification of the Arabidopsis homologs in our proteomics study may suggest a possible novel mechanism in HG and pectin biosynthesis, which includes *en bloc* transfer of oligosaccharides onto growing polymer chains. Alternatively, since mammalian ribophorin I has recently been shown to deliver selected potential substrates to the catalytic core of OST (Wilson et al., 2008), its GAUT1:GAUT7 complex-associated homolog (At1g76400) may function to regulate delivery of precursor proteins or growing poly/oligosaccharides to the catalytic core of the GalAT complex.

The identification of KOR1 as one of the GAUT1:GAUT7 complex putative interacting protein was surprising, since it is believed to be involved in cellulose biosynthesis. Recombinant KOR1 from *Brassica napus*, heterologously expressed in *Pichia pastoris*, has endo-1,4-β-glucanase activity towards carboxymethylcellulose and soluble, amorphous cellulose (Molhoj et

al., 2001). In *Agrobacterium tumefaciens*, an endoglucanase gene (*celC*) is present in the operon that contains the cellulose synthase gene (*celA*) and is required for cellulose synthesis in this bacterium (Matthysse et al., 1995). It is thought that KOR1 may play the same role as the *Agrobacterium* endoglucanase in cellulose synthesis. Nevertheless, the exact function of an endoglucanase in cellulose synthesis remains unclear, although it has been suggested that KOR1 may function to remove non-crystalline glucan chains or to release tension during cellulose synthesis (Molhoj et al., 2002).

Mutants of KOR1 have a significant reduction in crystalline cellulose (Molhoj et al., 2002), and the protein is co-expressed with the primary wall CesAs (Persson et al., 2005). On the other hand, KOR1 has been observed to localize subcellularly not only at the plasma membrane, where it supposedly assists in cellulose synthesis, but also in endosomes and the Golgi apparatus (Somerville, 2006). It is therefore possible that KOR1 has other function(s) intracellularly. The identification of KOR1 as a putative interacting protein of the GAUT1:GAUT7 HG:GalAT complex also suggests a potentially more intimate connection than previously thought, between the syntheses of cellulose and pectin. In support of this idea, it is noteworthy that reduction in cellulose content in the KOR1 mutants is accompanied by an increased pectin content in the wall (Molhoj et al., 2002). It has also been reported recently that the degree of methylesterification and distribution of pectin in the wall, which are determined by the actions of pectin methylesterase, polygalacturonase, and pectin-synthesizing enzyme activities, play roles in the deposition of cellulose microfibrils in conjunction with cortical microtubules (Yoneda et al., 2010).

Arabidopsis DIMINUTO/DWARF1 (DWF1) is involved in steroid biosynthesis. This integral membrane protein is thought to catalyze the conversion of 24-methylenecholesterol to

campesterol in the brassinosteroid biosynthetic pathway, and is most probably localized to the ER (Klahre et al., 1998; Choe et al., 1999). Genes involved in brassinosteroid synthesis were found to be co-expressed with the primary wall cellulose synthases *AtCesA1, 3*, and *6*, suggesting a co-regulation of the brassinosteroid synthesis pathway and primary wall biosynthesis (Persson et al., 2005), in which GAUT1 and GAUT7 are clearly involved.

DRP2A/ADL6 and DRP2B/ADL3 are two dynamin-related proteins that have been shown to localize to the cell plates during cytokinesis and to the Golgi, and are thought to be involved in clathrin-coated vesicular trafficking (Jin et al., 2001; Lam et al., 2002; Fujimoto et al., 2008). It is possible that they may have a role in vesicle trafficking of HG from the Golgi apparatus to the cell plate during cytokinesis, or to the extracellular space for incorporation into the cell wall. It is well established that different pectic epitopes are not distributed uniformly throughout the wall, as demonstrated by immunolabelling of plant cell walls using monoclonal antibodies raised against different pectic polysaccharides (Willats et al., 2001). It is tempting to speculate that the GAUT1:GAUT7 complex may synthesize a particular HG domain, which needs to be delivered to a specific site on the plasma membrane for incorporation into the wall. In this case, DRP2A/ADL6 and DRP2B/ADL3, by interacting with the GAUT1:GAUT7 complex, may have a role in regulating vesicular transport to ensure correct delivery of the HG domain for proper wall formation.

Connections between the GAUT1:GAUT7 complex and the four remaining putative interacting proteins are unclear. Two ATPases (At3g03060 and At2g18330) and the NAD(P)H dehydrogenase NDB4 (At2g20800) might play roles in providing energy that may be required during HG/pectin synthesis. TOC75-III, however, is one of the subunits of the translocon complexes at the outer chloroplast envelope membrane (Toc), which assist the importation of

pre-proteins into the chloroplast (Agne and Kessler, 2009). Interactions between these four proteins and the GAUT1:GAUT7 complex need to be validated, and their roles in pectin synthesis remain to be determined.

Conclusions

The research presented in this dissertation is a major step forward in our endeavor to understand how the complex structures of the plant cell wall are synthesized. Identification and characterization of the Arabidopsis GAUT1:GAUT7 HG:GalAT complex demonstrates for the first time the involvement of a functionally confirmed protein complex in the biosynthesis of plant cell wall matrix polysaccharides. This work also introduces several novel concepts in the plant wall biosynthesis research field, including the anchoring of a post-translationally processed, catalytically active glycosyltransferase to the Golgi membrane by its homolog, and the possible transfer of oligosaccharides *en bloc* onto growing polymer chains in pectin biosynthesis.

As is typical of scientific research, the results from this project lead to more questions that need to be answered. A more detailed characterization of the GAUT1:GAUT7 complex is certainly warranted, such as the determination of the sites of protein-protein interactions and the functional significance of the complex with regards to the HG:GalAT activity. The predicted tertiary structure of Arabidopsis GAUT1 indicates regions corresponding to segments in the resolved structure of rabbit glycogenin that are thought to be involved in homodimer formation (Gibbons et al., 2002; Yin et al., 2010). In the GAUT1:GAUT7 complex, these GAUT1 regions may be involved in the interaction with GAUT7 or with GAUT1 itself. It would be interesting to

see, for example, whether these regions contain any Cys residues that may engage in intermolecular disulfide bonds between GAUT1 and GAUT7.

Another important question is whether the presence of GAUT7 in the GAUT1:GAUT7 complex contributes to the GalAT activity of the complex. As mentioned earlier, GAUT7 may have a yet undetermined catalytic activity, or a regulatory role in HG polymerization by the complex. The GAUT1:GAUT7 complex may also have enhanced GalAT activity in comparison to that of GAUT1 alone. It is difficult to examine these possibilities in the immunoprecipitated native Arabidopsis complex, since GAUT1 and GAUT7 appear to always be in a complex with each other. Therefore, to test these hypotheses, it is essential to establish a heterologous expression system where GAUT1 and GAUT7 can be expressed in their active forms, either individually or together. Successful heterologous expression will also make it feasible to do site directed mutagenesis of residues predicted to be involved in substrate binding, catalysis, and protein-protein interaction, and to perform detailed biochemical characterization of the GalAT activity of the GAUT1:GAUT7 complex without possible interference from contaminating proteins that may be present in the immunoprecipitated native Arabidopsis complex.

It is also important to investigate the functional significance of the post-translational cleavage of GAUT1. One necessary experiment, the result of which may strengthen the conclusion of GAUT1 dependency on GAUT7 for Golgi retention, is a BiFC of GAUT1 and N-terminally truncated (i.e. without the transmembrane domain) GAUT7. This experiment could serve as a negative control for the BiFC data presented in this dissertation. Another possibility is to introduce mutations at the proposed cleavage site, express the resulting constructs by BiFC and/or fluorescence protein experiments as described in this dissertation, and observe whether the mutated GAUT1 remains intact and resides in the Golgi in the absence of GAUT7.

The identification of the GAUT1:GAUT7 complex-putative interacting proteins requires verification by other experimental methods. It would be ideal to perform reciprocal immunoprecipitation using antibodies against the putative interacting proteins, and verify the presence of GAUT1 and GAUT7 in the pulled-down proteins by mass spectrometry analyses. Polyclonal antibodies against several of the putative interacting proteins have been reported in the literature, such as those for KOR1 (Nicol et al., 1998; Szyjanowicz et al., 2004) and ADL6 (Lee et al., 2002), which could potentially be obtained and used to carry out these immunoprecipitations. However, no antibodies are available for most of the putative interacting proteins, and to make these antibodies would be very costly and time-consuming, making this approach not very feasible. A more feasible approach would be to perform BiFC and/or yeast two-hybrid analyses of these proteins. Between these two methods, BiFC is more appealing since it can give *in planta* evidence of the interactions, and can allow detection of indirect interactions between the putative interacting proteins and GAUT1 and/or GAUT7, while the yeast two-hybrid system can only show direct interactions between the bait and prey proteins. Once the association between the putative interacting proteins and the GAUT1:GAUT7 complex is verified, the roles of these proteins in pectin synthesis will need to be ascertained.

REFERENCES

Abdel-Massih RM, Rizkallah HD, Al-Din RS, Baydoun EA, Brett CT (2007) Nascent pectin formed in Golgi apparatus of pea epicotyls by addition of uronic acids has different properties from nascent pectin at the stage of galactan elongation. J Plant Physiol **164:** 1-10

Abu-Qarn M, Eichler J, Sharon N (2008) Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. Curr Opin Struct Biol **18:** 544-550

Agne B, Kessler F (2009) Protein transport in organelles: The Toc complex way of preprotein import. FEBS J **276:** 1156-1165

Akasaka-Manya K, Manya H, Nakajima A, Kawakita M, Endo T (2006) Physical and functional association of human protein *O*-mannosyltransferases 1 and 2. J Biol Chem **281:** 19339-19345

Akita K, Ishimizu T, Tsukamoto T, Ando T, Hase S (2002) Successive glycosyltransfer activity and enzymatic characterization of pectic polygalacturonate 4-α-galacturonosyltransferase solubilized from pollen tubes of *Petunia axillaris* using pyridylaminated oligogalacturonates as substrates. Plant Physiol **130:** 374-379

Albersheim P, Darvill A, Roberts K, Staehelin LA, Varner JE (1997) Do the structures of cell wall polysaccharides define their mode of synthesis? Plant Physiol **113:** 1-3

Almeida R, Levery SB, Mandel U, Kresse H, Schwientek T, Bennett EP, Clausen H (1999) Cloning and expression of a proteoglycan UDP-galactose:β-xylose β1,4-galactosyltransferase I. A seventh member of the human β4-galactosyltransferase gene family. J Biol Chem **274:** 26165-26171

Almond A (2007) Hyaluronan. Cell Mol Life Sci **64:** 1591-1596

Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP (1995) A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. Proc Natl Acad Sci U S A **92:** 9353-9357

Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga KS (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (*CesA*) gene family. Cellulose **11:** 287-299

Arend M, Muninger M, Fromm J (2008) Unique occurrence of pectin-like fibrillar cell wall deposits in xylem fibres of poplar. Plant Biol (Stuttg) **10:** 763-770

Arioli T, Peng L, Betzner AS, Burn J, Wittke W, Herth W, Camilleri C, Hofte H, Plazinski J, Birch R, Cork A, Glover J, Redmond J, Williamson RE (1998) Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science **279:** 717-720

Aryal RP, Ju T, Cummings RD (2010) The endoplasmic reticulum chaperone Cosmc directly promotes *in vitro* folding of T-synthase. J Biol Chem **285:** 2456-2462

Atanassov, II, Pittman JK, Turner SR (2009) Elucidating the mechanisms of assembly and subunit interaction of the cellulose synthase complex of Arabidopsis secondary cell walls. J Biol Chem **284:** 3833-3841

Bacic A, Harris PJ, Stone BA (1988) Structure and function of plant cell walls. *In* J Preiss, ed, The biochemistry of plants, Vol 14. Academic Press, Inc., New York, pp 297-371

Bai X, Zhou D, Brown JR, Crawford BE, Hennet T, Esko JD (2001) Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the β1,3-galactosyltransferase family (β3GalT6). J Biol Chem **276:** 48189-48195

Ball SG, Morell MK (2003) From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annu Rev Plant Biol **54:** 207-233

Baris I, Tuncel A, Ozber N, Keskin O, Kavakli IH (2009) Investigation of the interaction between the large and small subunits of potato ADP-glucose pyrophosphorylase. PLoS Comput Biol **5:** e1000546

Baron-Epel O, Gharyl PK, Schindler M (1988) Pectins as mediators of wall porosity in soybean cells. Planta **175:** 389-395

Behall K, Reiser S (1986) Effects of pectin on human metabolism. *In* ML Fishman, JJ Jen, eds, Chemistry and function of pectins. American Chemical Society, Washington, DC, pp 248-465

Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M (1999) Functions of cell surface heparan sulfate proteoglycans. Annu Rev Biochem **68:** 729-777

Bessueille L, Sindt N, Guichardant M, Djerbi S, Teeri TT, Bulone V (2009) Plasma membrane microdomains from hybrid aspen cells are involved in cell wall polysaccharide biosynthesis. Biochem J **420:** 93-103

Bethea HN, Xu D, Liu J, Pedersen LC (2008) Redirecting the substrate specificity of heparan sulfate 2-*O*-sulfotransferase by structurally guided mutagenesis. Proc Natl Acad Sci U S A **105:** 18724-18729

Bishop JR, Schuksz M, Esko JD (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature **446:** 1030-1037

Blennow A, Nielsen TH, Baunsgaard L, Mikkelsen R, Engelsen SB (2002) Starch phosphorylation: a new front line in starch research. Trends Plant Sci **7:** 445-450

Blumenkrantz N, Asboe-Hansen G (1973) New method for quantitative determination of uronic acids. Anal Biochem **54:** 484-489

Bodevin-Authelet S, Kusche-Gullberg M, Pummill PE, DeAngelis PL, Lindahl U (2005) Biosynthesis of hyaluronan: direction of chain elongation. J Biol Chem **280:** 8813-8818

Boehlein SK, Shaw JR, Hannah LC, Stewart JD (2010) Probing allosteric binding sites of the maize endosperm ADP-glucose pyrophosphorylase. Plant Physiol **152:** 85-95

Boehlein SK, Shaw JR, Stewart JD, Hannah LC (2008) Heat stability and allosteric properties of the maize endosperm ADP-glucose pyrophosphorylase are intimately intertwined. Plant Physiol **146:** 289-299

Boehlein SK, Shaw JR, Stewart JD, Hannah LC (2009) Characterization of an autonomously activated plant ADP-glucose pyrophosphorylase. Plant Physiol **149:** 318-326

Bolwell GP, Dalessandro G, Northcote DH (1985) Decrease of polygalacturonic acid synthase during xylem differentiation in sycamore. Phytochemistry **24:** 699-702

Bourlard T, Bruyant-Vannier MP, Schaumann A, Bruyant P, Morvan C (2001) Purification of several pectin methyltransferases from cell suspension cultures of flax (*Linum usitatissimum* L.). C R Acad Sci III **324:** 335-343

Bourlard T, Pellerin P, Morvan C (1997) Rhamnogalacturonan I and II are pectic substrates for flax-cell methyltransferases. Plant Physiol. Biochem. **35:** 623-629

Bourlard T, Schaumann-Gaudinet A, Bruyant-Vannier MP, Morvan C (1997) Various pectin methyltransferase activities with affinity for low and highly methylated pectins. Plant Cell Physiol **38:** 259-267

Bouton S, Leboeuf E, Mouille G, Leydecker MT, Talbotec J, Granier F, Lahaye M, Hofte H, Truong HN (2002) *QUASIMODO1* encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in Arabidopsis. Plant Cell **14:** 2577-2590

Bowles DJ, Northcote DH (1972) The sites of synthesis and transport of extracellular polysaccharides in the root tissues of maize. Biochem J **130:** 1133-1145

Bowling AJ, Brown RM, Jr. (2008) The cytoplasmic domain of the cellulose-synthesizing complex in vascular plants. Protoplasma **233:** 115-127

Brecht M, Mayer U, Schlosser E, Prehm P (1986) Increased hyaluronate synthesis is required for fibroblast detachment and mitosis. Biochem J **239:** 445-450

Brett CT (2000) Cellulose microfibrils in plants: biosynthesis, deposition, and integration into the cell wall. Int Rev Cytol **199:** 161-199

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell **17:** 2281-2295

Brown RM, Jr. (2004) Cellulose structure and biosynthesis: what is in store for the 21st century? J. Polymer Sci. A **42:** 487-495

Brown RM, Jr., Franke WW, Kleinig H, Falk H, Sitte P (1970) Scale formation in chrysophycean algae. I. Cellulosic and noncellulosic wall components made by the Golgi apparatus. J Cell Biol **45:** 246-271

Brown RM, Jr., Montezinos D (1976) Cellulose microfibrils: visualization of biosynthetic and orienting complexes in association with the plasma membrane. Proc Natl Acad Sci U S A **73:** 143-147

Brown RM, Jr., Willison JH, Richardson CL (1976) Cellulose biosynthesis in *Acetobacter xylinum*: visualization of the site of synthesis and direct measurement of the *in vivo* process. Proc Natl Acad Sci U S A **73:** 4565-4569

Brown RMJ (1996) The biosynthesis of cellulose. Pure Appl Chem **10:** 1345-1373

Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G (2010) A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. Proc Natl Acad Sci U S A **107:** 9452-9457

Bruyant-Vannier MP, Gaudinet-Schaumann A, Bourlard T, Morvan C (1996) Solubilization and partial characterization of pectin methyltransferase from flax cells. Plant Physiol Biochem **34:** 489-499

Buleon A, Colonna P, Planchot V, Ball S (1998) Starch granules: structure and biosynthesis. Int J Biol Macromol **23:** 85-112

Burda P, Aebi M (1999) The dolichol pathway of *N*-linked glycosylation. Biochim Biophys Acta **1426:** 239-257

Burda P, Jakob CA, Beinhauer J, Hegemann JH, Aebi M (1999) Ordered assembly of the asymmetrically branched lipid-linked oligosaccharide in the endoplasmic reticulum is ensured by the substrate specificity of the individual glycosyltransferases. Glycobiology **9:** 617-625

Burn JE, Hocart CH, Birch RJ, Cork AC, Williamson RE (2002) Functional analysis of the cellulose synthase genes *CesA1*, *CesA2*, and *CesA3* in Arabidopsis. Plant Physiol **129:** 797-807

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The *CesA* gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiol **134:** 224-236

Busse M, Feta A, Presto J, Wilen M, Gronning M, Kjellen L, Kusche-Gullberg M (2007) Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation. J Biol Chem **282:** 32802-32810

Busse M, Kusche-Gullberg M (2003) *In vitro* polymerization of heparan sulfate backbone by the EXT proteins. J Biol Chem **278:** 41333-41337

Caffall KH (2008) Expression and characterization of galacturonosyltransferase-6 (GAUT6) of the galacturonosyltransferase-1 (GAUT1)-related gene family of *Arabidopsis thaliana*. Ph.D Thesis. The University of Georgia, Athens

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res **344:** 1879-1900

Caffall KH, Pattathil S, Phillips SE, Hahn MG, Mohnen D (2009) *Arabidopsis thaliana* T-DNA mutants implicate *GAUT* genes in the biosynthesis of pectin and xylan in cell walls and seed testa. Mol Plant **2:** 1000-1014

Cano-Delgado A, Penfield S, Smith C, Catley M, Bevan M (2003) Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. Plant J **34:** 351-362

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res **37:** D233-238

Cao H, Imparl-Radosevich J, Guan H, Keeling PL, James MG, Myers AM (1999) Identification of the soluble starch synthase activities of maize endosperm. Plant Physiol **120:** 205-216

Cao H, James MG, Myers AM (2000) Purification and characterization of soluble starch synthases from maize endosperm. Arch Biochem Biophys **373:** 135-146

Carlsson P, Presto J, Spillmann D, Lindahl U, Kjellen L (2008) Heparin/heparan sulfate biosynthesis: processive formation of *N*-sulfated domains. J Biol Chem **283:** 20008-20014

Carpita N, Vergara C (1998) A recipe for cellulose. Science **279:** 672-673

Castro O, Movsichoff F, Parodi AJ (2006) Preferential transfer of the complete glycan is determined by the oligosaccharyltransferase complex and not by the catalytic subunit. Proc Natl Acad Sci U S A **103:** 14756-14760

Chatterjee M, Berbezy P, Vyas D, Coates SA, Barsby T (2005) Reduced expression of a protein homologous to glycogenin leads to reduction of starch content in *Arabidopsis* leaves. Plant Sci **168:** 501-509

Chavan M, Chen Z, Li G, Schindelin H, Lennarz WJ, Li H (2006) Dimeric organization of the yeast oligosaccharyl transferase complex. Proc Natl Acad Sci U S A **103:** 8947-8952

Chavan M, Yan A, Lennarz WJ (2005) Subunits of the translocon interact with components of the oligosaccharyl transferase complex. J Biol Chem **280:** 22917-22924

Chen S, Ehrhardt DW, Somerville CR (2010) Mutations of cellulose synthase (CESA1) phosphorylation sites modulate anisotropic cell expansion and bidirectional mobility of cellulose synthase. Proc Natl Acad Sci U S A **107:** 17188-17193

Choe S, Dilkes BP, Gregory BD, Ross AS, Yuan H, Noguchi T, Fujioka S, Takatsuto S, Tanaka A, Yoshida S, Tax FE, Feldmann KA (1999) The Arabidopsis *dwarf1* mutant is defective in the conversion of 24-methylenecholesterol to campesterol in brassinosteroid biosynthesis. Plant Physiol **119:** 897-907

Chu Z, Chen H, Zhang Y, Zhang Z, Zheng N, Yin B, Yan H, Zhu L, Zhao X, Yuan M, Zhang X, Xie Q (2007) Knockout of the *AtCESA2* gene affects microtubule orientation and causes abnormal cell expansion in Arabidopsis. Plant Physiol **143:** 213-224

Cifuentes C, Bulone V, Emons AM (2010) Biosynthesis of callose and cellulose by detergent extracts of tobacco cell membranes and quantification of the polymers synthesized *in vitro*. J Integr Plant Biol **52:** 221-233

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735-743

Coenen GJ, Bakx EJ, Verhoef RP, Schols HA, Voragen AG (2007) Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. Carbohydr. Polymers **70:** 224-235

Colombani A, Djerbi S, Bessueille L, Blomqvist K, Ohlsson A, Berglund T, Teeri TT, Bulone V (2004) *In vitro* synthesis of (1→3)-β-D-glucan (callose) and cellulose by detergent extracts of membranes from cell suspension cultures of hybrid aspen. Cellulose **11:** 313-327

Commuri PD, Keeling PL (2001) Chain-length specificities of maize starch synthase I enzyme: studies of glucan affinity and catalytic properties. Plant J **25:** 475-486

Cousins SK, Brown RMJ (1995) Cellulose I microfibril assembly: computational molecular mechanics energy analysis favours bonding by van der Waals forces as the initial step in crystallization. Polymer **36:** 3885-3888

Crawford BE, Olson SK, Esko JD, Pinhal MA (2001) Cloning, Golgi localization, and enzyme activity of the full-length heparin/heparan sulfate-glucuronic acid C5-epimerase. J Biol Chem **276:** 21538-21543

Crombie HJ, Scott C, Reid JSG (2003) Detergent-solubilisation of a homogalacturonan galacturonosyltransferase from mung bean. *In* AGT Voragen, H Schols, H Visser, eds, Advances in Pectin and Pectinase Research. Kluwer Academic Publishers, Dordrecht, pp 35-46

Crowell EF, Bischoff V, Desprez T, Rolland A, Stierhof YD, Schumacher K, Gonneau M, Hofte H, Vernhettes S (2009) Pausing of Golgi bodies on microtubules regulates secretion of cellulose synthase complexes in *Arabidopsis*. Plant Cell **21:** 1141-1154

Danon A, Rotari VI, Gordon A, Mailhac N, Gallois P (2004) Ultraviolet-C overexposure induces programmed cell death in *Arabidopsis*, which is mediated by caspase-like activities and which can be suppressed by caspase inhibitors, p35 and *Defender against Apoptotic Death*. J Biol Chem **279:** 779-787

Dardelle F, Lehner A, Ramdani Y, Bardor M, Lerouge P, Driouich A, Mollet JC (2010) Biochemical and immunocytological characterizations of *Arabidopsis thaliana* pollen tube cell wall. Plant Physiol **153:** 1563-1576

de Graffenried CL, Bertozzi CR (2004) The roles of enzyme localisation and complex formation in glycan assembly within the Golgi apparatus. Curr Opin Cell Biol **16:** 356-363

DeAngelis PL (1996) Enzymological characterization of the *Pasteurella multocida* hyaluronic acid synthase. Biochemistry **35:** 9768-9771

DeAngelis PL (1999) Hyaluronan synthases: fascinating glycosyltransferases from vertebrates, bacterial pathogens, and algal viruses. Cell Mol Life Sci **56:** 670-682

DeAngelis PL (2002) Microbial glycosaminoglycan glycosyltransferases. Glycobiology **12:** 9R-16R

DeAngelis PL, Jing W, Graves MV, Burbank DE, Van Etten JL (1997) Hyaluronan synthase of chlorella virus PBCV-1. Science **278:** 1800-1803

DeAngelis PL, Weigel PH (1994) Immunochemical confirmation of the primary structure of streptococcal hyaluronan synthase and synthesis of high molecular weight product by the recombinant enzyme. Biochemistry **33:** 9033-9039

Decreux A, Messiaen J (2005) Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. Plant Cell Physiol **46:** 268-278

Delatte T, Trevisan M, Parker ML, Zeeman SC (2005) Arabidopsis mutants *Atisa1* and *Atisa2* have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. Plant J **41:** 815-830

Delmer DP (1999) CELLULOSE BIOSYNTHESIS: Exciting Times for A Difficult Field of Study. Annu Rev Plant Physiol Plant Mol Biol **50:** 245-276

Derbyshire P, McCann MC, Roberts K (2007) Restricted cell elongation in *Arabidopsis* hypocotyls is associated with a reduced average pectin esterification level. BMC Plant Biol **7:** 31

Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylova Z, Parcy F, Hofte H, Gonneau M, Vernhettes S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A **104:** 15572-15577

Dhugga KS (2001) Building the wall: genes and enzyme complexes for polysaccharide synthases. Curr Opin Plant Biol **4:** 488-493

Diotallevi F, Mulder B (2007) The cellulose synthase complex: a polymerization driven supramolecular motor. Biophys J **92:** 2666-2673

Dirnberger D, Bencur P, Mach L, Steinkellner H (2002) The Golgi localization of *Arabidopsis thaliana* β1,2-xylosyltransferase in plant cells is dependent on its cytoplasmic and transmembrane sequences. Plant Mol Biol **50:** 273-281

Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol **43:** 1407-1420

Dolan L, Linstead P, Roberts K (1997) Developmental regulation of pectic polysaccharides in the root meristem of *Arabidopsis*. J Exp Bot **48:** 713-720

Doong RL, Liljebjelke K, Fralish G, Kumar A, Mohnen D (1995) Cell-free synthesis of pectin (Identification and partial characterization of polygalacturonate 4-α-galacturonosyltransferase and its products from membrane preparations of tobacco cell-suspension cultures). Plant Physiol **109:** 141-152

Doong RL, Mohnen D (1998) Solubilization and characterization of a galacturonosyltransferase that synthesizes the pectic polysaccharide homogalacturonan. The Plant Journal **13:** 363-374

DuBois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for determination of sugars and related substances. Anal Chem **28:** 350-356

Duncan MB, Liu M, Fox C, Liu J (2006) Characterization of the *N*-deacetylase domain from the heparan sulfate *N*-deacetylase/*N*-sulfotransferase 2. Biochem Biophys Res Commun **339:** 1232-1237

Dunkley TP, Hester S, Shadforth IP, Runions J, Weimar T, Hanton SL, Griffin JL, Bessant C, Brandizzi F, Hawes C, Watson RB, Dupree P, Lilley KS (2006) Mapping the Arabidopsis organelle proteome. Proc Natl Acad Sci U S A **103:** 6518-6523

Dunkley TP, Watson R, Griffin JL, Dupree P, Lilley KS (2004) Localization of organelle proteins by isotope tagging (LOPIT). Mol Cell Proteomics **3:** 1128-1134

Durand C, Vicre-Gibouin M, Follet-Gueye ML, Duponchel L, Moreau M, Lerouge P, Driouich A (2009) The organization pattern of root border-like cells of Arabidopsis is dependent on cell wall homogalacturonan. Plant Physiol **150:** 1411-1421

Eichler J, Adams MW (2005) Posttranslational protein modification in Archaea. Microbiol Mol Biol Rev **69:** 393-425

El-Battari A, Prorok M, Angata K, Mathieu S, Zerfaoui M, Ong E, Suzuki M, Lombardo D, Fukuda M (2003) Different glycosyltransferases are differentially processed for secretion, dimerization, and autoglycosylation. Glycobiology **13:** 941-953

Esko JD, Lindahl U (2001) Molecular diversity of heparan sulfate. J Clin Invest **108:** 169-173

Esko JD, Selleck SB (2002) Order out of chaos: assembly of ligand binding sites in heparan sulfate. Annu Rev Biochem **71:** 435-471

Esko JD, Zhang L (1996) Influence of core protein sequence on glycosaminoglycan assembly. Curr Opin Struct Biol **6:** 663-670

Ezaki N, Kido N, Takahashi K, Katou K (2005) The role of wall Ca^{2+} in the regulation of wall extensibility during the acid-induced extension of soybean hypocotyl cell walls. Plant Cell Physiol **46:** 1831-1838

Fagard M, Desnos T, Desprez T, Goubet F, Refregier G, Mouille G, McCann M, Rayon C, Vernhettes S, Hofte H (2000) *PROCUSTE1* encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of Arabidopsis. Plant Cell **12:** 2409-2424

Fetchko M, Stagljar I (2004) Application of the split-ubiquitin membrane yeast two-hybrid system to investigate membrane protein interactions. Methods **32:** 349-362

Fetrow JS, Siew N, Di Gennaro JA, Martinez-Yamout M, Dyson HJ, Skolnick J (2001) Genomic-scale comparison of sequence- and structure-based methods of function prediction: does structure provide additional insight? Protein Sci **10:** 1005-1014

Frydman RB, Cardini CE (1964) Soluble enzymes related to starch synthesis. Biochem Biophys Res Commun **17:** 407-411

Fujimoto M, Arimura S, Nakazono M, Tsutsumi N (2008) Arabidopsis dynamin-related protein DRP2B is co-localized with DRP1A on the leading edge of the forming cell plate. Plant Cell Rep **27:** 1581-1586

Fujita N, Taira T (1998) A 56-kDa protein is a novel granule-bound starch synthase existing in the pericarps, aleurone layers, and embryos of immature seed in diploid wheat (*Triticum monococcum* L.). Planta **207:** 125-132

Gao P, Xin Z, Zheng Z-L (2008) The *OSU1/QUA2/TSD2*-encoded putative methyltransferase is a critical modulator of carbon and nitrogen nutrient balance response in *Arabidopsis*. PLoS ONE **3:** e1387. doi:1310.1371/journal.pone.0001387

Gardiner JC, Taylor NG, Turner SR (2003) Control of cellulose synthase complex localization in developing xylem. Plant Cell **15:** 1740-1748

Gardner KH, Blackwell J (1974) The hydrogen bonding in native cellulose. Biochim Biophys Acta **343:** 232-237

Georgelis N, Shaw JR, Hannah LC (2009) Phylogenetic analysis of ADP-glucose pyrophosphorylase subunits reveals a role of subunit interfaces in the allosteric properties of the enzyme. Plant Physiol **151:** 67-77

Ghosh A, Kuppusamy H, Pilarski LM (2009) Aberrant splice variants of HAS1 (Hyaluronan Synthase 1) multimerize with and modulate normally spliced HAS1 protein: a potential mechanism promoting human cancer. J Biol Chem **284:** 18840-18850

Gibbons BJ, Roach PJ, Hurley TD (2002) Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. J Mol Biol **319:** 463-477

Goentzel BJ, Weigel PH, Steinberg RA (2006) Recombinant human hyaluronan synthase 3 is phosphorylated in mammalian cells. Biochem J **396:** 347-354

Gotting C, Kuhn J, Zahn R, Brinkmann T, Kleesiek K (2000) Molecular cloning and expression of human UDP-D-Xylose:proteoglycan core protein β-D-xylosyltransferase and its first isoform XT-II. J Mol Biol **304:** 517-528

Gotting C, Muller S, Schottler M, Schon S, Prante C, Brinkmann T, Kuhn J, Kleesiek K (2004) Analysis of the DXD motifs in human xylosyltransferase I required for enzyme activity. J Biol Chem **279:** 42566-42573

Goubet F, Council LN, Mohnen D (1998) Identification and partial characterization of the pectin methyltransferase "homogalacturonan-methyltransferase" from membranes of tobacco cell suspensions. Plant Physiol **116:** 337-347

Goubet F, Mohnen D (1999) Solubilization and partial characterization of homogalacturonan-methyltransferase from microsomal membranes of suspension-cultured tobacco cells. Plant Physiol **121:** 281-290

Goubet F, Mohnen D (1999) Subcellular localization and topology of homogalacturonan methyltransferase in suspension-cultured *Nicotiana tabacum* cells. Planta **209:** 112-117

Graves MV, Burbank DE, Roth R, Heuser J, DeAngelis PL, Van Etten JL (1999) Hyaluronan synthesis in virus PBCV-1-infected chlorella-like green algae. Virology **257:** 15-23

Grimaud F, Rogniaux H, James MG, Myers AM, Planchot V (2008) Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J Exp Bot **59:** 3395-3406

Grimson MJ, Haigler CH, Blanton RL (1996) Cellulose microfibrils, cell motility, and plasma membrane protein organization change in parallel during culmination in *Dictyostelium discoideum*. J Cell Sci **109 (Pt 13):** 3079-3087

Gu X, Bar-Peled M (2004) The biosynthesis of UDP-galacturonic acid in plants. Functional cloning and characterization of Arabidopsis UDP-D-glucuronic acid 4-epimerase. Plant Physiol **136:** 4256-4264

Gu Y, Kaplinsky N, Bringmann M, Cobb A, Carroll A, Sampathkumar A, Baskin TI, Persson S, Somerville CR (2010) Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. Proc Natl Acad Sci U S A **107:** 12866-12871

Guan H, Li P, Imparl-Radosevich J, Preiss J, Keeling P (1997) Comparing the properties of *Escherichia coli* branching enzyme and maize branching enzyme. Arch Biochem Biophys **342:** 92-98

Guerriero G, Fugelstad J, Bulone V (2010) What do we really know about cellulose biosynthesis in higher plants? J Integr Plant Biol **52:** 161-175

Guillaumie F, Sterling JD, Jensen KJ, Thomas OR, Mohnen D (2003) Solid-supported enzymatic synthesis of pectic oligogalacturonides and their analysis by MALDI-TOF mass spectrometry. Carbohydr Res **338:** 1951-1960

Gulberti S, Lattard V, Fondeur M, Jacquinet JC, Mulliert G, Netter P, Magdalou J, Ouzzine M, Fournel-Gigleux S (2005) Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human β1,4-galactosyltransferase 7 (GalT-I) and β1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J Biol Chem **280:** 1417-1425

Ha MA, Apperley DC, Evans BW, Huxham IM, Jardine WG, Vietor RJ, Reis D, Vian B, Jarvis MC (1998) Fine structure in cellulose microfibrils: NMR evidence from onion and quince. The Plant Journal **16:** 183-190

Habuchi H, Miyake G, Nogami K, Kuroiwa A, Matsuda Y, Kusche-Gullberg M, Habuchi O, Tanaka M, Kimata K (2003) Biosynthesis of heparan sulphate with diverse structures and functions: two alternatively spliced forms of human heparan sulphate 6-*O*-sulphotransferase-2 having different expression patterns and properties. Biochem J **371:** 131-142

Habuchi H, Tanaka M, Habuchi O, Yoshida K, Suzuki H, Ban K, Kimata K (2000) The occurrence of three isoforms of heparan sulfate 6-*O*-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted *N*-sulfoglucosamine. J Biol Chem **275:** 2859-2868

Hagner-McWhirter A, Li JP, Oscarson S, Lindahl U (2004) Irreversible glucuronyl C5-epimerization in the biosynthesis of heparan sulfate. J Biol Chem **279:** 14631-14638

Hahn MG, Darvill AG, Albersheim P (1981) Host-Pathogen Interactions : XIX. THE ENDOGENOUS ELICITOR, A FRAGMENT OF A PLANT CELL WALL POLYSACCHARIDE THAT ELICITS PHYTOALEXIN ACCUMULATION IN SOYBEANS. Plant Physiol **68:** 1161-1169

Haigler CH, Brown RM, Jr. (1986) Transport of rosettes from the Golgi apparatus to the plasma membrane in isolated mesophyll cells of *Zinnia elegans* during differentiation to tracheary elements in suspension culture. Protoplasma **134:** 111-120

Hanashiro I, Itoh K, Kuratomi Y, Yamazaki M, Igarashi T, Matsugasako J, Takeda Y (2008) Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. Plant Cell Physiol **49:** 925-933

Hannah LC, James M (2008) The complexities of starch biosynthesis in cereal endosperms. Curr Opin Biotechnol **19:** 160-165

Harada Y, Li H, Lennarz WJ (2009) Oligosaccharyltransferase directly binds to ribosome at a location near the translocon-binding site. Proc Natl Acad Sci U S A **106:** 6945-6949

Harholt J, Jensen JK, Sorensen SO, Orfila C, Pauly M, Scheller HV (2006) ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in Arabidopsis. Plant Physiol **140:** 49-58

Harris PJ, Northcote DH (1971) Polysaccharide formation in plant Golgi bodies. Biochim Biophys Acta **237:** 56-64

Harris PJ, Stone BA (2008) Chemistry and molecular organization of plant cell walls. *In* ME Himmel, ed, Biomass recalcitrance: deconstructing the plant cell wall for bioenergy. Blackwell Publishing Ltd, West Sussex, pp 61-93

Hascall VC (2000) Hyaluronan, a common thread. Glycoconj J **17:** 607-616

Hascall VC, Esko JD (2009) Hyaluronan. *In* A Varki, RD Cummings, JD Esko, HH Freeze, P Stanley, CR Bertozzi, GW Hart, ME Etzler, eds, Essentials of glycobiology, Ed 2nd. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Hascall VC, Majors AK, De La Motte CA, Evanko SP, Wang A, Drazba JA, Strong SA, Wight TN (2004) Intracellular hyaluronan: a new frontier for inflammation? Biochim Biophys Acta **1673:** 3-12

Hassinen A, Rivinoja A, Kauppila A, Kellokumpu S (2010) Golgi *N*-glycosyltransferases form both homo- and heterodimeric enzyme complexes in live cells. J Biol Chem **285:** 17771-17777

Heldermon C, DeAngelis PL, Weigel PH (2001) Topological organization of the hyaluronan synthase from *Streptococcus pyogenes*. J Biol Chem **276:** 2037-2046

Helenius A, Aebi M (2001) Intracellular functions of N-linked glycans. Science **291:** 2364-2369

Helenius A, Aebi M (2004) Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem **73:** 1019-1049

Hennen-Bierwagen TA, Lin Q, Grimaud F, Planchot V, Keeling PL, James MG, Myers AM (2009) Proteins from multiple metabolic pathways associate with starch biosynthetic enzymes in high molecular weight complexes: a model for regulation of carbon allocation in maize amyloplasts. Plant Physiol **149:** 1541-1559

Hennen-Bierwagen TA, Liu F, Marsh RS, Kim S, Gan Q, Tetlow IJ, Emes MJ, James MG, Myers AM (2008) Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes. Plant Physiol **146:** 1892-1908

Hese K, Otto C, Routier FH, Lehle L (2009) The yeast oligosaccharyltransferase complex can be replaced by STT3 from *Leishmania major*. Glycobiology **19:** 160-171

His I, Driouich A, Nicol F, Jauneau A, Hofte H (2001) Altered pectin composition in primary cell walls of *korrigan*, a dwarf mutant of Arabidopsis deficient in a membrane-bound endo-1,4-β-glucanase. Planta **212:** 348-358

Hizukuri S (1986) Polymodal distribution of the chain lengths of amylopectins, and its significance. Carbohydr Res **147:** 342-347

Hizukuri S, Takagi T (1984) Estimation of the distribution of molecular weight for amylose by the low-angle laser-light-scattering technique combined with high-performance gel chromatography. Carbohydr Res **134:** 1-10

Hu CD, Chinenov Y, Kerppola TK (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell **9:** 789-798

Hussain H, Mant A, Seale R, Zeeman S, Hinchliffe E, Edwards A, Hylton C, Bornemann S, Smith AM, Martin C, Bustos R (2003) Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133-149

Ibar C, Orellana A (2007) The import of *S*-adenosylmethionine into the Golgi apparatus is required for the methylation of homogalacturonan. Plant Physiol **145:** 504-512

Ishii T (1997) *O*-acetylated oligosaccharides from pectins of potato tuber cell walls. Plant Physiol **113:** 1265-1272

Ishii T (2002) A sensitive and rapid bioassay of homogalacturonan synthase using 2-aminobenzamide-labeled oligogalacturonides. Plant Cell Physiol **43:** 1386-1389

Ishii T, Ichita J, Matsue H, Ono H, Maeda I (2002) Fluorescent labeling of pectic oligosaccharides with 2-aminobenzamide and enzyme assay for pectin. Carbohydr Res **337:** 1023-1032

Ishii T, Matsunaga T (2001) Pectic polysaccharide rhamnogalacturonan II is covalently linked to homogalacturonan. Phytochemistry **57:** 969-974

Ishii T, Matsunaga T, Pellerin P, O'Neill MA, Darvill A, Albersheim P (1999) The plant cell wall polysaccharide rhamnogalacturonan II self-assembles into a covalently cross-linked dimer. J Biol Chem **274:** 13098-13104

Ishikawa M, Kuroyama H, Takeuchi Y, Tsumuraya Y (2000) Characterization of pectin methyltransferase from soybean hypocotyls. Planta **210:** 782-791

Itano N (2008) Simple primary structure, complex turnover regulation and multiple roles of hyaluronan. J Biochem **144:** 131-137

Itano N, Kimata K (2002) Mammalian hyaluronan synthases. IUBMB Life **54:** 195-199

Itano N, Sawai T, Yoshida M, Lenas P, Yamada Y, Imagawa M, Shinomura T, Hamaguchi M, Yoshida Y, Ohnuki Y, Miyauchi S, Spicer AP, McDonald JA, Kimata K (1999) Three isoforms of mammalian hyaluronan synthases have distinct enzymatic properties. J Biol Chem **274:** 25085-25092

Itoh K, Hirayama T, Takahashi A, Kubo W, Miyazaki S, Dairaku M, Togashi M, Mikami R, Attwood D (2007) In situ gelling pectin formulations for oral drug delivery at high gastric pH. Int J Pharm **335:** 90-96

Iwai H, Masaoka N, Ishii T, Satoh S (2002) A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem. Proc Natl Acad Sci U S A **99:** 16319-16324

Jackson CL, Dreaden TM, Theobald LK, Tran NM, Beal TL, Eid M, Gao MY, Shirley RB, Stoffel MT, Kumar MV, Mohnen D (2007) Pectin induces apoptosis in human prostate cancer cells: correlation of apoptotic function with pectin structure. Glycobiology **17:** 805-819

Jacob-Wilk D, Kurek I, Hogan P, Delmer DP (2006) The cotton fiber zinc-binding domain of cellulose synthase A1 from *Gossypium hirsutum* displays rapid turnover *in vitro* and *in vivo*. Proc Natl Acad Sci U S A **103:** 12191-12196

James MG, Denyer K, Myers AM (2003) Starch synthesis in the cereal endosperm. Curr Opin Plant Biol **6:** 215-222

Jarvis MC (1984) Structure and properties of pectin gels in plant cell walls. Plant Cell Environ. **7:** 153-164

Jarvis MC, Briggs SPH, Knox JP (2003) Intercellular adhesion and cell separation in plants. Plant, Cell and Environment **26:** 977-989

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J **6:** 3901-3907

Jenkins PJ, Donald AM (1995) The influence of amylose on starch granule structure. Int J Biol Macromol **17:** 315-321

Jeon JS, Ryoo N, Hahn TR, Walia H, Nakamura Y (2010) Starch biosynthesis in cereal endosperm. Plant Physiol Biochem **48:** 383-392

Jin JB, Kim YA, Kim SJ, Lee SH, Kim DH, Cheong GW, Hwang I (2001) A new dynamin-like protein, ADL6, is involved in trafficking from the *trans*-Golgi network to the central vacuole in Arabidopsis. Plant Cell **13:** 1511-1526

Jobling S (2004) Improving starch for food and industrial applications. Curr Opin Plant Biol **7:** 210-218

Johnson MA, von Besser K, Zhou Q, Smith E, Aux G, Patton D, Levin JZ, Preuss D (2004) Arabidopsis *hapless* mutations define essential gametophytic functions. Genetics **168:** 971-982

Ju T, Aryal RP, Stowell CJ, Cummings RD (2008) Regulation of protein *O*-glycosylation by the endoplasmic reticulum-localized molecular chaperone Cosmc. J Cell Biol **182:** 531-542

Ju T, Brewer K, D'Souza A, Cummings RD, Canfield WM (2002) Cloning and expression of human core 1 β1,3-galactosyltransferase. J Biol Chem **277:** 178-186

Ju T, Cummings RD, Canfield WM (2002) Purification, characterization, and subunit structure of rat core 1 β1,3-galactosyltransferase. J Biol Chem **277:** 169-177

Kakuta Y, Sueyoshi T, Negishi M, Pedersen LC (1999) Crystal structure of the sulfotransferase domain of human heparan sulfate *N*-deacetylase/*N*-sulfotransferase 1. J Biol Chem **274:** 10673-10676

Kampf M, Absmanner B, Schwarz M, Lehle L (2009) Biochemical characterization and membrane topology of Alg2 from *Saccharomyces cerevisiae* as a bifunctional α1,3- and 1,6-mannosyltransferase involved in lipid-linked oligosaccharide biosynthesis. J Biol Chem **284:** 11900-11912

Karamyshev AL, Kelleher DJ, Gilmore R, Johnson AE, von Heijne G, Nilsson I (2005) Mapping the interaction of the STT3 subunit of the oligosaccharyl transferase complex with nascent polypeptide chains. J Biol Chem **280:** 40489-40493

Karaoglu D, Kelleher DJ, Gilmore R (1997) The highly conserved Stt3 protein is a subunit of the yeast oligosaccharyltransferase and forms a subcomplex with Ost3p and Ost4p. J Biol Chem **272:** 32513-32520

Kataoka Y, Kondo T (1998) FT-IR microscopic analysis of changing cellulose crystalline structure during wood cell wall formation. Macromolecules **31:** 760-764

Kauss H, Hassid WZ (1967) Enzymic introduction of the methyl ester groups of pectin. J Biol Chem **242:** 3449-3453

Kauss H, Swanson AL (1969) Cooperation of enzymes responsible for polymerisation and methylation in pectin biosynthesis. Z Naturforsch B **24:** 28-33

Kauss H, Swanson AL, Hassid WZ (1967) Biosynthesis of the methyl ester groups of pectin by transmethylation from *S*-adenosyl-L-methionine. Biochem Biophys Res Commun **26:** 234-240

Kelleher DJ, Gilmore R (2006) An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology **16:** 47R-62R

Kelleher DJ, Karaoglu D, Mandon EC, Gilmore R (2003) Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. Mol Cell **12:** 101-111

Kim BT, Kitagawa H, Tamura J, Saito T, Kusche-Gullberg M, Lindahl U, Sugahara K (2001) Human tumor suppressor *EXT* gene family members *EXTL1* and *EXTL3* encode α1,4-*N*-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis. Proc Natl Acad Sci U S A **98:** 7176-7181

Kim H, Yan Q, Von Heijne G, Caputo GA, Lennarz WJ (2003) Determination of the membrane topology of Ost4p and its subunit interactions in the oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A **100:** 7460-7464

Kimura S, Itoh T (1996) New cellulose synthesizing complexes (terminal complexes) involved in animal cellulose biosynthesis in the tunicate *Metandrocarpa uedai*. Protoplasma **194:** 151-163

Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM, Jr. (1999) Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. Plant Cell **11:** 2075-2086

Kitagawa H, Izumikawa T, Uyama T, Sugahara K (2003) Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization. J Biol Chem **278:** 23666-23671

Kitagawa H, Shimakawa H, Sugahara K (1999) The tumor suppressor EXT-like gene *EXTL2* encodes an α1,4-*N*-acetylhexosaminyltransferase that transfers *N*-acetylgalactosamine

and *N*-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of heparan sulfate. J Biol Chem **274:** 13933-13937

Kitagawa H, Tone Y, Tamura J, Neumann KW, Ogawa T, Oka S, Kawasaki T, Sugahara K (1998) Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J Biol Chem **273:** 6615-6618

Klahre U, Noguchi T, Fujioka S, Takatsuto S, Yokota T, Nomura T, Yoshida S, Chua NH (1998) The Arabidopsis *DIMINUTO/DWARF1* gene encodes a protein involved in steroid synthesis. Plant Cell **10:** 1677-1690

Kleczkowski LA, Sokolov LN, Luo C, Villand P (1999) Molecular cloning and spatial expression of an ApL1 cDNA for the large subunit of ADP-glucose pyrophosphorylase from *Arabidopsis thaliana*. Z Naturforsch C **54:** 353-358

Knauer R, Lehle L (1999) The oligosaccharyltransferase complex from yeast. Biochim Biophys Acta **1426:** 259-273

Knox JP (1992) Cell-adhesion, cell-separation and plant morphogenesis. Plant Journal **2:** 137-141

Kobayashi S, Morimoto K, Shimizu T, Takahashi M, Kurosawa H, Shirasawa T (2000) Association of EXT1 and EXT2, hereditary multiple exostoses gene products, in Golgi apparatus. Biochem Biophys Res Commun **268:** 860-867

Kogan G, Soltes L, Stern R, Gemeiner P (2007) Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. Biotechnol Lett **29:** 17-25

Kohorn BD, Johansen S, Shishido A, Todorova T, Martinez R, Defeo E, Obregon P (2009) Pectin activation of MAP kinase and gene expression is WAK2 dependent. Plant J **60:** 974-982

Koike T, Izumikawa T, Tamura J, Kitagawa H (2009) FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan-protein linkage region. Biochem J **421:** 157-162

Kolpak FJ, Blackwell J (1976) Determination of the structure of cellulose II. Macromolecules **9:** 273-278

Kong Y, Zhou G, Avci U, Gu X, Jones C, Yin Y, Xu Y, Hahn MG (2009) Two poplar glycosyltransferase genes, *PdGATL1.1* and *PdGATL1.2*, are functional orthologs to *PARVUS/AtGATL1* in *Arabidopsis*. Mol Plant **2:** 1040-1050

Kotting O, Kossmann J, Zeeman SC, Lloyd JR (2010) Regulation of starch metabolism: the age of enlightenment? Curr Opin Plant Biol **13:** 321-329

Koyama M, Helbert W, Imai T, Sugiyama J, Henrissat B (1997) Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. Proc Natl Acad Sci U S A **94:** 9091-9095

Krupkova E, Immerzeel P, Pauly M, Schmulling T (2007) The *TUMOROUS SHOOT DEVELOPMENT2* gene of Arabidopsis encoding a putative methyltransferase is required for cell adhesion and co-ordinated plant development. Plant J **50:** 735-750

Kubo A, Colleoni C, Dinges JR, Lin Q, Lappe RR, Rivenbark JG, Meyer AJ, Ball SG, James MG, Hennen-Bierwagen TA, Myers AM (2010) Functions of heteromeric and homomeric isoamylase-type starch-debranching enzymes in developing maize endosperm. Plant Physiol **153:** 956-969

Kudlicka K, Brown RM, Jr., Li L, Lee JH, Shin H, Kuga S (1995) β-Glucan synthesis in the cotton fiber (IV. *In vitro* assembly of the cellulose I allomorph). Plant Physiol **107:** 111-123

Kudlicka K, Brown RMJ (1997) Cellulose and callose biosynthesis in higher plants. I. Solubilization and separation of (1→3)- and (1→4)-β-glucan synthase activities from mung bean. Plant Physiol **115:** 643-656

Kudo T, Iwai T, Kubota T, Iwasaki H, Takayma Y, Hiruma T, Inaba N, Zhang Y, Gotoh M, Togayachi A, Narimatsu H (2002) Molecular cloning and characterization of a novel UDP-Gal:GalNAcα peptide β1,3-galactosyltransferase (C1Gal-T2), an enzyme synthesizing a core 1 structure of *O*-glycan. J Biol Chem **277:** 47724-47731

Kuhn J, Gotting C, Schnolzer M, Kempf T, Brinkmann T, Kleesiek K (2001) First isolation of human UDP-D-xylose: proteoglycan core protein β-D-xylosyltransferase secreted from cultured JAR choriocarcinoma cells. J Biol Chem **276:** 4940-4947

Kumar GP, Sudheesh S, Vijayalakshmi NR (1993) Hypoglycaemic effect of *Coccinia indica*: mechanism of action. Planta Med **59:** 330-332

Kumar M, Thammannagowda S, Bulone V, Chiang V, Han KH, Joshi CP, Mansfield SD, Mellerowicz E, Sundberg B, Teeri T, Ellis BE (2009) An update on the nomenclature for the cellulose synthase genes in *Populus*. Trends Plant Sci **14:** 248-254

Kurek I, Kawagoe Y, Jacob-Wilk D, Doblin M, Delmer D (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. Proc Natl Acad Sci U S A **99:** 11109-11114

Kusche-Gullberg M, Kjellen L (2003) Sulfotransferases in glycosaminoglycan biosynthesis. Curr Opin Struct Biol **13:** 605-611

Lai-Kee-Him J, Chanzy H, Muller M, Putaux JL, Imai T, Bulone V (2002) *In vitro versus in vivo* cellulose microfibrils from plant primary wall synthases: structural differences. J Biol Chem **277:** 36931-36939

Lam BC, Sage TL, Bianchi F, Blumwald E (2002) Regulation of ADL6 activity by its associated molecular network. Plant J **31:** 565-576

Lane DR, Wiedemeier A, Peng L, Hofte H, Vernhettes S, Desprez T, Hocart CH, Birch RJ, Baskin TI, Burn JE, Arioli T, Betzner AS, Williamson RE (2001) Temperature-sensitive alleles of *RSW2* link the KORRIGAN endo-1,4-β-glucanase to cellulose synthesis and cytokinesis in Arabidopsis. Plant Physiol **126:** 278-288

Lapcik L, Jr., Lapcik L, De Smedt S, Demeester J, Chabrecek P (1998) Hyaluronan: preparation, structure, properties, and applications. Chem Rev **98:** 2663-2684

Laurent TC, Fraser JR (1992) Hyaluronan. FASEB J **6:** 2397-2404

Leboeuf E, Guillon F, Thoiron S, Lahaye M (2005) Biochemical and immunohistochemical analysis of pectic polysaccharides in the cell walls of *Arabidopsis* mutant *QUASIMODO 1* suspension-cultured cells: implications for cell adhesion. J Exp Bot **56:** 3171-3182

Lee C, Teng Q, Huang W, Zhong R, Ye ZH (2009) The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. Plant Cell Physiol **50:** 1982-1987

Lee C, Zhong R, Richardson EA, Himmelsbach DS, McPhail BT, Ye ZH (2007) The *PARVUS* gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. Plant Cell Physiol **48:** 1659-1672

Lee SH, Jin JB, Song J, Min MK, Park DS, Kim YW, Hwang I (2002) The intermolecular interaction between the PH domain and the C-terminal domain of Arabidopsis dynamin-like 6 determines lipid binding specificity. J Biol Chem **277:** 31842-31849

Leloir LF, De Fekete MA, Cardini CE (1961) Starch and oligosaccharide synthesis from uridine diphosphate glucose. J Biol Chem **236:** 636-641

Lennarz WJ (2007) Studies on oligosaccharyl transferase in yeast. Acta Biochim Pol **54:** 673-677

Lennon KA, Lord EM (2000) *In vivo* pollen tube cell of *Arabidopsis thaliana*. I. Tube cell cytoplasm and wall. Protoplasma **214:** 45-56

Lertpiriyapong K, Sung ZR (2003) The *elongation defective1* mutant of *Arabidopsis* is impaired in the gene encoding a serine-rich secreted protein. Plant Mol Biol **53:** 581-595

Li H, Chavan M, Schindelin H, Lennarz WJ (2008) Structure of the oligosaccharyl transferase complex at 12 Å resolution. Structure **16:** 432-440

Li JP, Gong F, El Darwish K, Jalkanen M, Lindahl U (2001) Characterization of the D-glucuronyl C5-epimerase involved in the biosynthesis of heparin and heparan sulfate. J Biol Chem **276:** 20069-20077

Liljebjelke K, Adolphson R, Baker K, Doong RL, Mohnen D (1995) Enzymatic synthesis and purification of uridine diphosphate [14C]galacturonic acid: a substrate for pectin biosynthesis. Anal Biochem **225:** 296-304

Lin TY, Elbein AD, Su JC (1966) Substrate specificity in pectin synthesis. Biochem Biophys Res Commun **22:** 650-657

Lind T, Tufaro F, McCormick C, Lindahl U, Lidholt K (1998) The putative tumor suppressors EXT1 and EXT2 are glycosyltransferases required for the biosynthesis of heparan sulfate. J Biol Chem **273:** 26265-26268

Lindeboom J, Mulder BM, Vos JW, Ketelaar T, Emons AM (2008) Cellulose microfibril deposition: coordinated activity at the plant plasma membrane. J Microsc **231:** 192-200

Liners F, Gaspar T, Van Cutsem P (1994) Acetyl- and methyl-esterification of pectins of friable and compact sugar-beet calli: consequences for intercellular adhesion. Planta **192:** 545-556

Linker A, Hoffman P, Sampson P, Meyer K (1958) Heparitin sulfate. Biochim Biophys Acta **29:** 443-444

Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci U S A **107:** 616-621

Liu F, Makhmoudova A, Lee EA, Wait R, Emes MJ, Tetlow IJ (2009) The *amylose extender* mutant of maize conditions novel protein-protein interactions between starch biosynthetic enzymes in amyloplasts. J Exp Bot **60:** 4423-4440

Liu L, Chen G, Fishman ML, Hicks KB (2005) Pectin gel vehicles for controlled fragrance delivery. Drug Deliv **12:** 149-157

Lomako J, Lomako WM, Whelan WJ (2004) Glycogenin: the primer for mammalian and yeast glycogen synthesis. Biochim Biophys Acta **1673:** 45-55

Lynch MA, Staehelin LA (1992) Domain-specific and cell type-specific localization of two types of cell wall matrix polysaccharides in the clover root tip. J Cell Biol **118:** 467-479

Macquet A, Ralet MC, Kronenberger J, Marion-Poll A, North HM (2007) In situ, chemical and macromolecular study of the composition of *Arabidopsis thaliana* seed coat mucilage. Plant Cell Physiol **48:** 984-999

Manderson K, Pinart M, Tuohy KM, Grace WE, Hotchkiss AT, Widmer W, Yadhav MP, Gibson GR, Rastall RA (2005) *In vitro* determination of prebiotic properties of oligosaccharides derived from an orange juice manufacturing by-product stream. Appl. Environ. Microbiol. **71:** 8383-8389

Manderson K, Pinart M, Tuohy KM, Grace WE, Hotchkiss AT, Widmer W, Yadhav MP, Gibson GR, Rastall RA (2005) *In vitro* determination of prebiotic properties of oligosaccharides derived from an orange juice manufacturing by-product stream. Appl Environ Microbiol **71:** 8383-8389

Manya H, Chiba A, Margolis RU, Endo T (2006) Molecular cloning and characterization of rat Pomt1 and Pomt2. Glycobiology **16:** 863-873

Manya H, Chiba A, Yoshida A, Wang X, Chiba Y, Jigami Y, Margolis RU, Endo T (2004) Demonstration of mammalian protein *O*-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. Proc Natl Acad Sci U S A **101:** 500-505

Master ER, Rudsander UJ, Zhou W, Henriksson H, Divne C, Denman S, Wilson DB, Teeri TT (2004) Recombinant expression and enzymatic characterization of PttCel9A, a KOR homologue from *Populus tremula x tremuloides*. Biochemistry **43:** 10080-10089

Mathieu Y, Guern J, Spiro MD, O'Neill MA, Kates K, Darvill A, Albersheim P (1998) The transient nature of the oligogalacturonide-induced ion fluxes of tobacco cells is not correlated with fragmentation of the oligogalacturonides. Plant J **16:** 305-311

Matthysse AG, White S, Lightfoot R (1995) Genes required for cellulose synthesis in Agrobacterium tumefaciens. J Bacteriol **177:** 1069-1075

McCann MC, Roberts K (1991) Architecture of the primary cell wall. *In* CW Lloyd, ed, The cytoskeletal basis of plant growth and form. Academic Press San Diego, pp 109-128

McCormick C, Duncan G, Goutsos KT, Tufaro F (2000) The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. Proc Natl Acad Sci U S A **97:** 668-673

McCormick C, Leduc Y, Martindale D, Mattison K, Esford LE, Dyer AP, Tufaro F (1998) The putative tumour suppressor EXT1 alters the expression of cell-surface heparan sulfate. Nat Genet **19:** 158-161

Meikle PJ, Ng KF, Johnson E, Hoogenraad NJ, Stone BA (1991) The β-glucan synthase from *Lolium multiflorum*. Detergent solubilization, purification using monoclonal antibodies, and photoaffinity labeling with a novel photoreactive pyrimidine analogue of uridine 5'-diphosphoglucose. J Biol Chem **266:** 22569-22581

Melotto E, Greve LC, Labavitch JM (1994) Cell wall metabolism in ripening fruit (VII. Biologically active pectin oligomers in ripening tomato (*Lycopersicon esculentum* Mill.) fruits). Plant Physiol **106:** 575-581

Messiaen J, Read ND, Van Cutsem P, Trewavas AJ (1993) Cell wall oligogalacturonides increase cytosolic free calcium in carrot protoplasts. J Cell Sci **104:** 365-371

Messiaen J, Van Cutsem P (1993) Defense gene transcription in carrot cells treated with oligogalacturonides. Plant Cell Physiol **34:** 1117-1123

Messner P (2004) Prokaryotic glycoproteins: unexplored but important. J Bacteriol **186:** 2517-2519

Meyer K, Palmer J (1934) The polysaccharide of the vitreous humor. J Biol Chem **107:** 629-634

Meyer MF, Kreil G (1996) Cells expressing the DG42 gene from early *Xenopus* embryos synthesize hyaluronan. Proc Natl Acad Sci U S A **93:** 4543-4547

Mochizuki H, Yoshida K, Shibata Y, Kimata K (2008) Tetrasulfated disaccharide unit in heparan sulfate: enzymatic formation and tissue distribution. J Biol Chem **283:** 31237-31245

Mohnen D (1999) Biosynthesis of pectins and galactomannans. *In* D Barton, K Nakanishi, O Meth-Cohn, eds, Comprehensive Natural Products Chemistry, Vol 3. Elsevier Science Ltd., Amsterdam, pp 497-527

Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol **11:** 266-277

Mohnen D, Bar-Peled M, Somerville CR (2008) Biosynthesis of plant cell walls. *In* M Himmel, ed, Biomass Recalcitrance. Blackwell Publishing, Oxford, pp 94-187

Molhoj M, Pagant S, Hofte H (2002) Towards understanding the role of membrane-bound endo-β-1,4-glucanases in cellulose biosynthesis. Plant Cell Physiol **43:** 1399-1406

Molhoj M, Ulvskov P, Dal Degan F (2001) Characterization of a functional soluble form of a *Brassica napus* membrane-anchored endo-1,4-β-glucanase heterologously expressed in *Pichia pastoris*. Plant Physiol **127:** 674-684

Mollet JC, Park SY, Nothnagel EA, Lord EM (2000) A lily stylar pectin is necessary for pollen tube adhesion to an in vitro stylar matrix. Plant Cell **12:** 1737-1750

Moore PJ, Swords KM, Lynch MA, Staehelin LA (1991) Spatial organization of the assembly pathways of glycoproteins and complex polysaccharides in the Golgi apparatus of plants. J Cell Biol **112:** 589-602

Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR, Dupree P (2010) Absence of branches from xylan in Arabidopsis *gux* mutants reveals potential for simplification of lignocellulosic biomass. Proc Natl Acad Sci U S A **107:** 17409-17414

Mouille G, Ralet MC, Cavelier C, Eland C, Effroy D, Hematy K, McCartney L, Truong HN, Gaudon V, Thibault JF, Marchant A, Hofte H (2007) Homogalacturonan synthesis in *Arabidopsis thaliana* requires a Golgi-localized protein with a putative methyltransferase domain. Plant J **50:** 605-614

Mueller SC, Brown RM, Jr. (1980) Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. J Cell Biol **84:** 315-326

Mukerjea R, Robyt JF (2005) Starch biosynthesis: further evidence against the primer nonreducing-end mechanism and evidence for the reducing-end two-site insertion mechanism. Carbohydr Res **340:** 2206-2211

Mukerjea R, Robyt JF (2005) Starch biosynthesis: the primer nonreducing-end mechanism versus the nonprimer reducing-end two-site insertion mechanism. Carbohydr Res **340:** 245-255

Mukerjea R, Yu L, Robyt JF (2002) Starch biosynthesis: mechanism for the elongation of starch chains. Carbohydr Res **337:** 1015-1022

Muller S, Disse J, Schottler M, Schon S, Prante C, Brinkmann T, Kuhn J, Kleesiek K, Gotting C (2006) Human xylosyltransferase I and N-terminal truncated forms: functional characterization of the core enzyme. Biochem J **394:** 163-171

Mutter M, Renard CM, Beldman G, Schols HA, Voragen AG (1998) Mode of action of RG-hydrolase and RG-lyase toward rhamnogalacturonan oligomers. Characterization of degradation products using RG-rhamnohydrolase and RG-galacturonohydrolase. Carbohydr Res **311:** 155-164

Mutwil M, Debolt S, Persson S (2008) Cellulose synthesis: a complex complex. Curr Opin Plant Biol **11:** 252-257

Nadanaka S, Kitagawa H (2008) Heparan sulphate biosynthesis and disease. J Biochem **144:** 7-14

Nagai N, Habuchi H, Esko JD, Kimata K (2004) Stem domains of heparan sulfate 6-*O*-sulfotransferase are required for Golgi localization, oligomer formation and enzyme activity. J Cell Sci **117:** 3331-3341

Nakamura A, Furuta H, Maeda H, Takao T, Nagamatsu Y (2002) Analysis of the molecular construction of xylogalacturonan isolated from soluble soybean polysaccharides. Biosci Biotechnol Biochem **66:** 1155-1158

Nakamura A, Furuta H, Maeda H, Takao T, Nagamatsu Y (2002) Structural studies by stepwise enzymatic degradation of the main backbone of soybean soluble polysaccharides consisting of galacturonan and rhamnogalacturonan. Biosci Biotechnol Biochem **66:** 1301-1313

Nakamura Y, Utsumi Y, Sawada T, Aihara S, Utsumi C, Yoshida M, Kitamura S (2010) Characterization of the reactions of starch branching enzymes from rice endosperm. Plant Cell Physiol **51:** 776-794

Nakata PA, Greene TW, Anderson JM, Smith-White BJ, Okita TW, Preiss J (1991) Comparison of the primary sequences of two potato tuber ADP-glucose pyrophosphorylase subunits. Plant Mol Biol **17:** 1089-1093

Nangia-Makker P, Hogan V, Honjo Y, Baccarini S, Tait L, Bresalier R, Raz A (2002) Inhibition of human cancer cell growth and metastasis in nude mice by oral intake of modified citrus pectin. J Natl Cancer Inst **94:** 1854-1862

Nasab FP, Schulz BL, Gamarro F, Parodi AJ, Aebi M (2008) All in one: *Leishmania major* STT3 proteins substitute for the whole oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. Mol Biol Cell **19:** 3758-3768

Navazio L, Moscatiello R, Bellincampi D, Baldan B, Meggio F, Brini M, Bowler C, Mariani P (2002) The role of calcium in oligogalacturonide-activated signalling in soybean cells. Planta **215:** 596-605

Negishi M, Dong J, Darden TA, Pedersen LG, Pedersen LC (2003) Glucosaminylglycan biosynthesis: what we can learn from the X-ray crystal structures of glycosyltransferases GlcAT1 and EXTL2. Biochem Biophys Res Commun **303:** 393-398

Nelson OE, Rines HW (1962) The enzymatic deficiency in the *waxy* mutant of maize. Biochem Biophys Res Commun **9:** 297-300

Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem **75:** 4646-4658

Nicol F, His I, Jauneau A, Vernhettes S, Canut H, Hofte H (1998) A plasma membrane-bound putative endo-1,4-β-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. EMBO J **17:** 5563-5576

Nilsson I, Kelleher DJ, Miao Y, Shao Y, Kreibich G, Gilmore R, von Heijne G, Johnson AE (2003) Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. J Cell Biol **161:** 715-725

Nilsson T, Hoe MH, Slusarewicz P, Rabouille C, Watson R, Hunte F, Watzele G, Berger EG, Warren G (1994) Kin recognition between medial Golgi enzymes in HeLa cells. Embo J **13:** 562-574

Nilsson T, Rabouille C, Hui N, Watson R, Warren G (1996) The role of the membrane-spanning domain and stalk region of *N*-acetylglucosaminyltransferase I in retention, kin recognition and structural maintenance of the Golgi apparatus in HeLa cells. J Cell Sci **109 (Pt 7):** 1975-1989

Nishi A, Nakamura Y, Tanaka N, Satoh H (2001) Biochemical and genetic analysis of the effects of *Amylose-Extender* mutation in rice endosperm. Plant Physiol **127:** 459-472

Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc **124:** 9074-9082

Nishiyama Y, Sugiyama J, Chanzy H, Langan P (2003) Crystal structure and hydrogen bonding system in cellulose Iα from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc **125:** 14300-14306

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2004) Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. Plant Cell **16:** 2394-2405

O'Neill MA, Eberhard S, Albersheim P, Darvill AG (2001) Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. Science **294:** 846-849

O'Neill MA, Ishii T, Albersheim P, Darvill AG (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annu Rev Plant Biol **55:** 109-139

O'Neill MA, Warrenfeltz D, Kates K, Pellerin P, Doco T, Darvill AG, Albersheim P (1996) Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. *In vitro* conditions for the formation and hydrolysis of the dimer. J Biol Chem **271:** 22923-22930

O'Reilly MK, Zhang G, Imperiali B (2006) In vitro evidence for the dual function of Alg2 and Alg11: essential mannosyltransferases in N-linked glycoprotein biosynthesis. Biochemistry **45:** 9593-9603

O'Sullivan AC (1997) Cellulose: the structure slowly unravels. Cellulose **4:** 173-207

Oecking C, Jaspert N (2009) Plant 14-3-3 proteins catch up with their mammalian orthologs. Curr Opin Plant Biol **12:** 760-765

Ogawa H, Shionyu M, Sugiura N, Hatano S, Nagai N, Kubota Y, Nishiwaki K, Sato T, Gotoh M, Narimatsu H, Shimizu K, Kimata K, Watanabe H (2010) Chondroitin sulfate synthase-2/chondroitin polymerizing factor has two variants with distinct function. J Biol Chem **285:** 34155-34167

Ogston AG, Stanier JE (1951) The dimensions of the particle of hyaluronic acid complex in synovial fluid. Biochem J **49:** 585-590

Ogston AG, Stanier JE (1952) Further observations on the preparation and composition of the hyaluronic acid complex of ox synovial fluid. Biochem J **52:** 149-156

Ohashi T, Ishimizu T, Akita K, Hase S (2007) In vitro stabilization and minimum active component of polygalacturonic acid synthase involved in pectin biosynthesis. Biosci Biotechnol Biochem **71:** 2291-2299

Ohdan T, Francisco PB, Jr., Sawada T, Hirose T, Terao T, Satoh H, Nakamura Y (2005) Expression profiling of genes involved in starch synthesis in sink and source organs of rice. J Exp Bot **56:** 3229-3244

Okada M, Nadanaka S, Shoji N, Tamura J, Kitagawa H (2010) Biosynthesis of heparan sulfate in EXT1-deficient cells. Biochem J **428:** 463-471

Okajima T, Yoshida K, Kondo T, Furukawa K (1999) Human homolog of *Caenorhabditis elegans sqv-3* gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J Biol Chem **274:** 22915-22918

Okita TW, Nakata PA, Anderson JM, Sowokinos J, Morell M, Preiss J (1990) The subunit structure of potato tuber ADP-glucose pyrophosphorylase. Plant Physiol **93:** 785-790

Okuda K, Li L, Kudlicka K, Kuga S, Brown RM, Jr. (1993) β-Glucan synthesis in the cotton fiber (I. Identification of β-1,4- and β-1,3-glucans synthesized *in vitro*). Plant Physiol **101:** 1131-1142

Opat AS, Houghton F, Gleeson PA (2000) Medial Golgi but not late Golgi glycosyltransferases exist as high molecular weight complexes. Role of luminal domain in complex formation and localization. J Biol Chem **275:** 11836-11845

Orfila C, Sorensen SO, Harholt J, Geshi N, Crombie H, Truong HN, Reid JS, Knox JP, Scheller HV (2005) *QUASIMODO1* is expressed in vascular tissue of *Arabidopsis thaliana* inflorescence stems, and affects homogalacturonan and xylan biosynthesis. Planta **222:** 613-622

Orzechowski S (2008) Starch metabolism in leaves. Acta Biochim Pol **55:** 435-445

Ouskova G, Spellerberg B, Prehm P (2004) Hyaluronan release from *Streptococcus pyogenes*: export by an ABC transporter. Glycobiology **14:** 931-938

Pagant S, Bichet A, Sugimoto K, Lerouxel O, Desprez T, McCann M, Lerouge P, Vernhettes S, Hofte H (2002) *KOBITO1* encodes a novel plasma membrane protein necessary for normal synthesis of cellulose during cell expansion in Arabidopsis. Plant Cell **14:** 2001-2013

Paredez AR, Somerville CR, Ehrhardt DW (2006) Visualization of cellulose synthase demonstrates functional association with microtubules. Science **312:** 1491-1495

Pattison RJ, Amtmann A (2009) N-glycan production in the endoplasmic reticulum of plants. Trends Plant Sci **14:** 92-99

Paulsen BS, Barsett H (2005) Bioactive pectic polysaccharides. Adv. Polym. Sci. **186:** 69-101

Pauly M, Scheller HV (2000) *O*-Acetylation of plant cell wall polysaccharides: identification and partial characterization of a rhamnogalacturonan *O*-acetyl-transferase from potato suspension-cultured cells. Planta **210:** 659-667

Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci U S A **93:** 12637-12642

Pedersen LC, Dong J, Taniguchi F, Kitagawa H, Krahn JM, Pedersen LG, Sugahara K, Negishi M (2003) Crystal structure of an α1,4-*N*-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. J Biol Chem **278:** 14420-14428

Pedrioli PG, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, Pratt B, Nilsson E, Angeletti RH, Apweiler R, Cheung K, Costello CE, Hermjakob H, Huang S, Julian RK, Kapp E, McComb ME, Oliver SG, Omenn G, Paton NW, Simpson R, Smith R, Taylor CF, Zhu W, Aebersold R (2004) A common open representation of mass spectrometry data and its application to proteomics research. Nat Biotechnol **22:** 1459-1466

Pena MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH (2007) *Arabidopsis irregular xylem8* and *irregular xylem9*: implications for the complexity of glucuronoxylan biosynthesis. Plant Cell **19:** 549-563

Peng L, Kawagoe Y, Hogan P, Delmer D (2002) Sitosterol-β-glucoside as primer for cellulose synthesis in plants. Science **295:** 147-150

Perrone P, Hewage CM, Thomson AR, Bailey K, Sadler IH, Fry SC (2002) Patterns of methyl and *O*-acetyl esterification in spinach pectins: new complexity. Phytochemistry **60:** 67-77

Persson S, Caffall KH, Freshour G, Hilley MT, Bauer S, Poindexter P, Hahn MG, Mohnen D, Somerville C (2007) The *Arabidopsis irregular xylem8* mutant is deficient in glucuronoxylan and homogalacturonan, which are essential for secondary cell wall integrity. Plant Cell **19:** 237-255

Persson S, Paredez A, Carroll A, Palsdottir H, Doblin M, Poindexter P, Khitrov N, Auer M, Somerville CR (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. Proc Natl Acad Sci U S A **104:** 15566-15571

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci U S A **102:** 8633-8638

Pienta KJ, Naik H, Akhtar A, Yamazaki K, Replogle TS, Lehr J, Donat TL, Tait L, Hogan V, Raz A (1995) Inhibition of spontaneous metastasis in a rat prostate cancer model by oral administration of modified citrus pectin. J Natl Cancer Inst **87:** 348-353

Pinhal MA, Smith B, Olson S, Aikawa J, Kimata K, Esko JD (2001) Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-*O*-sulfotransferase interact *in vivo*. Proc Natl Acad Sci U S A **98:** 12984-12989

Pippen EL, McCready RM, Owens HS (1950) Gelation properties of partially acetylated pectins. J. Am. Chem. Soc. **72:** 813-816

Popper ZA, Fry SC (2005) Widespread occurrence of a covalent linkage between xyloglucan and acidic polysaccharides in suspension-cultured angiosperm cells. Ann Bot (Lond) **96:** 91-99

Popper ZA, Fry SC (2008) Xyloglucan-pectin linkages are formed intra-protoplasmically, contribute to wall-assembly, and remain stable in the cell wall. Planta **227:** 781-794

Prehm P, Schumacher U (2004) Inhibition of hyaluronan export from human fibroblasts by inhibitors of multidrug resistance transporters. Biochem Pharmacol **68:** 1401-1410

Presto J, Thuveson M, Carlsson P, Busse M, Wilen M, Eriksson I, Kusche-Gullberg M, Kjellen L (2008) Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. Proc Natl Acad Sci U S A **105:** 4751-4756

Pummill PE, Achyuthan AM, DeAngelis PL (1998) Enzymological characterization of recombinant *Xenopus* DG42, a vertebrate hyaluronan synthase. J Biol Chem **273:** 4976-4981

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ, Jr., Hallett JP, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, Tschaplinski T (2006) The path forward for biofuels and biomaterials. Science **311:** 484-489

Ralet MC, Crepeau MJ, Buchholt HC, Thibault JF (2003) Polyelectrolyte behaviour and calcium binding properties of sugar beet pectins differing in their degrees of methylation and acetylation. Biochemical Engineering Journal **16:** 191-201

Ralet MC, Crepeau MJ, Lefebvre J, Mouille G, Hofte H, Thibault JF (2008) Reduced number of homogalacturonan domains in pectins of an *Arabidopsis* mutant enhances the flexibility of the polymer. Biomacromolecules **9:** 1454-1460

Ramesh HP, Tharanathan RN (2003) Carbohydrates--the renewable raw materials of high biotechnological value. Crit Rev Biotechnol **23:** 149-173

Regina A, Kosar-Hashemi B, Ling S, Li Z, Rahman S, Morell M (2010) Control of starch branching in barley defined through differential RNAi suppression of starch branching enzyme IIa and IIb. J Exp Bot **61:** 1469-1482

Renard CM, Lahaye M, Mutter M, Voragen FG, Thibault JF (1997) Isolation and structural characterisation of rhamnogalacturonan oligomers generated by controlled acid hydrolysis of sugar-beet pulp. Carbohydr Res **305:** 271-280

Renard CMGC, Jarvis MC (1999) Acetylation and methylation of homogalacturonans 1: optimisation of the reaction and characterisation of the products. Carbohydr Polymers **39:** 201-207

Richmond T (2000) Higher plant cellulose synthases. Genome Biol **1:** REVIEWS3001

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry **57:** 929-967

Rilla K, Siiskonen H, Spicer AP, Hyttinen JM, Tammi MI, Tammi RH (2005) Plasma membrane residence of hyaluronan synthase is coupled to its enzymatic activity. J Biol Chem **280:** 31890-31897

Rini J, Esko JD, Varki A (2008) Glycosyltransferases and glycan-processing enzymes. *In* A Varki, RD Cummings, JD Esko, HH Freeze, P Stanley, CR Bertozzi, GW Hart, ME Etzler, eds, Essentials of glycobiology, Ed 2nd. Cold Spring Harbor Laboratory Press, Plainview, NY.

Robert S, Bichet A, Grandjean O, Kierzkowski D, Satiat-Jeunemaitre B, Pelletier S, Hauser MT, Hofte H, Vernhettes S (2005) An *Arabidopsis* endo-1,4-β-D-glucanase involved in cellulose synthesis undergoes regulated intracellular cycling. Plant Cell **17:** 3378-3389

Roldan I, Wattebled F, Mercedes Lucas M, Delvalle D, Planchot V, Jimenez S, Perez R, Ball S, D'Hulst C, Merida A (2007) The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. Plant J **49:** 492-504

Rong J, Habuchi H, Kimata K, Lindahl U, Kusche-Gullberg M (2001) Substrate specificity of the heparan sulfate hexuronic acid 2-*O*-sulfotransferase. Biochemistry **40:** 5548-5555

Rongine De Fekete MA, Leloir LF, Cardini CE (1960) Mechanism of starch biosynthesis. Nature **187:** 918-919

Roudier F, Fernandez AG, Fujita M, Himmelspach R, Borner GH, Schindelman G, Song S, Baskin TI, Dupree P, Wasteneys GO, Benfey PN (2005) COBRA, an Arabidopsis extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. Plant Cell **17:** 1749-1763

Ruiz-Canada C, Kelleher DJ, Gilmore R (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. Cell **136:** 272-283

Sakuragi Y, Nørholm MHH, Scheller HV (In print) Visual Mapping of Cell Wall Biosynthesis. *In* ZA Popper, ed, The Plant Cell Wall: Methods and Protocols. Springer, New York

Sato S, Kato T, Kakegawa K, Ishii T, Liu YG, Awano T, Takabe K, Nishiyama Y, Kuga S, Nakamura Y, Tabata S, Shibata D (2001) Role of the putative membrane-bound endo-1,4-β-glucanase KORRIGAN in cell elongation and cellulose synthesis in *Arabidopsis thaliana*. Plant Cell Physiol **42:** 251-263

Satoh H, Shibahara K, Tokunaga T, Nishi A, Tasaki M, Hwang SK, Okita TW, Kaneko N, Fujita N, Yoshida M, Hosaka Y, Sato A, Utsumi Y, Ohdan T, Nakamura Y (2008) Mutation of the plastidial α-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. Plant Cell **20:** 1833-1849

Saurin AJ, Borden KL, Boddy MN, Freemont PS (1996) Does this have a familiar RING? Trends Biochem Sci **21:** 208-214

Sawada T, Francisco PB, Jr., Aihara S, Utsumi Y, Yoshida M, Oyama Y, Tsuzuki M, Satoh H, Nakamura Y (2009) Chlorella starch branching enzyme II (BEII) can complement the function of BEIIb in rice endosperm. Plant Cell Physiol **50:** 1062-1074

Saxena IM, Brown RM, Jr. (1997) Identification of cellulose synthase(s) in higher plants: sequence analysis of processive β-glycosyltransferases with the common motif 'D,D,D35Q(R,Q)XRW'. Cellulose **4:** 33-49

Saxena IM, Brown RM, Jr. (2005) Cellulose biosynthesis: current views and evolving concepts. Ann Bot **96:** 9-21

Saxena IM, Brown RM, Jr., Fevre M, Geremia RA, Henrissat B (1995) Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. J Bacteriol **177:** 1419-1424

Schaumann A, Bruyant-Vannier MP, Goubet F, Morvan C (1993) Pectic metabolism in suspension-cultured cells of flax, *Linum usitatissimum*. Plant Cell Physiol. **34:** 891-897

Scheible WR, Eshed R, Richmond T, Delmer D, Somerville C (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis Ixr1* mutants. Proc Natl Acad Sci U S A **98:** 10079-10084

Scheller HV, Doong RL, Ridley BL, Mohnen D (1999) Pectin biosynthesis: a solubilized α1,4-galacturonosyltransferase from tobacco catalyzes the transfer of galacturonic acid from UDP-galacturonic acid onto the non-reducing end of homogalacturonan. Planta **207:** 512-517

Schindelman G, Morikami A, Jung J, Baskin TI, Carpita NC, Derbyshire P, McCann MC, Benfey PN (2001) *COBRA* encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis*. Genes Dev **15:** 1115-1127

Schols HA, Geraeds CCJM, Searle-van Leeuwen MF, Kormelink FJM, Voragen AG (1990) Rhamnogalacturonase: a novel enzyme that degrades the hairy regions of pectins. Carbohydr Res **206:** 105-115

Schon S, Prante C, Bahr C, Kuhn J, Kleesiek K, Gotting C (2006) Cloning and recombinant expression of active full-length xylosyltransferase I (XT-I) and characterization of subcellular localization of XT-I and XT-II. J Biol Chem **281:** 14224-14231

Schulz BL, Aebi M (2009) Analysis of glycosylation site occupancy reveals a role for Ost3p and Ost6p in site-specific *N*-glycosylation efficiency. Mol Cell Proteomics **8:** 357-364

Schulz BL, Stirnimann CU, Grimshaw JP, Brozzo MS, Fritsch F, Mohorko E, Capitani G, Glockshuber R, Grutter MG, Aebi M (2009) Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. Proc Natl Acad Sci U S A **106:** 11061-11066

Schulz T, Schumacher U, Prehm P (2007) Hyaluronan export by the ABC transporter MRP5 and its modulation by intracellular cGMP. J Biol Chem **282:** 20999-21004

Schwarz M, Knauer R, Lehle L (2005) Yeast oligosaccharyltransferase consists of two functionally distinct sub-complexes, specified by either the Ost3p or Ost6p subunit. FEBS Lett **579:** 6564-6568

Sehnke PC, Chung HJ, Wu K, Ferl RJ (2001) Regulation of starch accumulation by granule-associated plant 14-3-3 proteins. Proc Natl Acad Sci U S A **98:** 765-770

Senay C, Lind T, Muguruma K, Tone Y, Kitagawa H, Sugahara K, Lidholt K, Lindahl U, Kusche-Gullberg M (2000) The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis. EMBO Rep **1:** 282-286

Shedletzky E, Shmuel M, Trainin T, Kalman S, Delmer D (1992) Cell wall structure in cells adapted to growth on the cellulose-synthesis inhibitor 2,6-dichlorobenzonitrile : a comparison between two dicotyledonous plants and a Graminaceous monocot. Plant Physiol **100:** 120-130

Sherrier DJ, VandenBosch KA (1994) Secretion of cell wall polysaccharides in *Vicia* root hairs. The Plant Journal **5:** 185-195

Shibatani T, David LL, McCormack AL, Frueh K, Skach WR (2005) Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits. Biochemistry **44:** 5982-5992

Shure M, Wessler S, Fedoroff N (1983) Molecular identification and isolation of the *Waxy* locus in maize. Cell **35:** 225-233

Shworak NW, Liu J, Petros LM, Zhang L, Kobayashi M, Copeland NG, Jenkins NA, Rosenberg RD (1999) Multiple isoforms of heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase. Isolation, characterization, and expression of human cDNAs and identification of distinct genomic loci. J Biol Chem **274:** 5170-5184

Singh B, Avci U, Eichler Inwood SE, Grimson MJ, Landgraf J, Mohnen D, Sorensen I, Wilkerson CG, Willats WG, Haigler CH (2009) A specialized outer layer of the

primary cell wall joins elongating cotton fibers into tissue-like bundles. Plant Physiol **150:** 684-699

Skidmore MA, Guimond SE, Rudd TR, Fernig DG, Turnbull JE, Yates EA (2008) The activities of heparan sulfate and its analogue heparin are dictated by biosynthesis, sequence, and conformation. Connect Tissue Res **49:** 140-144

Smeds E, Feta A, Kusche-Gullberg M (2010) Target selection of heparan sulfate hexuronic acid 2-*O*-sulfotransferase. Glycobiology **20:** 1274-1282

Smith AM (2001) The biosynthesis of starch granules. Biomacromolecules **2:** 335-341

Somerville C (2006) Cellulose synthesis in higher plants. Annu Rev Cell Dev Biol **22:** 53-78

Spicer AP, Augustine ML, McDonald JA (1996) Molecular cloning and characterization of a putative mouse hyaluronan synthase. J Biol Chem **271:** 23400-23406

Spicer AP, McDonald JA (1998) Characterization and molecular evolution of a vertebrate hyaluronan synthase gene family. J Biol Chem **273:** 1923-1932

Spicer AP, Olson JS, McDonald JA (1997) Molecular cloning and characterization of a cDNA encoding the third putative mammalian hyaluronan synthase. J Biol Chem **272:** 8957-8961

Spirig U, Bodmer D, Wacker M, Burda P, Aebi M (2005) The 3.4-kDa Ost4 protein is required for the assembly of two distinct oligosaccharyltransferase complexes in yeast. Glycobiology **15:** 1396-1406

Spiro MD, Kates KA, Koller AL, O'Neill MA, Albersheim P, Darvill AG (1993) Purification and characterization of biologically active 1,4-linked α-oligogalacturonides after partial digestion of polygalacturonic acid with endopolygalacturonase. Carbohydr Res **247:** 9-20

Spiro MD, Ridley BL, Eberhard S, Kates KA, Mathieu Y, O'Neill MA, Mohnen D, Guern J, Darvill A, Albersheim P (1998) Biological activity of reducing-end-derivatized oligogalacturonides in tobacco tissue cultures. Plant Physiol **116:** 1289-1298

Staehelin LA, Moore I (1995) The plant Golgi apparatus: structure, functional organization and trafficking mechanisms. Annu Rev Plant Physiol Plant Mol Biol **46:** 261-288

Stanley P, Schachter H, Taniguchi N (2008) N-glycans. *In* A Varki, JP Cummings, JD Esko, HH Freeze, P Stanley, CR Bertozzi, GW Hart, ME Etzler, eds, Essentials of glycobiology., Ed 2nd. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Sterling JD (2004) Characterization and identification of a galacturonosyltransferase involved in pectin biosynthesis. Ph.D Thesis. The University of Georgia, Athens

Sterling JD, Atmodjo MA, Inwood SE, Kumar Kolli VS, Quigley HF, Hahn MG, Mohnen D (2006) Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. Proc Natl Acad Sci U S A **103:** 5236-5241

Sterling JD, Lemons JA, Forkner IF, Mohnen D (2005) Development of a filter assay for measuring homogalacturonan: alpha-(1,4)-Galacturonosyltransferase activity. Anal Biochem **343:** 231-236

Sterling JD, Lemons JA, Forkner IF, Mohnen D (2005) Development of a filter assay for measuring homogalacturonan: α-(1,4)-galacturonosyltransferase activity. Anal Biochem **343:** 231-236

Sterling JD, Quigley HF, Orellana A, Mohnen D (2001) The catalytic site of the pectin biosynthetic enzyme α-1,4-galacturonosyltransferase is located in the lumen of the Golgi. Plant Physiol **127:** 360-371

Streb S, Delatte T, Umhang M, Eicke S, Schorderet M, Reinhardt D, Zeeman SC (2008) Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. Plant Cell **20:** 3448-3466

Sudheesh S, Vijayalakshmi NR (1999) Lipid-lowering action of pectin from *Cucumis sativus*. Food Chemistry **67:** 281-286

Sugahara K, Kitagawa H (2002) Heparin and heparan sulfate biosynthesis. IUBMB Life **54:** 163-175

Sugiyama J, Vuong R, Chanzy H (1991) Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall. Macromolecules **24:** 4168-4175

Szydlowski N, Ragel P, Raynaud S, Lucas MM, Roldan I, Montero M, Munoz FJ, Ovecka M, Bahaji A, Planchot V, Pozueta-Romero J, D'Hulst C, Merida A (2009) Starch granule initiation in *Arabidopsis* requires the presence of either class IV or class III starch synthases. Plant Cell **21:** 2443-2457

Szyjanowicz PM, McKinnon I, Taylor NG, Gardiner J, Jarvis MC, Turner SR (2004) The *irregular xylem 2* mutant is an allele of *korrigan* that affects the secondary cell wall of *Arabidopsis thaliana*. Plant J **37:** 730-740

Takeo S, Fujise M, Akiyama T, Habuchi H, Itano N, Matsuo T, Aigaki T, Kimata K, Nakato H (2004) *In vivo* hyaluronan synthesis upon expression of the mammalian hyaluronan synthase gene in *Drosophila*. J Biol Chem **279:** 18920-18925

Takeuchi Y, Tsumuraya Y (2001) *In vitro* biosynthesis of homogalacturonan by a membrane-bound galacturonosyltransferase from epicotyls of azuki bean. Biosci Biotechnol Biochem **65:** 1519-1527

Talhaoui I, Bui C, Oriol R, Mulliert G, Gulberti S, Netter P, Coughtrie MW, Ouzzine M, Fournel-Gigleux S (2010) Identification of key functional residues in the active site of human β1,4-galactosyltransferase 7: a major enzyme in the glycosaminoglycan synthesis pathway. J Biol Chem

Tammi MI, Day AJ, Turley EA (2002) Hyaluronan and homeostasis: a balancing act. J Biol Chem **277:** 4581-4584

Tanaka K, Murata K, Yamazaki M, Onosato K, Miyao A, Hirochika H (2003) Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. Plant Physiol **133:** 73-83

Taylor NG (2007) Identification of cellulose synthase AtCesA7 (IRX3) in vivo phosphorylation sites--a potential role in regulating protein degradation. Plant Mol Biol **64:** 161-171

Taylor NG (2008) Cellulose biosynthesis and deposition in higher plants. New Phytol **178:** 239-252

Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci U S A **100:** 1450-1455

Taylor NG, Laurie S, Turner SR (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis. Plant Cell **12:** 2529-2540

Taylor NG, Scheible WR, Cutler S, Somerville CR, Turner SR (1999) The *irregular xylem3* locus of Arabidopsis encodes a cellulose synthase required for secondary cell wall synthesis. Plant Cell **11:** 769-780

Tetlow IJ, Beisel KG, Cameron S, Makhmoudova A, Liu F, Bresolin NS, Wait R, Morell MK, Emes MJ (2008) Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. Plant Physiol **146:** 1878-1891

Tetlow IJ, Morell MK, Emes MJ (2004) Recent developments in understanding the regulation of starch metabolism in higher plants. J Exp Bot **55:** 2131-2145

Tetlow IJ, Wait R, Lu Z, Akkasaeng R, Bowsher CG, Esposito S, Kosar-Hashemi B, Morell MK, Emes MJ (2004) Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein-protein interactions. Plant Cell **16:** 694-708

Thakur BR, Singh RK, Handa AK (1997) Chemistry and uses of pectin--a review. Crit Rev Food Sci Nutr **37:** 47-73

Theuwissen E, Mensink RP (2008) Water-soluble dietary fibers and cardiovascular disease. Physiol Behav **94:** 285-292

Tickle P, Burrell MM, Coates SA, Emes MJ, Tetlow IJ, Bowsher CG (2009) Characterization of plastidial starch phosphorylase in *Triticum aestivum* L. endosperm. J Plant Physiol **166:** 1465-1478

Timmers J, Vernhettes S, Desprez T, Vincken JP, Visser RG, Trindade LM (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett **583:** 978-982

Toole BP (2004) Hyaluronan: from extracellular glue to pericellular cue. Nat Rev Cancer **4:** 528-539

Turley EA, Noble PW, Bourguignon LY (2002) Signaling properties of hyaluronan receptors. J Biol Chem **277:** 4589-4592

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell **9:** 689-701

Utsumi Y, Nakamura Y (2006) Structural and enzymatic characterization of the isoamylase1 homo-oligomer and the isoamylase1-isoamylase2 hetero-oligomer from rice endosperm. Planta **225:** 75-87

Vainonen JP, Sakuragi Y, Stael S, Tikkanen M, Allahverdiyeva Y, Paakkarinen V, Aro E, Suorsa M, Scheller HV, Vener AV, Aro EM (2008) Light regulation of CaS, a novel phosphoprotein in the thylakoid membrane of *Arabidopsis thaliana*. FEBS J **275:** 1767-1777

Vannier MP, Thoiron B, Morvan C, Demarty M (1992) Localization of methyltransferase activities throughout the endomembrane system of flax (*Linum usitatissimum* L) hypocotyls. Biochem J **286 (Pt 3):** 863-868

Ventriglia T, Kuhn ML, Ruiz MT, Ribeiro-Pedro M, Valverde F, Ballicora MA, Preiss J, Romero JM (2008) Two Arabidopsis ADP-glucose pyrophosphorylase large subunits (APL1 and APL2) are catalytic. Plant Physiol **148:** 65-76

Vergara CE, Carpita NC (2001) β-D-glycan synthases and the *CesA* gene family: lessons to be learned from the mixed-linkage (1→3),(1→4)β-D-glucan synthase. Plant Mol Biol **47:** 145-160

Vigetti D, Genasetti A, Karousou E, Viola M, Clerici M, Bartolini B, Moretto P, De Luca G, Hascall VC, Passi A (2009) Modulation of hyaluronan synthase activity in cellular membrane fractions. J Biol Chem **284:** 30684-30694

Villand P, Olsen OA, Kleczkowski LA (1993) Molecular characterization of multiple cDNA clones for ADP-glucose pyrophosphorylase from *Arabidopsis thaliana*. Plant Mol Biol **23:** 1279-1284

Villemez CL, Lin TY, Hassid WZ (1965) Biosynthesis of the polygalacturonic acid chain of pectin by a particulate enzyme preparation from *Phaseolus aureus* seedlings. Proc Natl Acad Sci U S A **54:** 1626-1632

Villemez CL, Swanson AL, Hassid WZ (1966) Properties of a polygalacturonic acid-synthesizing enzyme system from *Phaseolus aureus* seedlings. Arch Biochem Biophys **116:** 446-452

Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J **33:** 949-956

Volpi N, Schiller J, Stern R, Soltes L (2009) Role, metabolism, chemical modifications and applications of hyaluronan. Curr Med Chem **16:** 1718-1745

Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol **122:** 255-264

Vu NT, Shimada H, Kakuta Y, Nakashima T, Ida H, Omori T, Nishi A, Satoh H, Kimura M (2008) Biochemical and crystallographic characterization of the starch branching enzyme I (BEI) from *Oryza sativa* L. Biosci Biotechnol Biochem **72:** 2858-2866

Wagner GK, Pesnot T (2010) Glycosyltransferases and their assays. Chembiochem **11:** 1939-1949

Wallsgrove RM, Lea PJ, Miflin BJ (1983) Intracellular localization of aspartate kinase and the enzymes of threonine and methionine biosynthesis in green leaves. Plant Physiol **71:** 780-784

Wang J, Elliott JE, Williamson RE (2008) Features of the primary wall CESA complex in wild type and cellulose-deficient mutants of *Arabidopsis thaliana*. J Exp Bot **59:** 2627-2637

Wang J, Howles PA, Cork AH, Birch RJ, Williamson RE (2006) Chimeric proteins suggest that the catalytic and/or C-terminal domains give CesA1 and CesA3 access to their specific sites in the cellulose synthase of primary walls. Plant Physiol **142:** 685-695

Weatherly DB, Atwood JA, 3rd, Minning TA, Cavola C, Tarleton RL, Orlando R (2005) A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. Mol Cell Proteomics **4:** 762-772

Weerapana E, Imperiali B (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology **16:** 91R-101R

Wei G, Bai X, Gabb MM, Bame KJ, Koshy TI, Spear PG, Esko JD (2000) Location of the glucuronosyltransferase domain in the heparan sulfate copolymerase EXT1 by analysis of Chinese hamster ovary cell mutants. J Biol Chem **275:** 27733-27740

Wei G, Bai X, Sarkar AK, Esko JD (1999) Formation of HNK-1 determinants and the glycosaminoglycan tetrasaccharide linkage region by UDP-GlcUA:Galactose β1,3-glucuronosyltransferases. J Biol Chem **274:** 7857-7864

Weigel PH, Hascall VC, Tammi M (1997) Hyaluronan synthases. J Biol Chem **272:** 13997-14000

Western TL, Young DS, Dean GH, Tan WL, Samuels AL, Haughn GW (2004) *MUCILAGE-MODIFIED4* encodes a putative pectin biosynthetic enzyme developmentally regulated by *APETALA2, TRANSPARENT TESTA GLABRA1*, and *GLABRA2* in the Arabidopsis seed coat. Plant Physiol **134:** 296-306

Whitcombe AJ, O'Neill MA, Steffan W, Albersheim P, Darvill AG (1995) Structural characterization of the pectic polysaccharide, rhamnogalacturonan-II. Carbohydr Res **271:** 15-29

Whitelock JM, Iozzo RV (2005) Heparan sulfate: a complex polymer charged with biological activity. Chem Rev **105:** 2745-2764

Wightman R, Marshall R, Turner SR (2009) A cellulose synthase-containing compartment moves rapidly beneath sites of secondary wall synthesis. Plant Cell Physiol **50:** 584-594

Wightman R, Turner S (2010) Trafficking of the cellulose synthase complex in developing xylem vessels. Biochem Soc Trans **38:** 755-760

Wightman R, Turner SR (2008) The roles of the cytoskeleton during cellulose deposition at the secondary cell wall. Plant J **54:** 794-805

Willats WG, McCartney L, Mackie W, Knox JP (2001) Pectin: cell biology and prospects for functional analysis. Plant Mol Biol **47:** 9-27

Willats WG, Orfila C, Limberg G, Buchholt HC, van Alebeek GJ, Voragen AG, Marcus SE, Christensen TM, Mikkelsen JD, Murray BS, Knox JP (2001) Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls. Implications for pectin methyl esterase action, matrix properties, and cell adhesion. J Biol Chem **276:** 19404-19413

Willer T, Amselgruber W, Deutzmann R, Strahl S (2002) Characterization of POMT2, a novel member of the PMT protein *O*-mannosyltransferase family specifically localized to the acrosome of mammalian spermatids. Glycobiology **12:** 771-783

Willison JH, Brown RM, Jr. (1978) Cell wall structure and deposition in *Glaucocystis*. J Cell Biol **77:** 103-119

Wilson CM, High S (2007) Ribophorin I acts as a substrate-specific facilitator of N-glycosylation. J Cell Sci **120:** 648-657

Wilson CM, Kraft C, Duggan C, Ismail N, Crawshaw SG, High S (2005) Ribophorin I associates with a subset of membrane proteins after their integration at the sec61 translocon. J Biol Chem **280:** 4195-4206

Wilson CM, Roebuck Q, High S (2008) Ribophorin I regulates substrate delivery to the oligosaccharyltransferase core. Proc Natl Acad Sci U S A **105:** 9534-9539

Wolf S, Mouille G, Pelloux J (2009) Homogalacturonan methyl-esterification and plant development. Mol Plant **2:** 851-860

Wolf S, Rausch T, Greiner S (2009) The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. Plant J **58:** 361-375

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MM, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan S, Shinozaki K, Davis RW, Theologis A, Ecker JR (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science **302:** 842-846

Yan A, Lennarz WJ (2005) Two oligosaccharyl transferase complexes exist in yeast and associate with two different translocons. Glycobiology **15:** 1407-1415

Yan A, Wu E, Lennarz WJ (2005) Studies of yeast oligosaccharyl transferase subunits using the split-ubiquitin system: topological features and *in vivo* interactions. Proc Natl Acad Sci U S A **102:** 7121-7126

Yang B, Yang BL, Savani RC, Turley EA (1994) Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein. EMBO J **13:** 286-296

Yapo BM, Lerouge P, Thibault JF, Ralet MC (2007) Pectins from citrus peel cell walls contain homogalacturonans homogenous with respect to molar mass, rhamnogalacturonan I and rhamnogalacturonan II. Carbohydr Polymers **69:** 426-435

Yin Y, Chen H, Hahn MG, Mohnen D, Xu Y (2010) Evolution and function of the plant cell wall synthesis-related Glycosyltransferase Family 8. Plant Physiol **153:** 1729-1746

Yin Y, Mohnen D, Gelineo-Albersheim I, Xu Y, Hahn MG (In press) Glycosyltransferases of family 8 (GT8). *In* Plant Cell Wall Polysaccharides: Biosynthesis and Bioengineering Blackwell, Oxford, UK

Yoneda A, Ito T, Higaki T, Kutsuna N, Saito T, Ishimizu T, Osada H, Hasezawa S, Matsui M, Demura T (2010) Cobtorin target analysis reveals that pectin functions in the deposition of cellulose microfibrils in parallel with cortical microtubules. The Plant Journal **64:** 657-667

York WS, O'Neill MA (2008) Biochemical control of xylan biosynthesis - which end is up? Curr Opin Plant Biol **11:** 258-265

Yoshida M, Itano N, Yamada Y, Kimata K (2000) *In vitro synthesis* of hyaluronan by a single protein derived from mouse HAS1 gene and characterization of amino acid residues essential for the activity. J Biol Chem **275:** 497-506

Young WW, Jr. (2004) Organization of Golgi glycosyltransferases in membranes: complexity via complexes. J Membr Biol **198:** 1-13

Zandleven J, Sorensen SO, Harholt J, Beldman G, Schols HA, Scheller HV, Voragen AJ (2007) Xylogalacturonan exists in cell walls from various tissues of *Arabidopsis thaliana*. Phytochemistry **68:** 1219-1226

Zeeman SC, Kossmann J, Smith AM (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol **61:** 209-234

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol **154:** 78-97

Zhang GF, Staehelin LA (1992) Functional compartmentation of the Golgi apparatus of plant cells. Plant Physiol **99:** 1070-1083

Zhang X, Szydlowski N, Delvalle D, D'Hulst C, James MG, Myers AM (2008) Overlapping functions of the starch synthases SSII and SSIII in amylopectin biosynthesis in *Arabidopsis*. BMC Plant Biol **8:** 96

Zheng H, Kunst L, Hawes C, Moore I (2004) A GFP-based assay reveals a role for RHD3 in transport between the endoplasmic reticulum and Golgi apparatus. Plant J **37:** 398-414

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol **136:** 2621-2632

Zuo J, Niu QW, Nishizawa N, Wu Y, Kost B, Chua NH (2000) KORRIGAN, an Arabidopsis endo-1,4-β-glucanase, localizes to the cell plate by polarized targeting and is essential for cytokinesis. Plant Cell **12:** 1137-1152

APPENDIX A

CHARACTERIZATION OF ANTI-GAUT1, ANTI-GAUT7, AND ANTI-GAUT4 POLYCLONAL ANTIBODIES

The identification of GAUT1 and GAUT7 from the proteomics analyses of partially purified, GalAT-enriched Arabidopsis solubilized enzyme preparations, made clear the importance of generating polyclonal antibodies against these proteins for use as molecular tools for future research (Sterling, 2004). Considering the high level of amino acid sequence similarity among the 15 members of the GAUT1-related gene family, the approach taken was to raise the antibodies against unique peptide regions instead of against whole proteins. Three peptide sequences were selected from each protein based on the following criteria: low hydrophobicity, high beta-turn character, and predictions that the selected sequence resided in the loop regions of the proteins. The selected peptides (19-25 a.a. long) were synthesized as tetrameric multiple antigenic peptides (MAPs) on a tri-lysine core, and each was used to immunize two New Zealand White rabbits. In addition to GAUT1 and GAUT7, anti-peptide antibodies were also raised against JS36L/GAUT4 (At5g47780), another member of the GAUT gene family with significant homology to GAUT1. For each of the three proteins, GAUT1, GAUT7, and GAUT4, six sets of antibodies against 3 peptide regions were produced (Table A-1, Figure A-1) that were characterized as hereby described.

Table A-1. Antigenic peptide sequences of GAUT1, GAUT7, and GAUT4 used to raise the polyclonal antibodies.

Protein / Peptides	Amino acid position	Peptide sequence	Peptide length (a.a.)	Antibody #
GAUT1/JS36/At3g61130				
PIM006	341-365	IDYYLLSPEKRKFPRSENLENPNLY	25	738-739
PIM007	448-472	AMREYYFKADHPTSGSSNLKYRNPK	25	740-741
PIM008	132-154	NLNEKRDSISKDSIHQKVETPTK	23	742-743
GAUT7/JS33/At2g38650				
PIM003	515-540	LGVSETYQKYYKEMSSGDESSEAIL	25	732-733
PIM004	302-326	LTVEHFKSDSLEDPISEKFSDPSLL	25	734-735
PIM005	82-103	DEVLQKINPVLPKKSDINVGSR	22	736-737
GAUT4/JS36L/At5g47780				
PIM009	128-150	LTQQTSEKVDEQPEPNAFGAKKD	23	744-745
PIM010	93-116	ATDDDTHSHTDISIKQVTHDAASD	24	746-747
PIM011	288-306	ALNSSEQQFPNQEKLEDTQ	19	748-749

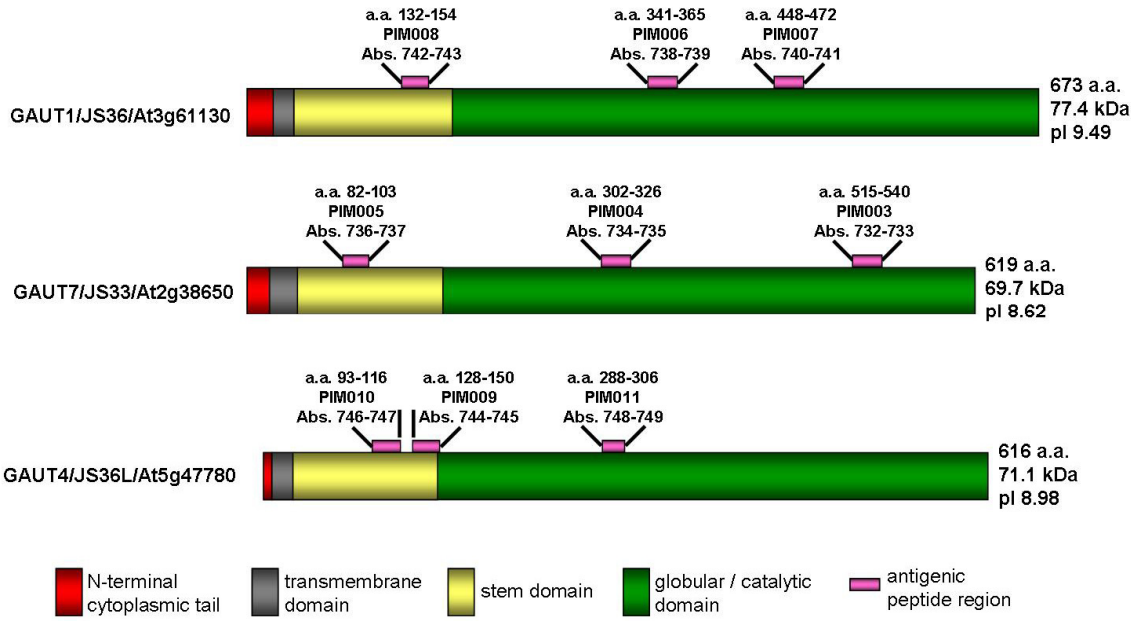


Figure A-1. Schematic of the antigenic peptides of GAUT1, GAUT7, and GAUT4. N-terminal cytoplasmic tails and transmembrane domains were predicted using the HMMTOP v.2 program (http://www.enzim.hu/hmmtop/index.html); stem domains are estimated as approximately 25% of the total length of the protein. The positions of the different domains and of the antigenic peptides are drawn to scale based on the proteins' amino acid lengths.

Indirect enzyme-linked immunosorbent assays (ELISA) against the respective MAPs were employed to determine the serum titer and working dilution (to be used in Western blotting) for different bleeds of each antibody (Figures A-2, A-3, A-4; Table A-2). Serum titer is defined as the highest dilution factor that still gives a detectable response, and is a reflection of the amount of the specific antibody against the target epitope; thus the higher the serum titer, the stronger the antibody is. In general, both GAUT1 and GAUT7 antibodies have relatively higher titer than the anti-GAUT4 antibodies (at least for the second bleeds), suggesting that the selected GAUT1 and GAUT7 peptides are more antigenic than those of GAUT4. The working dilutions calculated from ELISA data were tested by Western blotting to probe the respective MAPs resolved by reducing SDS-PAGE. In general, immunoblots performed at the calculated working dilutions of the antibodies could clearly recognize protein bands corresponding to respective MAPs without giving to much background signal. However, for some of the antibodies, the dilution factors needed to be adjusted slightly in subsequent Western blotting experiments to obtain optimum results.

The antibodies were further characterized to determine if they could recognize their respective GAUTs in their denatured and/or intact (folded) forms. This was analyzed by Western blotting and immunoprecipitation, respectively, using both recombinant proteins and native GAUTs present in Arabidopsis extracts. Native Arabidopsis protein preparations have the advantage that the GAUTs are presumably in the correctly folded and enzymatically active (at least for HG:GalAT activity) form. One consideration, which was important especially in regards to GAUT1, is that analysis by Western blotting provides protein bands that electrophoresed based on their molecular masses. Apriori, only the GAUT predicted sizes, based on amino acid sequences, were known and these predicted masses did not account for any

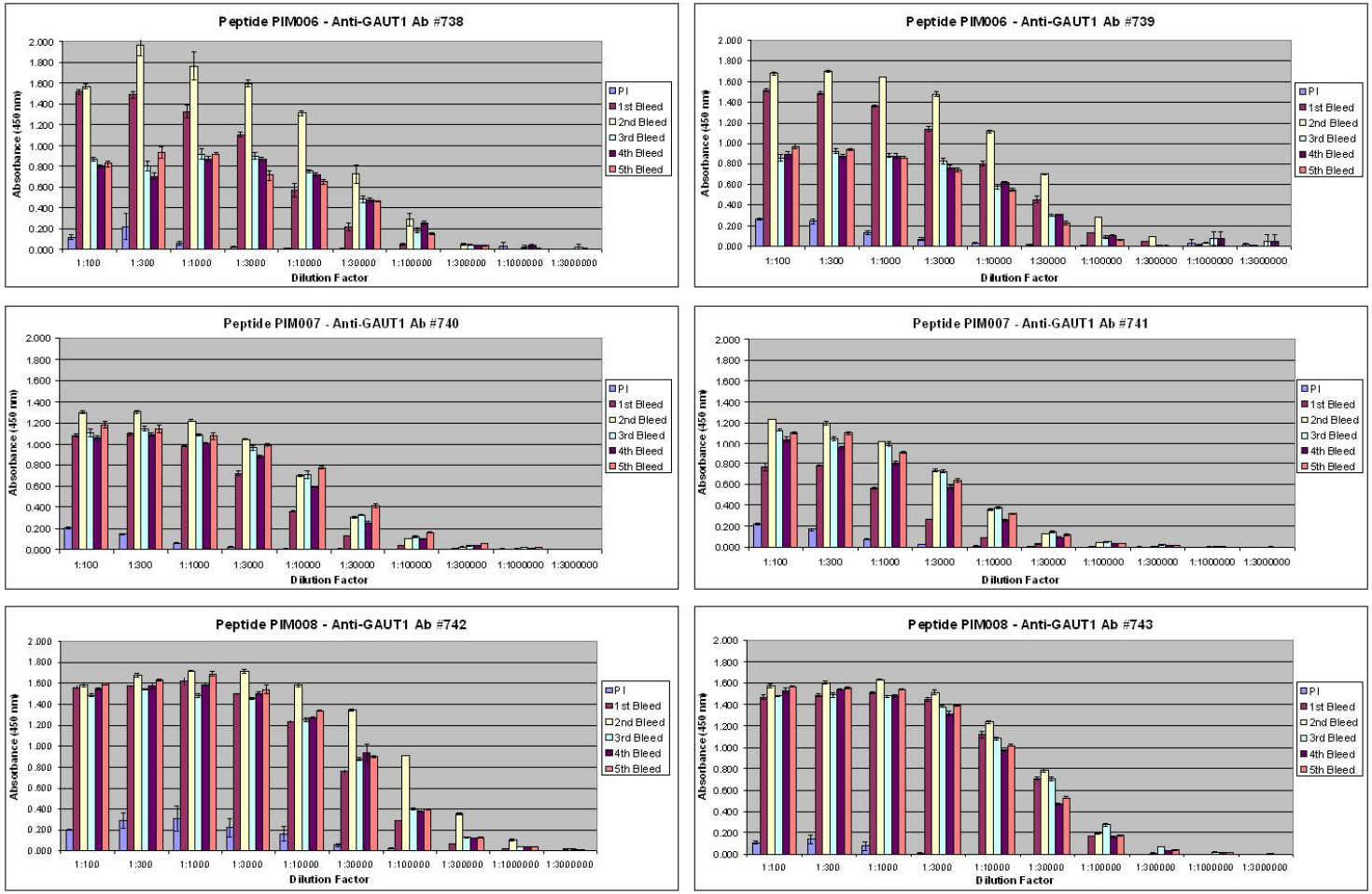


Figure A-2. ELISA of anti-GAUT1 antibodies against respective GAUT1 antigenic peptides.

See Table A-1 for information on the antigenic peptides. Data are average from one experiment done in triplicate ± standard error.

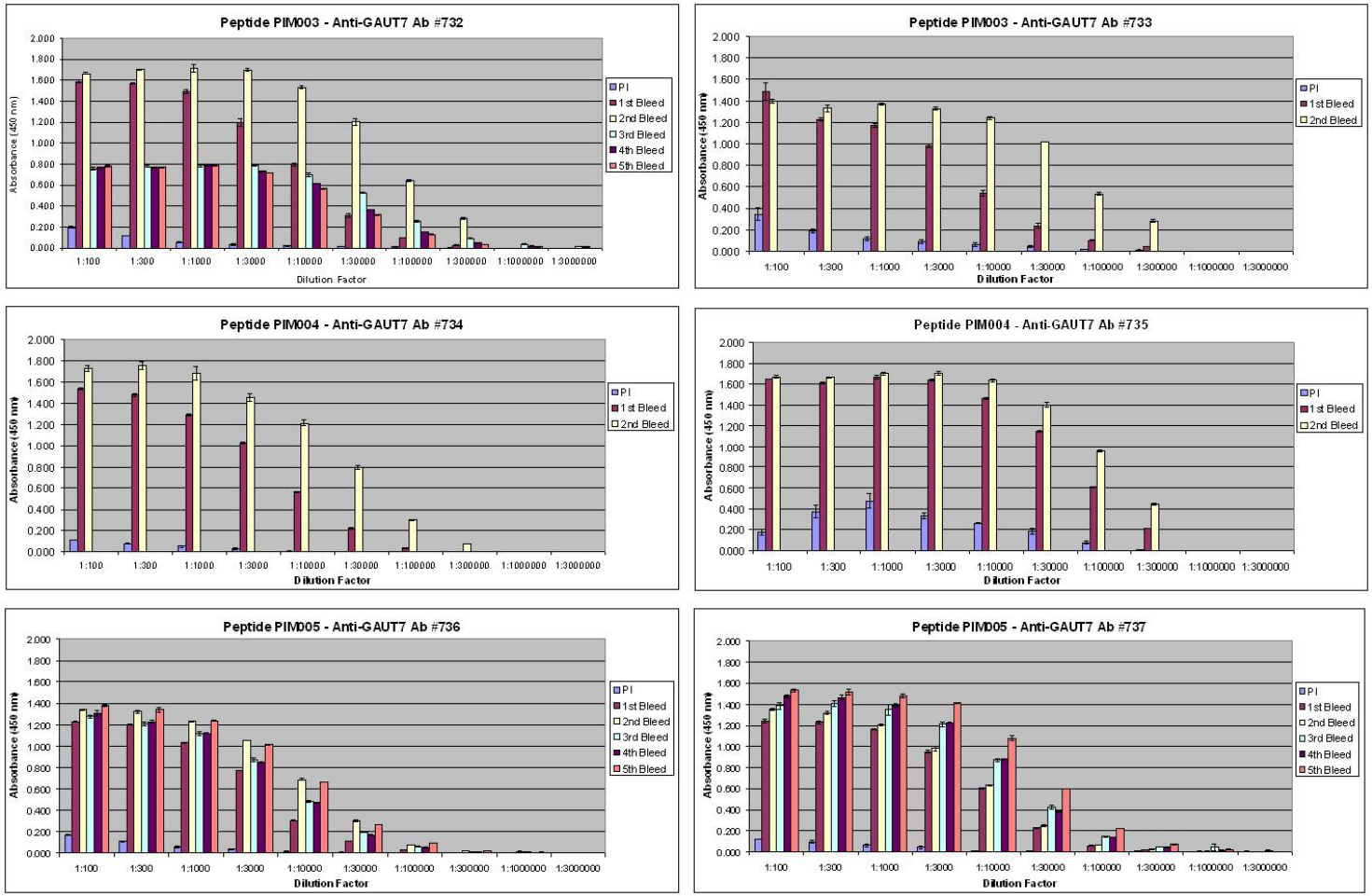


Figure A-3. ELISA of anti-GAUT7 antibodies against respective GAUT7 antigenic peptides.

See Table A-1 for information on the antigenic peptides. Data are average from one experiment done in triplicate ± standard error.

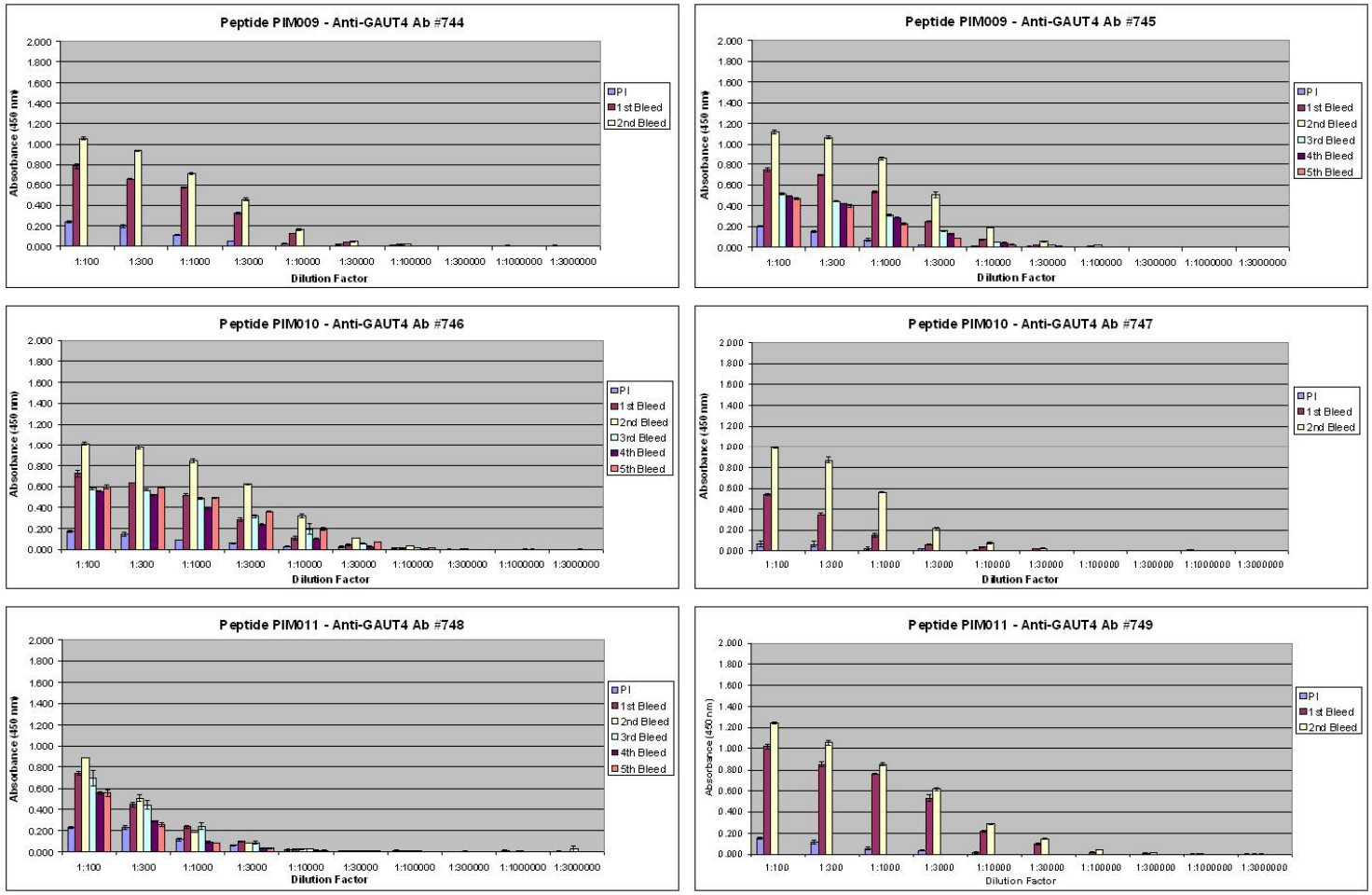


Figure A-4. ELISA of anti-GAUT4 antibodies against respective GAUT4 antigenic peptides.

See Table A-1 for information on the antigenic peptides. Data are average from one experiment done in triplicate ± standard error.

Table A-2. Serum titer and deduced working dilutions of the second bleeds of anti-GAUT1, anti-GAUT7, and anti-GAUT4 antibodies based on the ELISA data.

Titer is presented here as the serum highest dilution (see Figures A-2, A-3, A-4) that still provides a measurable signal. Working dilutions were estimated as the dilution factors giving approximately 75% of the maximum response signal.

Protein / Peptides	Antibody #	Serum titer	Deduced working dilution
GAUT1/JS36/At3g61130			
PIM006	738	1 : 300,000	1 : 10,000
	739	1 : 1,000,000	1 : 5,000
PIM007	740	1 : 300,000	1 : 5,000
	741	1 : 100,000	1 : 2,000
PIM008	742	1 : 1,000,000	1 : 50,000
	743	1 : 100,000	1 : 10,000
GAUT7/JS33/At2g38650			
PIM003	732	1 : 300,000	1 : 20,000
	733	1 : 300,000	1 : 30,000
PIM004	734	1 : 300,000	1 : 5,000
	735	1 : 300,000	1 : 40,000
PIM005	736	1 : 300,000	1 : 3,000
	737	1 : 300,000	1 : 3,000
GAUT4/JS36L/At5g47780			
PIM009	744	1 : 100,000	1 : 300
	745	1 : 100,000	1 : 1,000
PIM010	746	1 : 100,000	1 : 100
	747	1 : 30,000	1 : 100
PIM011	748	1 : 10,000	1 : 100
	749	1 : 300,000	1 : 1,000

potential post-translational modifications. Both Arabidopsis solubilized membrane protein preparations and SP-Sepharose-purified, GalAT activity-containing fractions (SP fractions), were used as native Arabidopsis protein samples. In contrast, recombinant proteins, usually tagged with epitope tags, allow for identification and analysis of the expressed proteins using an anti-tag antibody. GAUT1 and GAUT7 had previously been stably-expressed in human embryonic kidney (HEK293) cells as truncated versions missing the first 41 and 43 a.a., respectively, of the predicted N-terminal cytoplasmic tail and transmembrane domain. These were fused to poly-histidine and hemagluttinin (HA) epitopes, and anti-HA antibody was shown to bind to the recombinant proteins both in Western blots and in immunoprecipitations (Sterling et al., 2006). The recombinant GAUT1 and GAUT7 were used as positive controls for the characterization of the anti-GAUT1 and anti-GAUT7 antibodies. There was, unfortunately, no recombinant GAUT4 available to perform the same assessment for the anti-GAUT4 antibodies.

Western blot analyses of Arabidopsis solubilized enzyme preparations identified variable, complex protein band patterns recognized by the antibodies. Even antibodies generated against the same peptides yielded diverse protein band patterns against solubilized enzyme preparations. Thus, it was virtually impossible to identify to a particular protein band as GAUT1, GAUT7, or GAUT4 (data not shown and Figure A-5). Anti-GAUT1 and anti-GAUT7 antibodies were, however, able to recognize the anti-HA-immunoprecipitated recombinant GAUT1 and GAUT7, respectively (Figure A-5). Furthermore, several of the anti-GAUT1 and anti-GAUT7 antibodies recognized common protein bands in Western blots of the SP fractions (Figure A-5). Four of the anti-GAUT1 antibodies (#738, 739, 740, and 741) recognize protein bands of ~60-70 kDa, a size smaller than the 77.4 kDa expected for GAUT1 based on its amino acid sequence. These results suggest a possible proteolytic-processing of Arabidopsis GAUT1, which is confirmed in the

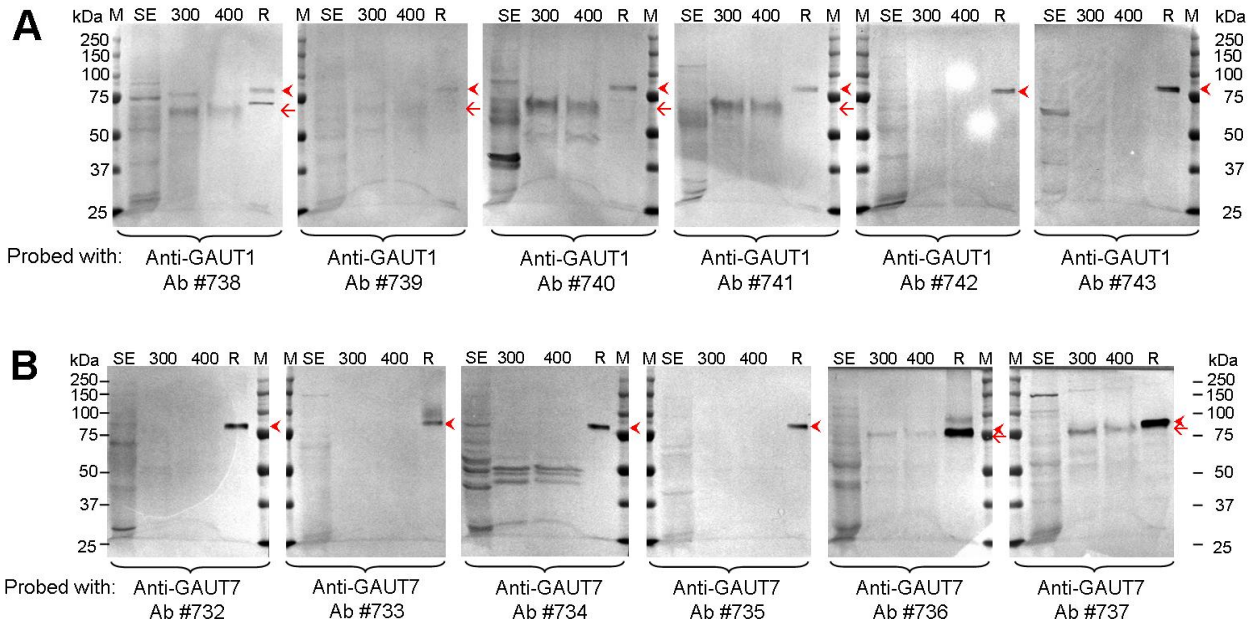


Figure A-5. (A) Anti-GAUT1 and (B) anti-GAUT7 antibodies recognize the denatured form of GAUT1 and GAUT7, respectively, in Western blots of native Arabidopsis and recombinant protein preparations.

Protein samples include: (SE) Arabidopsis solubilized membrane protein preparation; (300) and (400) SP-Sepharose column-purified Arabidopsis solubilized membrane proteins that were eluted using buffers containing 300 mM or 400 mM NaCl, respectively; (R) recombinant GAUT1 (in panel A) or GAUT7 (in panel B) that were immunoprecipitated from lysates of HEK293 cells stably expressing the respective constructs (cell lines 36-25II and 33c6, respectively). The recombinant constructs encode truncated proteins, missing the N-terminal cytoplasmic tail and transmembrane domain, which are tagged with hemaglutinin (HA) epitope to allow immunoprecipitation using a monoclonal anti-HA antibody.

The anti-GAUT1 and anti-GAUT7 antibodies used were the 2nd bleed serum. Samples were loaded at 51 µg/lane for (SE); 5.1 µg/lane for (300) and (400); and 10 µl/lane for (R) (equivalent to the total recombinant protein immunoprecipitated from 1 ml cell lysate using 40 µl conjugate

anti-HA:protein A agarose beads). Red arrows indicate Arabidopsis GAUT1 or GAUT7, red arrow heads indicate recombinant GAUT1 or GAUT7. The indicated antibody numbers represent serum from independent rabbits immunized with GAUT1 (#738 through 743) or GAUT7 (#732 through 737) as described in Figure A-1.

research presented in this dissertation. Two of the anti-GAUT7 antibodies (#736 and 737) recognize protein bands of ~75 kDa, which is larger than the expected 69.7 kDa for GAUT7 based on its amino acid sequence. This larger size of Arabidopsis GAUT7 could be due to post-translational modifications such as N- or *O*-glycosylation and phosphorylation. These ~60 and ~75 kDa protein bands were not recognized by the preimmune serum of the respective antibodies (Figure A-6 and data not shown), suggesting specific binding by the immune serum. It is noteworthy that Arabidopsis solubilized protein preparations seem to be too complex and contain proteins that the antibodies may non-specifically recognize, in comparison to the SP fractions. It therefore is important to consider the purity of the protein fraction when using the anti-GAUT antibodies. The rest of the antibodies, i.e. anti-GAUT1 #742 and 743, and anti-GAUT7 #732 through 735, could recognize the recombinant, but not the native Arabidopsis, proteins on the Western blots. As described in Chapter 3 in this dissertation, in the case of GAUT1, further experiments revealed that this is due to the loss by post-translational protein processing, of N-terminal region of GAUT1 that contains the target peptide epitope. However, the reason for this observation is still not clear for the anti-GAUT7 antibodies.

Anti-GAUT1 and anti-GAUT7 antibodies #740 and #737, respectively, consistently gave the best results in Western analyses, and thus were selected for further use. These antibodies, but not the preimmune sera, were shown to immunoprecipitate the target proteins from both lysates of recombinant HEK293 cell lines and from Arabidopsis SP fraction (Figure A-6, Figure 3.1-A in Chapter 3). Anti-GAUT1 antibody #741 was also found to immunoprecipitate GAUT1 from SP fraction, while antibodies #742 and 743 could not (data not shown).

To ensure that the anti-GAUT1 and anti-GAUT7 antibodies are specific for their target proteins and do not cross-react, antibodies #740 and #737 were tested in reciprocal

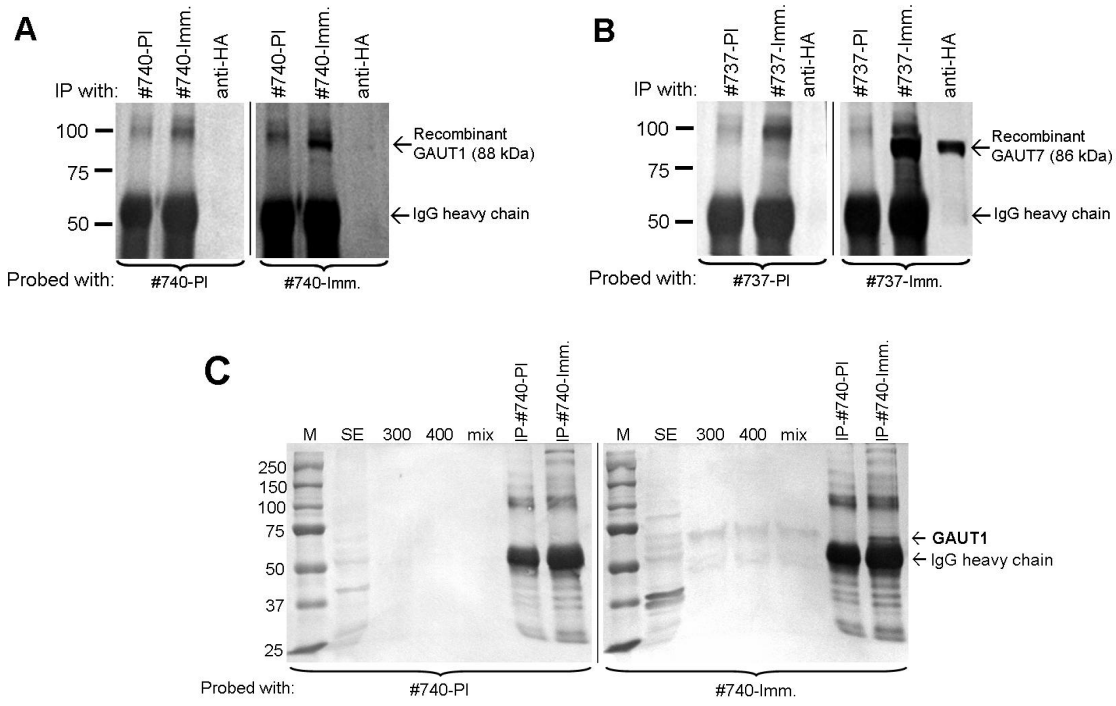


Figure A-6. Anti-GAUT1 antibody #740 (A and C) and anti-GAUT7 antibody #737 (B) recognize intact recombinant (A and B) and native mature, processed Arabidopsis (C) proteins. (A) and (B) Lysates of HEK293 cell line 36-25II (recombinant GAUT1) and 33c6 (recombinant GAUT7), respectively, were immunoprecipitated by antibodies and analyzed by Western blotting as indicated. Immunoprecipitations by anti-HA antibody served as positive controls since the recombinant proteins were tagged with a HA-epitope. Western blots of the same samples probed with anti-GAUT1 antibody #741 and with anti-GAUT7 antibody #736 gave results similar to those in (A) and (B), respectively. (C) Antibody #740 immunoprecipitates the ~60 kDa protein band from Arabidopsis SP fraction.

Abbreviations: (IP) immunoprecipitation; (PI) preimmune serum; (Imm.) immune serum. Protein samples include (SE) Arabidopsis solubilized enzyme; (300) and (400) - SP-Sepharose column-purified Arabidopsis solubilized membrane proteins that were eluted using buffers

containing 300 mM or 400 mM NaCl, respectively; (mix) – a mixture of equal volumes of the SP Sepharose column fractions (300) and (400).

immunoprecipitations of recombinant GAUT1 and GAUT7 followed by Western blot analyses using both antibodies. The results (Figure A-7) showed that each antibody recognizes only its target protein in both intact and denatured forms. The specificity of the antibodies was also demonstrated by the fact that they recognize different protein bands in the Western blot of Arabidopsis SP fractions (Figure A-5).

To verify the identity of the ~60 kDa and the ~75 kDa protein bands recognized in the Arabidopsis SP fractions by the anti-GAUT1 and anti-GAUT7 antibodies, respectively, the protein bands were excised from SDS-PAGE gels and subjected to trypsin in-gel digestion followed by liquid chromatography - tandem mass spectrometry (LC-MS/MS) analysis on a Thermo Finnigan LTQ to obtain peptide sequences (kindly performed by Drs. Jae-Min Lim and Lance Wells, CCRC, GA). Searches against the database of Arabidopsis proteins identified the ~60 kDa protein band as GAUT1 by nine peptides, and the ~75 kDa band as GAUT7 by three peptides, thus confirming the specificity of the anti-GAUT1 and anti-GAUT7 antibodies.

Overall results of the antibody characterization are summarized in Table A-3. Anti-GAUT1 antibodies #740 and 741, and anti-GAUT7 antibody #737, have since been used in multiple Western blotting and immunoprecipitation experiments with consistent results. Anti-GAUT4 antibodies, on the other hand, gave multiple protein band patterns in Western blots of SP fractions (one experiment; data not shown). Availability of a recombinant GAUT4 may help to determine the usefulness of the anti-GAUT4 antibodies.

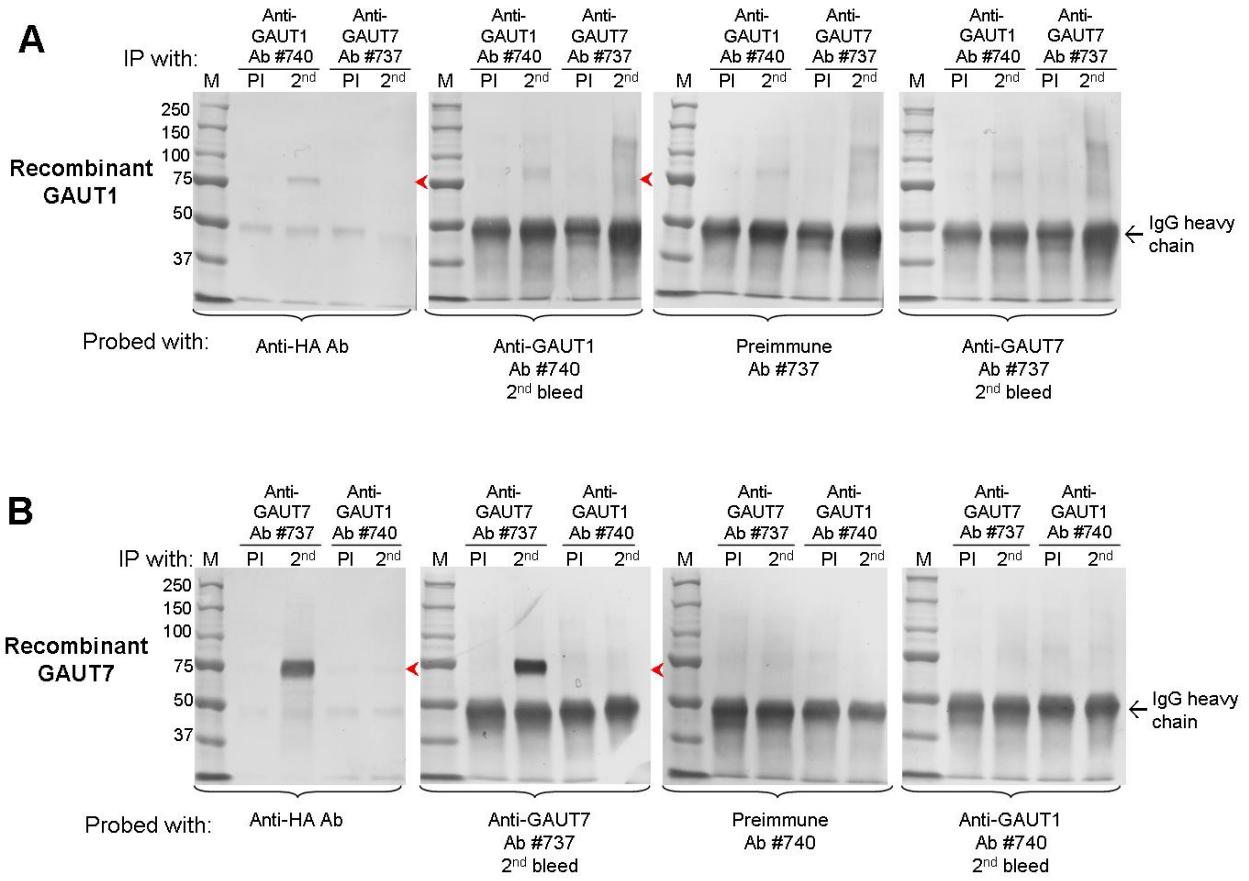


Figure A-7. Anti-GAUT1 (#740) and anti-GAUT7 (#737) antibodies are specific for their respective GAUT proteins.

Protein samples used are the recombinant GAUT1 (in panel A) and recombinant GAUT7 (in panel B). These recombinant proteins were immunoprecipitated using either preimmune, anti-GAUT1, or anti-GAUT7 antibodies, from HEK293 cell lines expressing the respective recombinant constructs (cell lines 36-25II and 33c6, respectively; see Figure A-5 legend and text for more detail on the recombinant constructs).

The immunoprecipitants were analyzed by Western blotting using antibodies as indicated. The left-most Western blots in panel A and B were probed with anti-hemaglutinin (anti-HA) antibody, serving as the positive controls for the immunoprecipitation of the recombinant

GAUT1 and GAUT7. Abbreviations: (IP) immunoprecipitation; (M) molecular weight protein marker; (PI) preimmune serum; (2^{nd}) immune serum 2^{nd} bleed; (HA) hemaglutinin.

Table A-3. Summary of the reactivity of anti-GAUT1, anti-GAUT7, and anti-GAUT4 antibodies against recombinant and native Arabidopsis GAUT1, GAUT7 and GAUT4, respectively, as judged by Western blot and immunoprecipitation analyses.

Yes – The antibody recognizes the respective GAUT protein in the experiment tested.

No – The antibody does not recognize the respective GAUT protein in the experiment tested.

n/a – Result is not available due to the unavailability of recombinant GAUT4.

n/d – The antibody reactivity has not been experimentally determined.

Protein / Peptides	Antibody #	Recognition of recombinant GAUT protein		Recognition of Native Arabidopsis GAUT in SP fraction	
		Western blot	Immuno-precipitation	Western blot	Immuno-precipitation
GAUT1/JS36/At3g61130					
PIM006	738	Yes	n/d	Yes	n/d
	739	Yes	n/d	Yes	n/d
PIM007	740	Yes	Yes	Yes	Yes
	741	Yes	n/d	Yes	Yes
PIM008	742	Yes	n/d	No	No
	743	Yes	n/d	No	No
GAUT7/JS33/At2g38650					
PIM003	732	Yes	n/d	No	n/d
	733	Yes	n/d	No	n/d
PIM004	734	Yes	n/d	No	n/d
	735	Yes	n/d	No	n/d
PIM005	736	Yes	n/d	Yes	n/d
	737	Yes	Yes	Yes	Yes
GAUT4/JS36L/At5g47780					
PIM009	744	n/a	n/a	n/d	n/d
	745	n/a	n/a	n/d	n/d
PIM010	746	n/a	n/a	n/d	n/d
	747	n/a	n/a	n/d	n/d
PIM011	748	n/a	n/a	n/d	n/d
	749	n/a	n/a	n/d	n/d

APPENDIX B

IMMUNOABSORPTION OF GAUT1 AND ASSOCIATED GALAT ACTIVITY FROM ARABIDOPSIS SP FRACTION BY ANTI-GAUT1 ANTIBODY

Transient expression of an N-terminally truncated version of GAUT1 in HEK293 cells and recovery of GalAT activity in immunoprecipitated protein from the culture media surrounding the cells, served as the first evidence that GAUT1 is a HG:GalAT (Sterling, 2004). Stable transformation of the same construct in the same system, however, did not give active GAUT1, with the recombinant protein found in cell lysates instead of getting secreted into the culture media as expected (Sterling, 2004). With the availability of anti-GAUT1 antibodies able to immunoprecipitate GAUT1, we show here that removal of GAUT1 from Arabidopsis protein preparations by immunoprecipitation using anti-GAUT1 antibody results in the depletion of GalAT activity from the supernatant and its recovery in the pellet fraction (Sterling et al., 2006). These results provided another line of evidence that associate GAUT1 with GalAT activity, and thus, strengthen the identity of GAUT1 as a HG:GalAT (Sterling et al., 2006).

Anti-GAUT1 antibody #740, having been shown able to bind to Arabidopsis GAUT1 in both denatured and intact forms (see Appendix A), was used to immunoprecipitate GAUT1 from Arabidopsis SP fraction (SP-Sepharose-purified Arabidopsis solubilized enzyme preparation). This was followed by analyses of the supernatant and pellet fractions by Western blotting to confirm that GAUT1 was precipitated. GalAT activity was also assayed using either radioactive

(UDP-[^{14}C]GalA) or non-radioactive (UDP-GalA) donor substrate. Compositions of typical radioactive and non-radioactive GalAT assays are presented in Table B-1. In the quantitative filter assay (Sterling et al., 2005) radioactive method, GalAT activity is measured as the amount of [^{14}C]-GalA incorporated onto the acceptor substrate oligogalacturonides (OGA, DP 7-23). In contrast, the non-radioactive assay is qualitative or semi-quantitative, with the reaction products (OGAs with higher DP) separated on high percentage polyacrylamide gel electrophoresis (OGA-PAGE) and visualized by alcian-blue/silver nitrate staining as described (Sterling et al., 2006).

Initially, immunoprecipitation was done using protein A agarose beads (Sigma, St. Louis, MO) as the solid support onto which the antibody was conjugated. However, pre-incubation of the SP fraction with the protein A agarose, at an amount usually used in immunoprecipitation, dramatically reduced by ~90% the GalAT activity in SP fraction (Figure B-1-A). As a comparison, SP fraction diluted to approximately the same final volume had only ~10% reduction in GalAT activity (data not shown). One possible explanation for the negative effect of protein A agarose on GalAT activity is that GalAT may bind to the agarose, which has some structural similarity to homogalacturonan (Figure B-1-B). Another observation by Dr. Kerry Caffall in our laboratory on GAUT6, another member of the GAUT gene family, was that *E. coli*-expressed, His-tagged recombinant GAUT6 bound irreversibly to a 6% agarose-based nickel affinity column during purification and could only be released using denaturing agent such as 6 M guanidine, suggesting that the agarose molecule may act as a transition-state analog at the active site of GAUT6, or of GAUT proteins in general (Caffall, 2008). On the other hand, during SP-Sepharose purification of Arabidopsis solubilized membrane proteins, GalAT activity as well as GAUT1 and GAUT7 could routinely be eluted using buffers containing 300 mM and

Table B-1. Compositions of the radioactive and non-radioactive GalAT activity assays.

	Radioactive assay	**Non-radioactive assay**
Assay type	Quantitative	Qualitative / semi-quantitative
Total reaction volume	60 µl	15 µl
Amount of enzyme	30 µl[a]	5 µl[a]
Concentration of donor substrate	7.66 µM UDP-[^{14}C]GalA[b]	3 mM UDP-GalA
Concentration of acceptor substrate	1.5 µg/µl OGAs DP 7-23 = 564.21 µM[d]	0.33 µg/µl OGAs DP 13[c] = 143.07 µM[e]
Molar ratio of donor : acceptor substrate	0.014 : 1	20.969 : 1
Other components: $MnCl_2$ HEPES, pH 7.3 Sucrose KCl BSA	 1.25 mM 50 mM 200 mM 25 mM 0.05% (w/v)	 1.9 mM 50 mM 200 mM 25 mM 0.05% (w/v)
Incubation temperature	30°C	30°C

([a]) Protein concentration of Arabidopsis SP fraction is typically ~ 0.2 - 0.3 mg/ml. ([b]) Assuming use of 8 µl UDP-[^{14}C]GalA stock solution prepared at approx. 20,000 cpm/µl and assuming specific activity of 196 mCi/mmol (1 Ci = 37 GBq) and MW of 580.4. ([c]) OGAs enriched for a single DP are used in non-radioactive GalAT assays (DP 13 is used in this particular example). ([d]) Molar calculation is based on OGAs DP 15 (MW = 2658.6). ([e]) Molar calculation is based on OGAs DP 13 (MW = 2306.52).

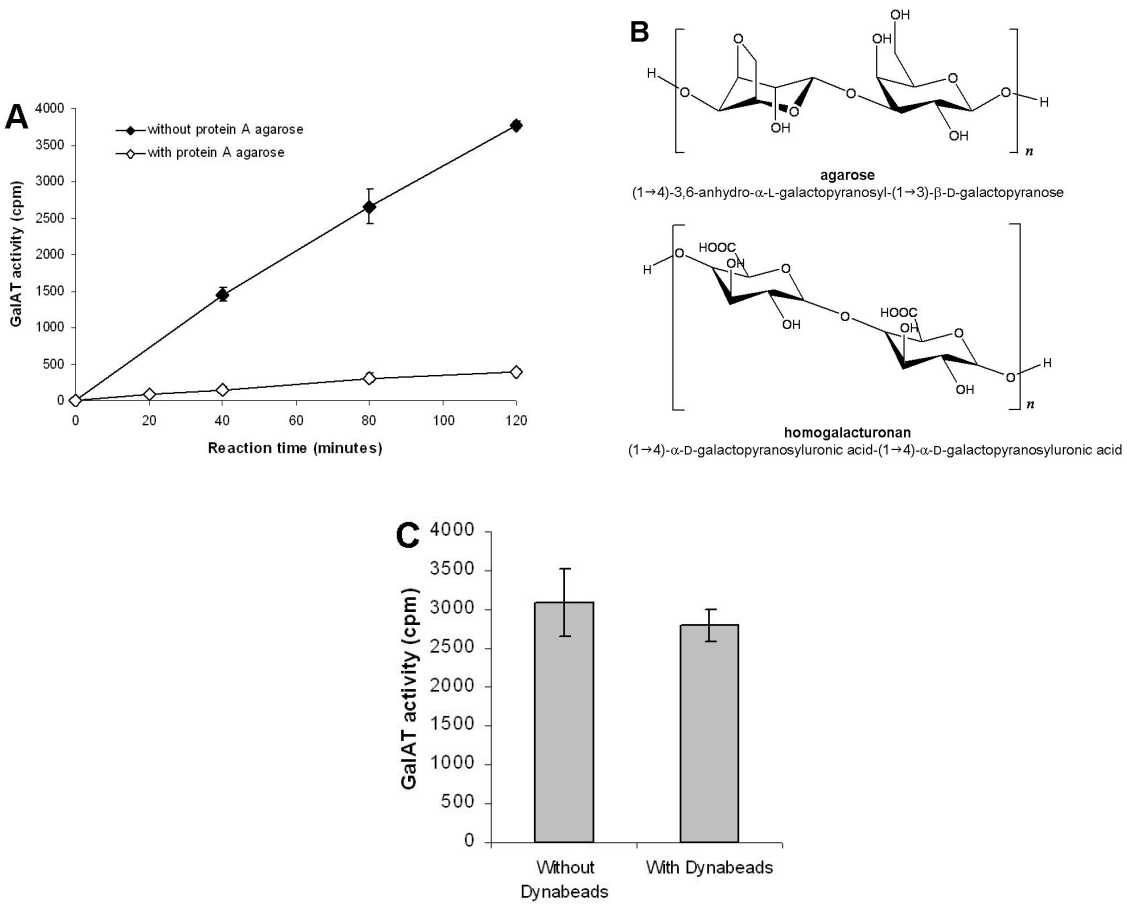


Figure B-1. GalAT activity is negatively affected by agarose beads, but not by magnetic Dynabeads.

(A) Time course assay of GalAT activity in SP fraction following a 2-hour incubation in the absence (◆) or presence (◇) of protein A agarose beads. SP fraction was incubated with protein A agarose at a ratio of 2:1 of SP fraction:drained beads (or 1:1 of SP fraction:50% slurry of beads) at 4°C for 2 hours with end-to-end rotation. After removal of the beads, 30 µl supernatant was subjected to a GalAT activity assay in a 60 µl reaction volume, alongside control SP fraction pre-incubated under the same conditions but without the protein A agarose beads. Data are T0-subtracted average cpm incorporated per reaction ± standard deviation of duplicate reactions.

(B) Structures of the disaccharide units of agarose and homogalacturonan.

(C) GalAT activity assay of SP fraction following a 2-hour incubation in the absence (with dilution to the same final volume as with addition of beads) or presence of Dynabeads M280 Sheep anti-rabbit IgG. GalAT activity assay was carried out in 60 µl reaction volumes for 2 hours. Data shown are T0-subtracted average cpm incorporated per reaction ± standard deviation of duplicate reactions.

400 mM NaCl. This indicates that these proteins do not bind irreversibly to Sepharose beads. It is possible that the irreversible binding of GAUT proteins may occur only with agarose, but not with Sepharose, which is highly cross-linked agarose, or that some GAUT proteins (e.g. GAUT6) may bind more tightly to agarose than other GAUTs (e.g. GAUT1 and 7).

Despite the underlying cause, the above observations rendered protein A agarose not suitable for research with GalAT enzymes, and therefore another alternative solid support was needed that was inert and did not interfere with the GalAT activity. Magnetic Dynabeads M280 Sheep anti-rabbit IgG (Invitrogen, Carlsbad, CA), made from superparamagnetic polymer particles, were found not to significantly affect GalAT activity (Figure B-1-C), and were thus chosen for use in further immunoprecipitation experiments.

To support the transient expression data, anti-GAUT1 antibody-conjugated Dynabeads were used to specifically immunoabsorb GalAT activity from SP fraction (Figure B-2). The use of increasing amounts of anti-GAUT1 serum immunodepleted GalAT activity (up to 88.4%) from the SP fraction, while the activity was recovered (up to 35.3%) in the immunoabsorbed material (Figure B-2-A). This depletion coincided with the gradual disappearance of the GAUT1 protein band from the immunodepleted fractions as shown by Western blot analysis (Figure B-2-B), while the immunoabsorbtion of GalAT activity coincided with the gradual appearance of the same protein band in the immunoabsorbed fractions (Figure B-2-C). Control immunoprecipitation done using the preimmune serum did not immunodeplete/immunoabsorb GAUT1 nor GalAT activity (Figure B-2), verifying the above results as specifically due to the anti-GAUT1 antibody.

The product made by the immunoabsorbed proteins was characterized by enzymatic digestion with a HG-specific exopolygalacturonase (EPG; from *Aspergillus tubingensis*), which

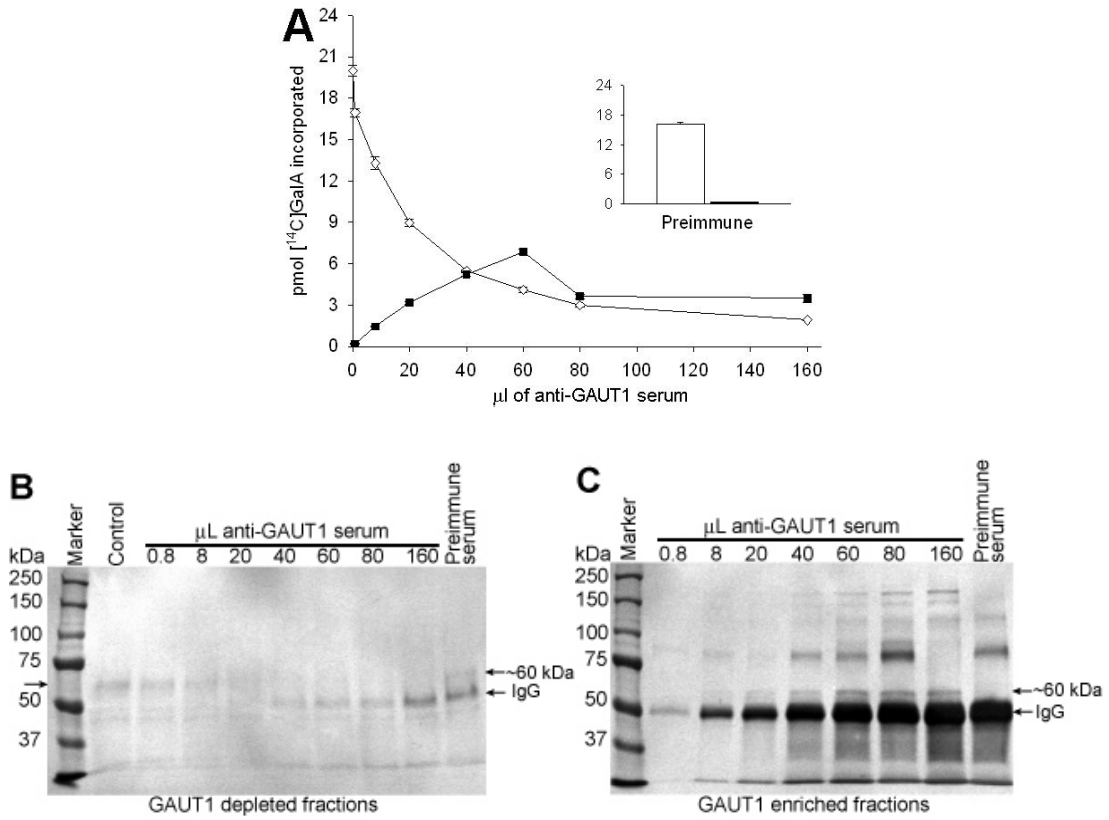


Figure B-2. Anti-GAUT1 serum immunoabsorbs GAUT1 and associated GalAT activity from Arabidopsis SP fraction. (Re-printed from Sterling et al., 2006).

(A) GalAT activity is immunodepleted from SP fractions (◇) by increasing amounts of GAUT1 antiserum (0-160 µl of antiserum applied per 240 µl SP fraction), while the activity is retained in the immunoabsorbed proteins (■). Inset: similar fractions treated with 160 µl preimmune serum (open bar = immunodepleted fraction; closed bar = immunoabsorbed fraction). (B) and (C) Western blots of GAUT1 antiserum-depleted fractions and of GAUT1 antiserum-immunoabsorbed (GAUT1-enriched) fractions, showing the gradual disappearance and appearance, respectively, of the ~60 kDa GAUT1 protein band with higher amounts of antiserum. IgG: IgG heavy chain detected by the secondary antibody used in Western blotting. The control was a fraction treated using160 µl preimmune serum. Samples were loaded at 30 µl/lane.

specifically cleaves GalA residues in α-1,4-linkages. The synthesized product was sensitive to the EPG treatment, resulting in the bulk of product (98.9%) being digested, in comparison to controls treated with mock (water) or boiled EPG (Figure B-3, compare bars 7 and 8 to bar 9). This result indicates that the synthesized product is HG, and therefore confirms that the activity immunoprecipitated by the anti-GAUT1 serum is indeed a GalAT.

Similar results were obtained using the non-radioactive GalAT assay on the immunoabsorbed proteins. SP fraction, GAUT1-depleted fraction, and GAUT1-enriched fraction were incubated in a GalAT reaction for 0, 3, and 22 hours, at 30°C with donor substrate UDP-GalA and acceptor substrates DP-13-enriched OGAs. Electrophoresis separation of the GalAT reaction products on a 30% polyacrylamide gel showed that the anti-GAUT1 immunoabsorbed material (GAUT1-enriched fraction) incorporated multiple GalA residues onto the acceptor substrates in a time-dependent manner, resulting in oligomers of increasing sizes (up to DP 21 after 22-hr incubation) which were digestible by EPG (Figure B-4, lanes 10-13). In contrast, the GAUT1-depleted fraction gave less GalA incorporation onto the OGA substrates (up to DP 18) in comparison to the initial SP fraction (up to DP 25) (Figure B-4, compare lanes 6-9 to lanes 2-5), indicating reduction of the GalAT activity in the immunodepleted fraction. These results substantiate the data from the radioactive GalAT assay showing that the anti-GAUT1 serum immunoabsorbed GalAT activity, and further confirm that GAUT1 is a HG:GalAT.

It is noteworthy that the degree of GalA incorporation onto the elongating OGA acceptors by solubilized GalAT appears to be dependent on the molar ratio between the donor substrate UDP-GalA and the acceptor substrate OGAs. When acceptor OGAs are present in molar excess in comparison to UDP-GalA, a condition used in the quantitative radioactive

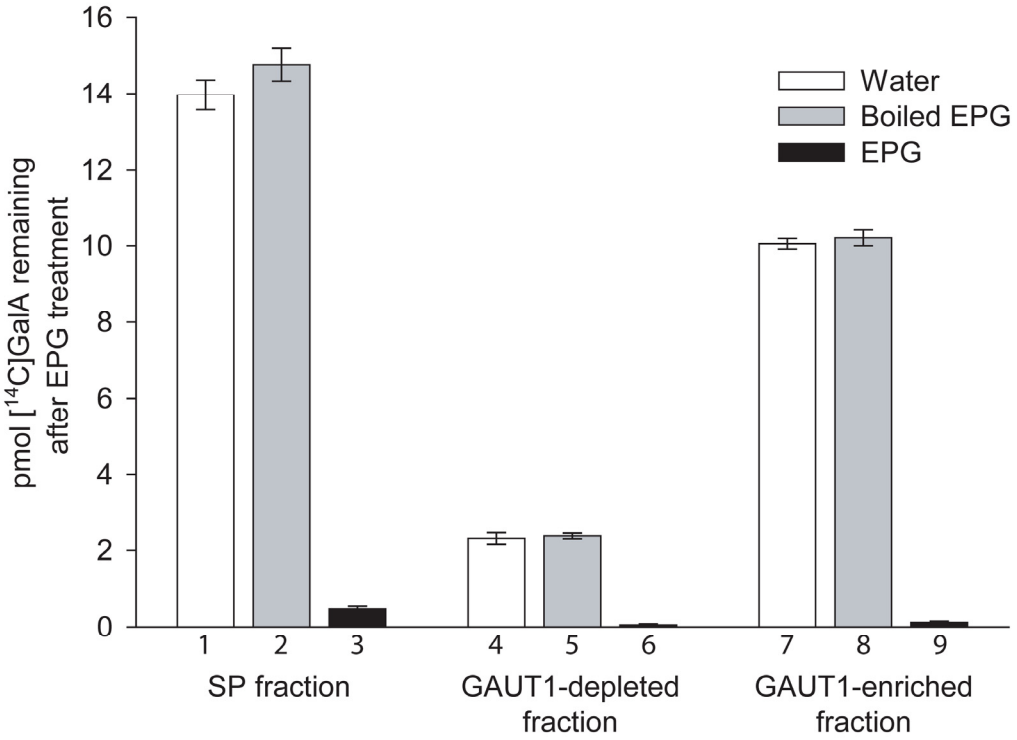


Figure B-3. Exopolygalacturonase (EPG) digestion of the products made by anti-GAUT1-immunoabsorbed GalAT. (Re-printed with modification from Sterling et al., 2006).

Products were synthesized in 2-hour GalAT reactions consisting of buffer, UDP-[^{14}C]GalA, OGA acceptors, and Arabidopsis SP fraction (lanes 1-3), SP fraction following GAUT1 immunodepletion (GAUT1-depleted fraction; lanes 4-6), or anti-GAUT1-immunoabsorbed material (GAUT1-enriched fraction; lanes 7-9). Following termination of GalAT reactions, the reaction mixture was adjusted to pH 4.5 by addition of 10 µl of 2 M acetic acid and 4.2 µl of 1 M sodium acetate buffer, pH 4.2. Each product sample was incubated overnight at 30°C with 2 µl of water, purified EPG (from *Aspergillus tubingensis,* EC 3.2.1.67, 0.54 mg/ml, 262 U/mg, 1 U = 1 µmol of reducing sugar produced/minute, a kind gift from Dr. Carl Bergmann, CCRC), or boiled EPG. The digestion reaction was terminated by addition of 23 µl of 1 M NaOH, and the final mixture was processed using the filter assay method (Sterling et al., 2005).

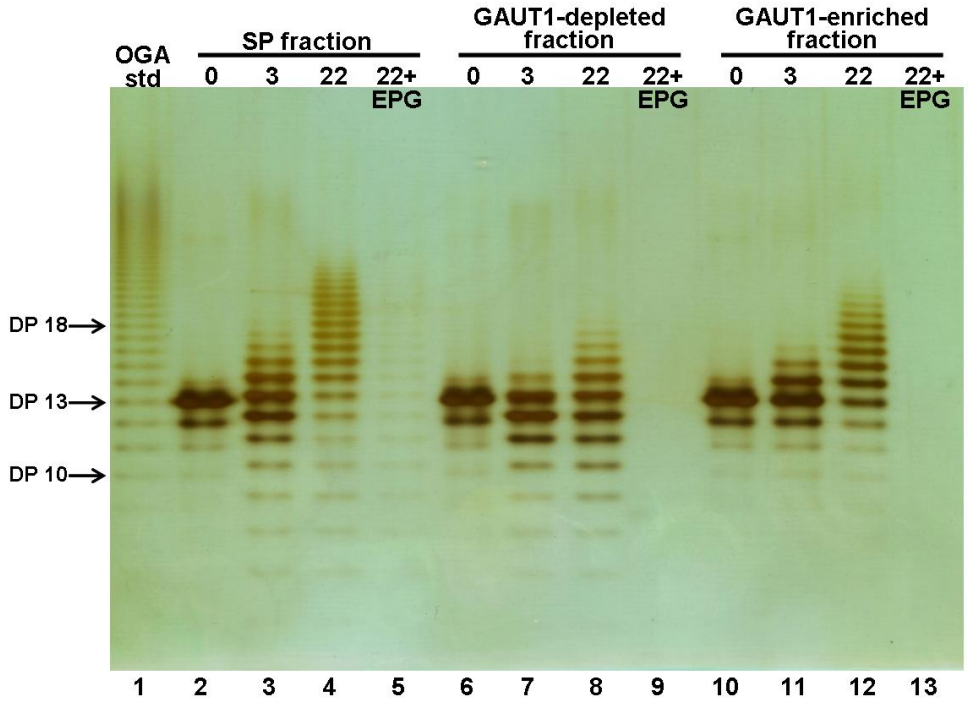


Figure B-4. Separation of non-radioactive GalAT reaction products synthesized by anti-GAUT1-immunoabsorbed GalAT using 30% polyacrylamide gel electrophoresis (OGA-PAGE) (GalAT reactions and OGA-PAGE were performed by Sarah E.I. Inwood). (Re-printed from Sterling et al., 2006).

Products were synthesized in GalAT reactions containing buffer, UDP-GalA, OGA acceptors, and enzyme samples of Arabidopsis SP fraction (lanes 2-5), SP fraction following GAUT1 immunodepletion (GAUT1-depleted fraction, lanes 6-9), or anti-GAUT1 immunoabsorbed material (GAUT1-enriched fraction, lanes 10-13) for 0, 3, or 22 hrs as noted. Lanes 5, 9 and 13 are the respective 22-hr reaction products after digestion with EPG. Lane 1 is a standard mixture of OGAs DP 7-23.

GalAT assay (Table B-1), solubilized GalAT is able to add only one GalA residue onto OGA acceptors, as previously reported for tobacco enzyme (Doong and Mohnen, 1998). In contrast, when UDP-GalA is used in molar excess compared to OGAs, as employed in the non-radioactive GalAT assay (Table B-1), solubilized Arabidopsis enzymes are able to successively transfer several (up to >10) GalA residues onto the OGAs. Similar observations have also been reported for GalAT from petunia pollen tubes (Akita et al., 2002).

A closer look on the band patterns of the separated GalAT reaction products on the OGA-PAGE revealed interesting observations that may otherwise be overlooked. While incubation with the GalAT activity-containing samples clearly resulted in elongating OGAs, OGA bands of smaller DP (as small as DP 7; Figure B-4 lanes 3-4, 7-8) than the starting OGA supplied in the GalAT reactions (mixture of DP 9-14, but particularly enriched for DP 13; Figure B-4 lanes 2 and 6) were also observed following 3- and 22-hr reactions with the SP fraction and the GAUT1-depleted fraction, suggesting that some kind of OGA degrading activity is present in these samples. Moreover, after the incubations the OGAs appear to have a slight increase in size in comparison to the starting OGAs and to their OGA standard counterparts of the same DPs (e.g. in Figure B-4, OGA band of DP 12 in lane 3 does not align parallel to those in lanes 1 and 2, but appears to resolve at a position in between OGAs DP 12 and DP 13 in those lanes), suggesting that the OGAs may have become modified in someway. This size increase is estimated on average to be approximately 1/3 of the distance between bands of OGA standard, which equals to one GalA difference of 176.04 Da (MW of GalA monomer (194.04 Da) minus MW of H_2O (18 Da) removed in the formation of glycosidic bonds), suggesting a possible size increase of ~60 Da. Both the degradation and modification phenomena were only observed following incubation with SP fraction or in the GAUT1-depleted fraction, but not with the GAUT1-

enriched fraction, indicating that the immunoprecipitation with the anti-GAUT1 antibody managed to separate the GAUT1 GalAT activity from the activities that degrade and/or modify the OGAs. It is not clear if these two latter activities are related or independent from each other. Future elucidation of the identities of these OGA modifying activities shall provide a better understanding of the HG synthesis mechanism.

APPENDIX C

N-TERMINAL SEQUENCING OF GAUT1

Two lines of evidence led to a hypothesis that Arabidopsis GAUT1 is post-translationally modified by proteolytic cleavage at the N-terminus. Deduced from its gene sequence, GAUT1 is predicted to be a protein of 673 a.a. with a molecular weight of 77.4 kDa. However, Arabidopsis GAUT1 was consistently observed by Western blotting to be a ~60 kDa protein band, presenting ~17 kDa difference in molecular weight from the predicted to the observed size. In addition, all three sets of anti-GAUT1 polyclonal antibodies raised against three peptide regions of GAUT1 were able to recognize the truncated recombinant GAUT1 (missing the first 41 a.a. at the N-terminus) stably expressed in HEK293 cells, but only two of them (generated against GAUT1 peptide regions at a.a. positions 341-365 and 448-472, respectively) could detect the Arabidopsis GAUT1 in SP fraction (see Appendix A). The other antibody set, raised against GAUT1 peptide at a.a. positions 132-154, could not recognize the Arabidopsis protein, probably due to the loss of epitope from the hypothesized proteolytic cleavage of Arabidopsis GAUT1.

In order to test the above hypothesis, an attempt was made to sequence the N-terminus of Arabidopsis GAUT1 by subjecting the GAUT1 protein band to N-terminal sequencing (Edman degradation). GAUT1-enriched sample was prepared by immunoprecipitation of the GAUT1:GAUT7 protein complex from Arabidopsis SP fraction using anti-GAUT7-specific IgG-beads (anti-GAUT7 specific IgG covalently coupled to magnetic Dynabeads M270 Epoxy

(Invitrogen, Carlsbad, CA). Anti-GAUT7-specific-IgG beads were used instead of that of anti-GAUT1, since anti-GAUT7 antiserum was about 3 times stronger than anti-GAUT1 antiserum based on ELISA). The immunoabsorbed proteins were subsequently eluted from the beads. To obtain enough GAUT1 protein (minimum of 10 pmol required) for N-terminal sequencing, this immunoprecipitation-elution process was repeated 7 times (90 μg of IgG bead was applied to 3 ml of fresh SP fraction in each cycle of immunoprecipitation). The eluted protein solutions were pooled, concentrated, and resolved on reducing SDS-PAGE alongside pre-stained protein markers and a small aliquot of un-concentrated eluted proteins. The separated proteins were subsequently transferred onto a PVDF membrane (PVDF-PSQ, Millipore, Billerica, MA). The lane containing the concentrated sample was stained by Coomassie Brilliant Blue R-250, while the lane containing the un-concentrated sample was probed with anti-GAUT1 serum to detect GAUT1 protein band. By comparison to this immunoblot, the position of the GAUT1 protein band on the Coomassie-stained membrane was determined. A PVDF membrane piece containing the GAUT1 protein band was excised and sent for N-terminal sequencing at the W.M. Keck Facility at Yale University (New Haven, CT).

The 8-cycle N-terminal sequencing analysis yielded a complex mixture of very low level amino acids in each cycle that did not give a unique definitive N-terminal sequence for GAUT1 (Figure C-1A). Upon an additional amino acid analysis of the sample PVDF membrane, it was found that the submitted sample contained ~8 pmol of proteins of ~60 kDa (based on the observed ~60 kDa size of the mature Arabidopsis GAUT1). Since the average yield of N-terminal sequencing from PVDF membrane is 15%, the ~8 pmol sample would have given a level of sequencing at ~1.2 pmol that was in agreement with the N-terminal sequencing yields of 0.97 pmol for methionine in cycle 1 and 0.74 pmol for asparagine in cycle 3 (Figure C-1). Thus,

Figure C-1. (A) The result of 8-cycle N-terminal sequence analysis of the GAUT1 protein band, and (B) GAUT1 full-length protein sequence showing several possible N-terminal sequences of mature Arabidopsis GAUT1 based on the N-terminal sequencing result.

(A) In each cycle, the amino acid listed at the bottom is the most abundant. The higher the amino acid in each column, the less it is in abundance. Amino acids in parentheses are all present at approximately the same level, but much lower than those not in parentheses.

(B) Underlined in black (labeled a, b, and g) are the three matching sequences suggested by the W.M. Keck Facility; underlined in blue (labeled c through f) are the four matching sequences based on manual comparison of the N-terminal sequencing result and the GAUT1 protein

sequence. Thicker underline indicates amino acids detected in the sequencing analysis, while thinner underline refers to those deduced from the GAUT1 protein sequence. Highlighted in yellow are the first 41 a.a. of GAUT1, including predicted transmembrane domain; in green are the cysteine residues; in red is the lysine residue at a.a. position 100; in pink are the three peptide regions against which anti-GAUT1 antibodies were generated.

Note: According to the W.M. Keck Facility, Arg and His are often underestimated during sequence analysis of protein samples on PVDF membranes. Therefore, it is likely that the Arg residue N-terminal of the sequence ANELVQ (labeled 'g' in the figure) was not detected in the first cycle of the analysis. We therefore propose that the N-terminal sequence of GAUT1 starts at Arg_{167}, and not at Ala_{168}.

the data verified that the low yield of the N-terminal sequencing analysis was due to a low amount of sample, and not because of possible blockage by acetylation of the N-termini of the analyzed proteins. Moreover, the amino acid composition of the sample did not completely match that of GAUT1, suggesting that the sample indeed contained a mixture of proteins (Table C-1). It was apparent that the scaled-up immunoaffinity purification effort failed to sufficiently purify the very low abundance GAUT1 in enough quantity and away from other contaminating proteins of similar molecular weight, and thus complicated the sequencing analysis.

The W.M. Keck Facility, by using software GPMAW, suggested three possible matching sequences from the N-terminal sequencing result (Figure C-1B, labeled a, b, and g). However, each one of these sequences was not a consecutive match (having gap(s) within the proposed sequence). We also tried to deduce the possible GAUT1 N-terminus by manually comparing various permutations of the amino acid residues detected in the N-terminal sequencing analysis to the GAUT1 protein sequence (within the first ~200 a.a., where the hypothesized proteolytic cleavage is likely to take place) and then calculating the molecular weights of the resulting "truncated" GAUT1 proteins having the matching sequences as the N-termini. From this exercise, four sequences were found (Figure C-1B, labeled c, d, e, f) that match GAUT1 protein sequence without gaps within each sequence. Of all these possible N-terminal sequences (Table C-2), sequence RANELVQ at a.a. position 168-174 of GAUT1 would give a truncated protein that is closest in size (58.6 kDa) to the experimentally observed GAUT1. We therefore proposed this sequence as the N-terminal sequence of Arabidopsis GAUT1, and the proteolytic cleavage to occur between the methionine and arginine residues at a.a. position 167 and 168 of GAUT1, respectively. It is important to bear in mind that this N-terminal sequencing result is not by itself conclusive, since none of the sequences was a 100% match, and therefore should only be used as

Table C-1. Amino acid analysis of the sample submitted for N-terminal sequencing.

The analysis result is also compared to the amino acid profile (in mol %) of full length or truncated (with proteolytic cleavage assumed to occur between a.a. position 167 and 168) versions of GAUT1 (the last two columns). The acid hydrolysis method used for this analysis converts Asn and Gln into Asp and Glu, respectively. In the subsequent HPLC, Cys co-elutes with Pro, and Met sulfoxide (a common oxidation product in peptides/proteins) co-elutes with Asp, thus the observed values of Met may be low and those of Asp and Pro may be somewhat high. n/a – data not available.

Amino acid	Quantity				mol % in GAUT1	
	nmol	μg	mol %	# residues	Full length	Truncated
Ala	0.259	0.018	7.4	40.6	6.2	6.9
Cys	n/a	n/a	n/a	n/a	1.2	1.2
Asp + Asn	0.269	0.031	7.6	42.1	13.3	14.0
Glu + Gln	0.610	0.079	17.4	95.6	7.5	7.7
Phe	0.174	0.026	4.9	27.2	3.7	3.4
Gly	0.330	0.019	9.4	51.8	4.6	4.0
His	0.037	0.005	1	5.8	3.0	3.4
Ile	0.235	0.027	6.7	36.8	4.5	4.2
Lys	0.104	0.013	2.9	16.2	7.9	8.3
Leu	0.441	0.050	12.5	69.1	11.4	12.1
Met	0.052	0.007	1.5	8.1	2.8	3.2
Pro	0.139	0.013	3.9	21.7	4.5	4.7
Arg	0.088	0.014	2.5	13.8	6.4	5.3
Ser	0.223	0.019	6.3	34.9	7.7	5.9
Thr	0.155	0.016	4.4	24.2	3.7	3.8
Val	0.296	0.029	8.4	46.4	5.8	5.1
Trp	n/a	n/a	n/a	n/a	1.9	2.2
Tyr	0.105	0.017	3	16.5	3.9	4.7

Table C-2. The possible N-terminal sequences of GAUT1 and the properties of the resulting truncated proteins. Underlined are the amino acids identified in the N-terminal sequencing analysis.

N-terminal peptide sequence	a.a. position	Resulting protein		
		Length (a.a.)	MW (kDa)	pI
a - MALKRGL	1 - 7	673	77.4	9.49
b - ALKRGLSG	2 - 9	672	77.2	9.50
c - VLVLLI	23 - 28	651	75.2	9.35
d - KEILDVIA	78 - 85	596	68.8	9.39
e - GPLSL	92 - 96	582	67.4	9.46
f - RANELVQ	168 - 174	506	58.6	9.30
g - LENAAI	183 - 188	491	56.9	9.35

supporting evidence. It is nevertheless in agreement with other lines of evidence such as the mass spectrometry sequence coverage and the transient expression of truncated-GAUT1 GFP-fusion as presented in Chapter 3, in support of a proteolytically-processed Arabidopsis GAUT1.

APPENDIX D

CHARACTERIZATION OF ARABIDOPSIS SP FRACTION: EFFECTS OF REDUCING AGENTS, OGA SUBSTRATE CONCENTRATION, AND DIVERSE PECTIC SUBSTRATES ON GALAT ACTIVITY, AND THE PRESENCE OF PECTIN METHYLTRANSFERASE ACTIVITY

As part of the effort to study GAUT1 and GAUT7 proteins described in this thesis, GalAT activity from Arabidopsis solubilized membrane protein preparation was routinely purified over a SP-Sepharose column based on the protocol developed by Dr. Jason Sterling (Sterling, 2004; Sterling et al., 2006). This cation-exchange chromatography procedure was modified to a 10 ml column, onto which 15 ml of Arabidopsis solubilized membrane proteins are loaded and the bound proteins eluted using a NaCl step gradient (10 ml each of 100 mM, 200 mM, 300 mM, 400 mM, 500 mM, and 600 mM NaCl) in SP buffer (50 mM HEPES pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 25% (v/v) glycerol, and 1% (v/v) Triton X-100). Consistent results have been repeatedly obtained, with GalAT activity as well as the GAUT1 and GAUT7 proteins being typically recovered in the fraction eluted with buffer containing 300 mM (300 mM (I) and (II) fractions) and in the first half of the fraction eluted with buffer containing 400 mM NaCl (400 mM (I) fraction) (Figure D-1). Following desalting on PD-10 columns (GE Healthcare, Piscataway, NJ) to remove the salt, these fractions are pooled and named Arabidopsis SP fraction, which has been used as starting material for various experiments described in this

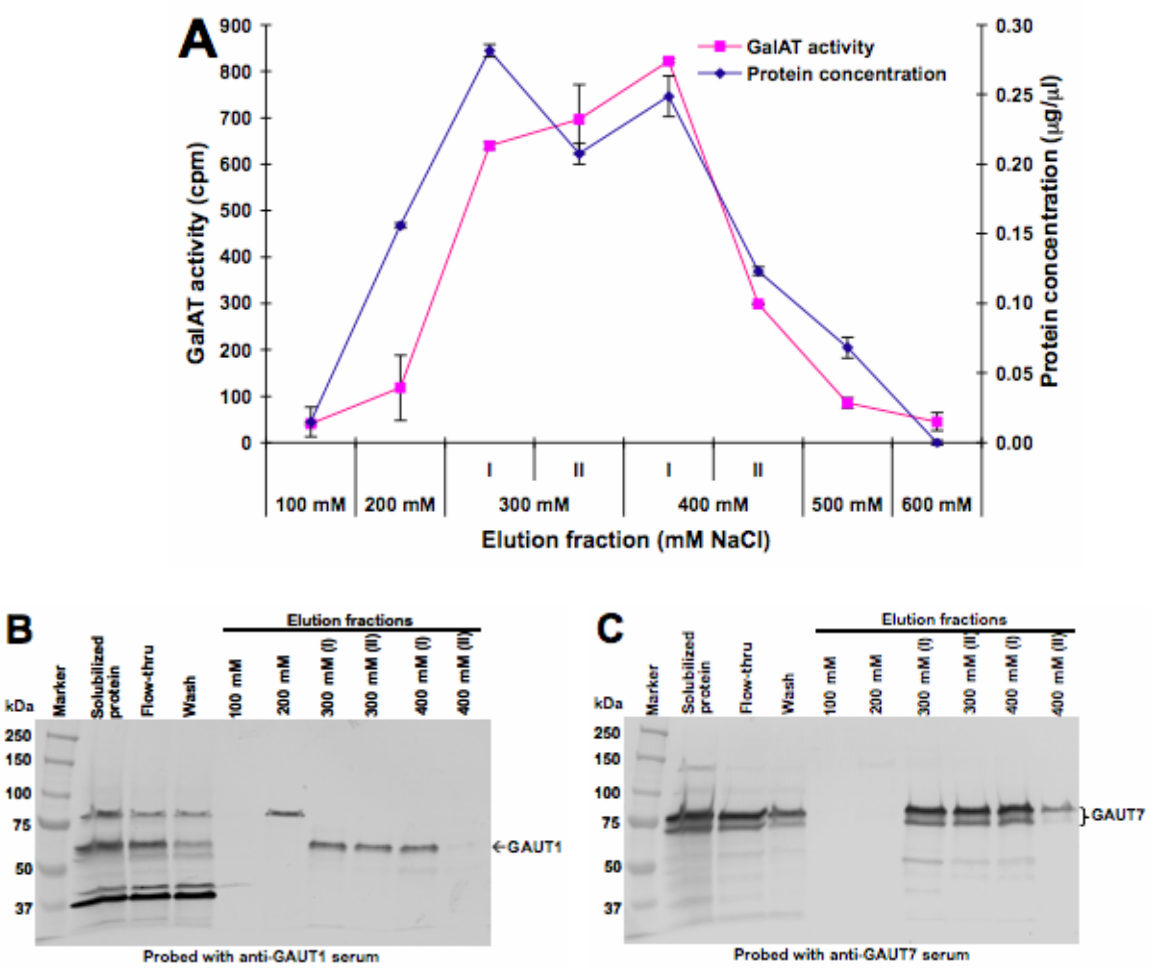


Figure D-1. Typical elution profile of SP-Sepharose column purification of Arabidopsis solubilized enzyme. (A) GalAT activity and protein concentration. GalAT activity was assayed in a 30 µl reaction containing buffer, 80 µg of OGA (DP 7-23), 5.75 µM UDP-[^{14}C]GalA, and 10 µl of enzyme. Reactions were incubated at 30°C for 40 minutes, and subsequently processed using the filter assay method (Sterling et al., 2005). Data shown are background subtracted values ± standard error of duplicate reactions for the GalAT assay and of triplicate reactions for protein concentration assay. (B) and (C) Western blots probed with anti-GAUT1 and anti-GAUT7 serum, respectively. An equal volume of each fraction (30 µl) was loaded in each lane.

dissertation, including immunoprecipitations and proteomics analyses. Here several characteristics of the SP fraction are described, especially regarding its GalAT activity, which were observed during the course of optimization for the reported experiments.

Effect of reducing agents on GalAT activity

GAUT1 and GAUT7 have both been shown to resolve at a higher molecular weight on non-reducing SDS-PAGE, suggesting that these two proteins form a heteromeric core complex held together by disulfide bond(s) (see Chapter 3). With regards to the GalAT activity, this finding led to the question of whether these disulfide bond(s) are necessary for the enzymatic activity of the GAUT1:GAUT7 GalAT complex. In an attempt to answer this question, the effect of two reducing agents, β-mercaptoethanol (2-ME) and dithiothreitol (DTT), on GalAT activity was tested in the Arabidopsis SP fraction. These reducing agents are generally used to prepare samples for reducing SDS-PAGE to break intramolecular disulfide bonds so as to ensure that the sample proteins are completely denatured, and are typically applied at final concentrations of 1% (v/v) (128 mM) and 25 mM for 2-ME and DTT, respectively. On the other hand, 2-ME is also routinely used in the homogenizing buffer at 0.1% (v/v) concentration to prevent action of proteases during preparation of Arabidopsis microsomal membranes, with no apparent effect on GalAT activity (Guillaumie et al., 2003).

In this experiment, the GalAT assay reaction mixture was supplemented with a range of concentrations of 1, 5, 25 and 100 mM DTT and 0.01, 0.1, 1, and 5 % (w/v) 2-ME (equal to 1.28, 12.8, 128, and 640 mM 2-ME, respectively), and incubated with the enzyme sample for 2 hrs. The results indicated that the reducing agents have a negative effect on GalAT activity (Figure D-2). At 0.01% (1.28 mM) and 0.1% (12.8 mM), 2-ME did not show any effect, which is in

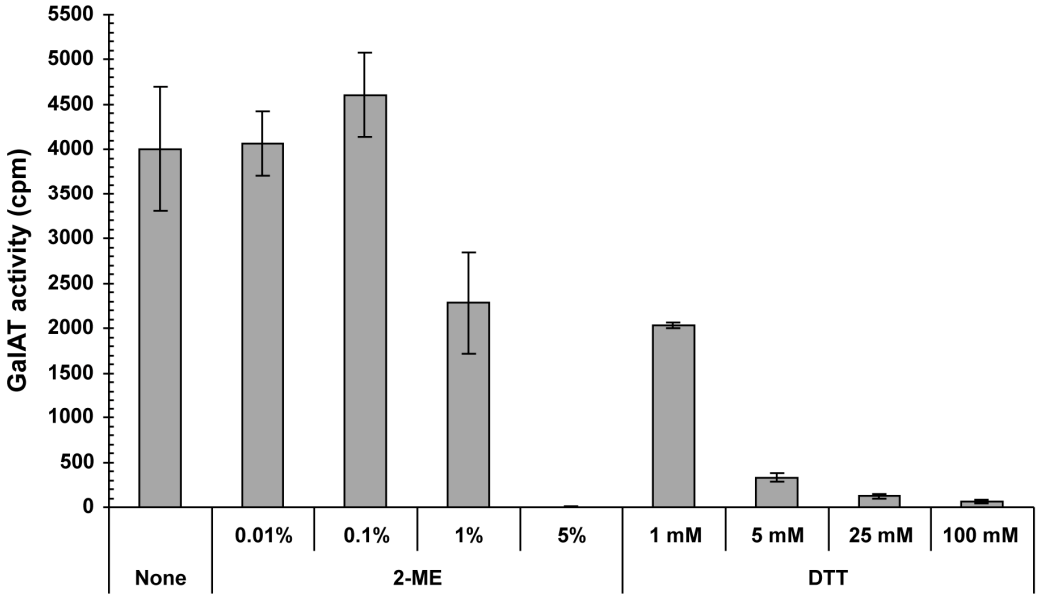


Figure D-2. Reducing agents have a negative effect on GalAT activity.

GalAT activity was assayed in a 60 µl reaction containing buffer, 1 µM OGA (DP 7-23), 5.75 µM UDP-[^{14}C]GalA, 30 µl of Arabidopsis SP fraction, and reducing agents as indicated. Reactions were incubated for 2 hrs at 30°C, and subsequently processed using the filter assay method (Doong et al., 1995; Sterling et al., 2005). 2-ME: β-mercaptoethanol; DTT: dithiothreitol. Data shown are the average of background-subtracted values ± standard deviation of duplicate reactions.

agreement with its use in the microsomal membrane preparation. However, 2-ME reduced GalAT activity by half at 1% (128 mM) and virtually abolished the activity at 5% (640 mM). In contrast, the effect of DTT was more severe, since at 1 mM it readily decreased the GalAT activity by approximately half and at 100 mM by ~99%.

On the one hand the experiment suggested the importance of disulfide bond(s) for the GalAT activity in the SP fraction and presumably for the GAUT1:GAUT7 GalAT complex. However, it is necessary to interpret this result with caution. It is not clear at this point if the reduction in GalAT activity was caused by disruption of disulfide bond(s) or due to the reduced condition of the GalAT reaction that may have influenced the catalysis. Moreover, disulfide bond(s) may exist both intramolecularly within GAUT1 and/or GAUT7 (each protein has 8 cysteine residues), and intermolecularly between the two proteins. If disruption of these disulfide bond(s) indeed resulted in the decreased GalAT activity, it is unclear if this was due to the disintegration of the GAUT1:GAUT7 complex, or simply because protein folding of the catalytic subunit of the complex was perturbed such that it was no longer enzymatically active. The former case may indicate that complex formation may be essential for the GalAT activity of the GAUT1:GAUT7 complex. Alternatively, GAUT7 may function solely to anchor the proteolytically-cleaved GAUT1 to the Golgi membrane for its Golgi retention (see Chapter 3 in this dissertation). Further experiments are needed to address these questions.

Effect of OGA acceptor substrate concentration on GalAT activity of Arabidopsis SP fraction

Solubilized GalAT requires the presence of exogenously-added acceptor substrate for its enzymatic activity. While it has been shown with solubilized tobacco enzyme that GalAT may prefer OGAs of DP ≥10 (Doong and Mohnen, 1998; Akita et al., 2002; Ishii, 2002), for most

experiments involving permeabilized or solubilized GalAT in our laboratory, an OGA mixture of DP 7-23 is usually applied in GalAT activity assays at 80 μg for a 30 μl reaction or at 90 μg for a 60 μl reaction, which equals 1003.03 μM and 564.21 μM, respectively, in the presence of ~5.75 – 7.66 μM UDP-[^{14}C]GalA. However, the kinetics of the Arabidopsis GalAT in the SP fraction, and of the GAUT1:GAUT7 complex (see Chapter 3), for acceptor substrates have not been determined.

As knowledge of an optimum acceptor substrate concentration would be an invaluable basis for many further studies, an experiment was carried out to test the effect of various concentrations of OGA on [^{14}C]GalA incorporation by Arabidopsis SP fraction. In this case, concentration in molarity (μM OGA) was chosen over mass per volume (μg OGA per reaction) to anticipate future experiments involving different acceptor substrates that may have different molecular weights (see next section). Concentrations of OGA mixture DP 7-23 at a range of 0 – 391 μM were tested in 60 μl GalAT reactions in the presence of 5.75 μM UDP-[^{14}C]GalA and 30 μl SP fraction. It was hypothesized that as the OGA concentration increases the GalAT activity would increase and eventually reach saturation. On the contrary, the experimental results showed that while GalAT activity did increase at OGA concentrations up to 1 μM, beyond 1 μM OGA the GalAT activity decreased considerably (Figure D-3). In a separate experiment, OGA at 0.78 – 391 μM was shown not to have a significant quenching effect on the scintillation measurement (data not shown), confirming that the trend illustrated in Figure D-3 is due to inhibition of GalAT activity and ruling out possible interference of scintillation counting by high amounts of OGA supplied in the reactions. The same trend was obtained regardless of the method used to process the GalAT reaction products, i.e. whether the precipitation method or the

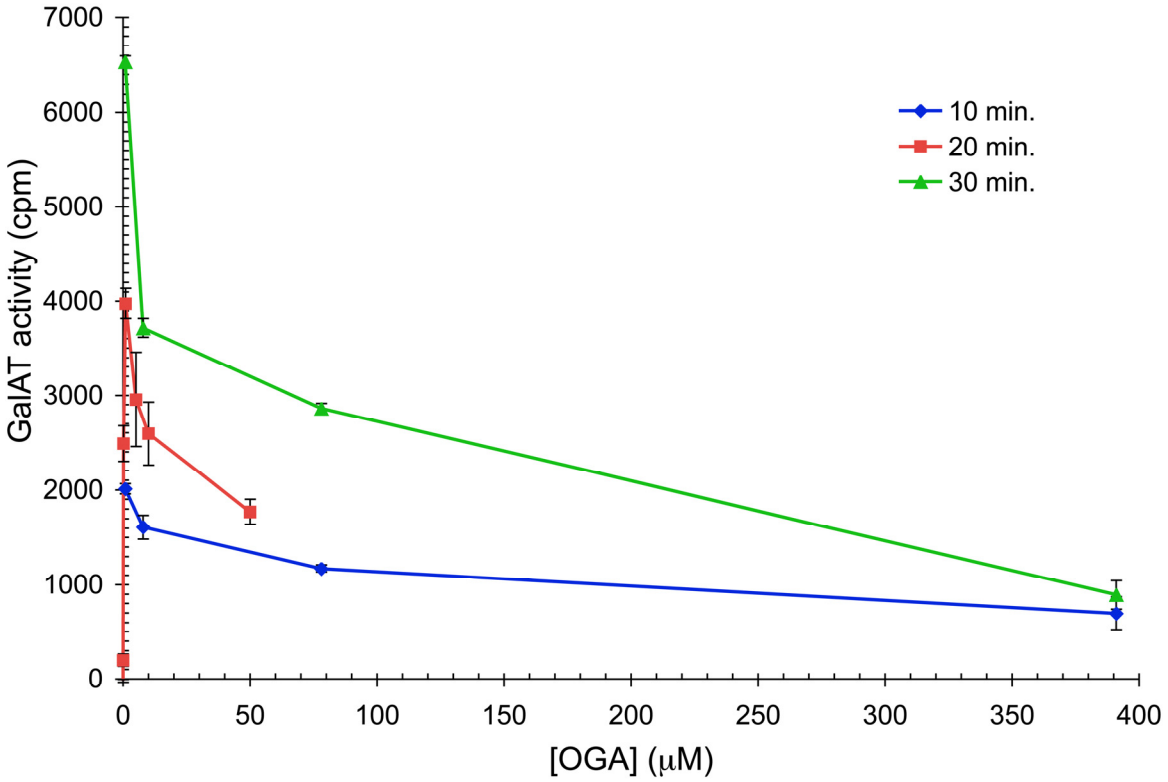


Figure D-3. High concentration of OGA acceptor substrate inhibits GalAT activity of Arabidopsis SP fraction.

GalAT activity was assayed in a 60 µl reaction containing buffer, 5.75 µM UDP-[^{14}C]GalA, 30 µl of Arabidopsis SP fraction, and OGA (DP 7-23) at a concentration range of 0 - 391 µM (0, 0.01, 0.1, 0.78, 1, 7.8, 10, 50, 78, 391 µM). Reactions were incubated for 10, 20, or 30 minutes at 30°C, and subsequently processed using the precipitation method (Doong et al., 1995). Data shown are the average of background-subtracted values ± standard deviation of duplicate reactions.

filter assay method was used (data not shown), although the filter assay method did appear to give overall lower scintillation signals than the precipitation method.

The above results clearly show that OGA at 1 μM OGA is optimum for the GalAT activity of SP fraction in the described reaction settings, and suggest that OGA at high concentrations is actually inhibitory to the GalAT activity of SP fraction. On the other hand, other preliminary data indicated that GalAT of permeabilized Arabidopsis microsomes gave higher activity at 1003.03 μM (80 μg in 30 μl reaction) than at 1 μM OGA (data not shown), suggesting that different enzyme samples may behave differently and that optimum reaction conditions need to be developed for each sample. Further research will be necessary to study the kinetics of OGA inhibition on GalAT activity, which may contribute to an understanding of the GalAT reaction mechanism.

Pectic acceptor substrate specificity of GalAT activity of Arabidopsis SP fraction

As many as six different GalA transferase activities, including three α-1,4-GalATs, have been hypothesized to be required to construct the different kinds of pectic polysaccharides (Caffall and Mohnen, 2009). Arabidopsis SP fraction has been shown to have HG:GalAT activity as it transfers GalA from UDP-GalA onto OGA acceptors in α-1,4-linkage (Sterling et al., 2006), however, it is not known if this GalAT activity is specific for only OGA acceptors or if it is also active on other types of pectic oligo/polysaccharides. We therefore tried to probe the substrate specificity of the GalAT activity in SP fraction by using different types of pectic oligo/polymers as acceptor substrates in the GalAT activity assays. Since the SP fraction showed high GalAT activity at very low molarity of OGAs (see Figure D-3), it was very important to pay extra attention to the purity of the different pectic oligomers/polymers and to

the way they were prepared. As will be described in more detail below, we discovered from our initial experiment that Arabidopsis SP fraction showed GalAT activity onto pectic RG-I oligomers and polymers, which were prepared using size exclusion chromatography, and which thus likely contained some contaminating HG. The GalAT activity observed with these RG-I substrates was therefore likely to be artifactual due to the presence of OGA or a stretch of GalA residues in these substrates, resulting in false positive results. As presented in Chapter 3 in this dissertation, subsequent experiments using a pure, HPAEC-purified RG-I mixture of DP 6-26 (see Appendix E) reproducibly showed that GalAT activity in the SP fraction and from the GAUT1:GAUT7 GalAT complex specifically uses OGAs, thus proving it to be HG:GalAT since it strongly prefers OGAs as acceptor substrates at the 0.1 and 1 μM concentrations tested.

In the initial experiment, pectic oligo/polymers of different types and from different sources were tested in GalAT activity assays as acceptor substrates in comparison to OGA: RG-I crude polymers from tobacco and sycamore, RG-I DP 10-14 (G-(R-G)$_n$) with 4,5-unsaturated GalA at the non-reducing ends, putative RG-I DP 12 ((R-G)$_6$ with unmodified non-reducing ends), RG-II dimers from tobacco and wine, and RG-II monomer from wine. These pectic substrates were obtained from several colleagues at Complex Carbohydrate Research Center, UGA. The purity of these substrates was therefore not ascertained. The estimated molecular weights of these acceptor substrates were calculated as follows: OGA DP 7-23 with average DP 15 – 2650.6 Da, tobacco and sycamore crude RG-I – 500 kDa, RG-I oligomers with average DP 12 – 1951.2 Da, RG-II dimer – 10,000 Da, RG-II monomer – 5,000 Da. Each substrate was supplied at 1 μM in a 60 μl GalAT reaction in the presence of 5.75 μM UDP-[14C]GalA and 30 μl of SP fraction. The reactions were incubated for 20 minute at 30°C.

Reaction results as presented in Figure D-4 indicated that at 1 μM substrate concentration, SP fraction gave the highest GalAT activity with OGA as substrate. In this experiment, GalAT also appeared to be able to use RG-I polymer or oligomers, provided that the non-reducing ends were not modified, with preference for oligomers (DP 12) rather than the polymer. No activity was observed with the 4,5-unsaturated RG-I oligomers, in agreement with the previous report with tobacco solubilized GalAT which was not active on OGA substrate containing 4,5-unsaturated GalA at the non-reducing end (Scheller et al., 1999). Very low amounts of activity were observed with RG-II oligomers, with slight preference for monomer rather than dimer forms. Further experiments testing higher concentrations of RG-II may confirm whether or not the GalAT activity in SP fraction can use RG-II as an acceptor substrate. At least for RG-I and RG-II, the different sources of the pectic substrates seemed to have no effect on the GalAT activity of SP fraction.

The GalAT activity observed with the above mentioned RG-I polymers and oligomers needs to be interpreted with caution considering the way these substrates were prepared. The RG-I polymers used in Figure D-4 were purified by size exclusion chromatography from cell wall fractions released by endopolygalacturonase (EndoPG) digestion. Since it is believed that the backbones of HG and RG-I are connected (O'Neill et al., 1996), the above EndoPG digestion may still leave stretches of GalA residues at the ends of the resulting RG-I backbones. The putative RG-I DP 12 oligomers were prepared by RG-hydrolase digestion of Arabidopsis seed mucilage followed by isolation of the DP 12 by size exclusion chromatography (Y. Kong, pers. comm.). Since size exclusion chromatography separates molecule based on size, it is likely not able to separate RG-I from OGA of the same DP. It is therefore possible that the preparation

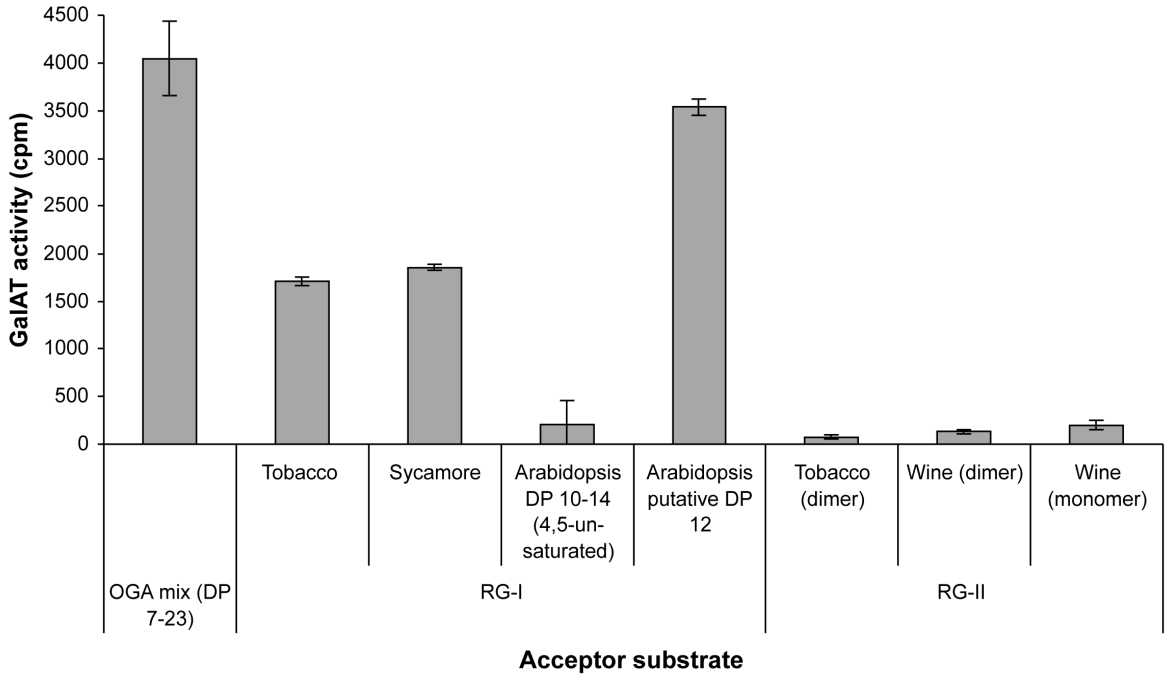


Figure D-4. Investigation of the acceptor substrate specificity of the GalAT in Arabidopsis SP fraction indicates possible OGA contamination in RG-I polymer/oligomer preparations.

GalAT activity was assayed in a 60 µl reaction containing buffer, 5.75 µM UDP-[^{14}C]GalA, 30 µl of Arabidopsis SP fraction, and 1 µM acceptor substrate as indicated. Reactions were incubated for 20 minutes at 30°C, and subsequently processed using the precipitation method (Doong et al., 1995). Data shown are the average of background-subtracted values ± standard deviation of duplicate reactions. All the pectic materials tested, except for OGA mix, were kind gifts from colleagues at CCRC, Athens, GA: tobacco RG-I and RG-II and sycamore RG-I were from Stefan Eberhard, RG-I of DP 10-14 (4,5-un-saturated) from Dr. Cheng Hua Deng, putative RG-I of DP 12 from Dr. Yingzhen Kong, wine RG-II (monomer and dimers) from Dr. Malcolm O'Neill.

used here was not 100% pure RG-I oligomers but also contained some DP-12-oligomers with stretches of GalA residues. As described in Appendix E, it is more appropriate to use HPAEC to prepare pure RG-I oligomers, since OGA oligomers carry a higher negative charge and elute at a later retention time in the HPAEC, compared to the RG-I oligomers of the same DP, thus allowing purification of the specific type of pectic oligomers.

Pectin methyltransferase activity of Arabidopsis SP fraction and GAUT1:GAUT7 GalAT complex

The identification of two putative pectin methyltransferases as candidate members of the GAUT1:GAUT7 GalAT holocomplex by proteomics analyses of anti-GAUT1 and anti-GAUT7 immunoabsorbed proteins as described in Chapter 2 suggested a close relationship between the synthesis and the methylesterification of HG. In light of this finding, the homogalacturonan methyltransferase (HG-MT) activity in SP fraction and of the GAUT1:GAUT7 complex was tested. Here several results from preliminary experiments are presented, which may be basis for further, more detailed study.

A methyltransferase (MT) assay was conducted essentially as described (Goubet and Mohnen, 1999) with slight modifications in the reaction compositions where necessary (Figure D-5, D-6, and D-7). Incubation of Arabidopsis SP fraction with [^{14}C]methyl-*S*-adenosyl-L-methionine ([^{14}C]methyl-SAM) yielded radioactivity incorporation into OGA and PGA substrates (Figure D-5A). This reflects the total MT activity present in the enzyme sample. The HG-MT activity was then assessed by treating the methylesterified products from the MT reactions with purified pectin methylesterase (PME), which removes the [^{14}C]methyl groups incorporated onto the OGA or polygalacturonic acid (PGA) substrate (Figure D-5B). A

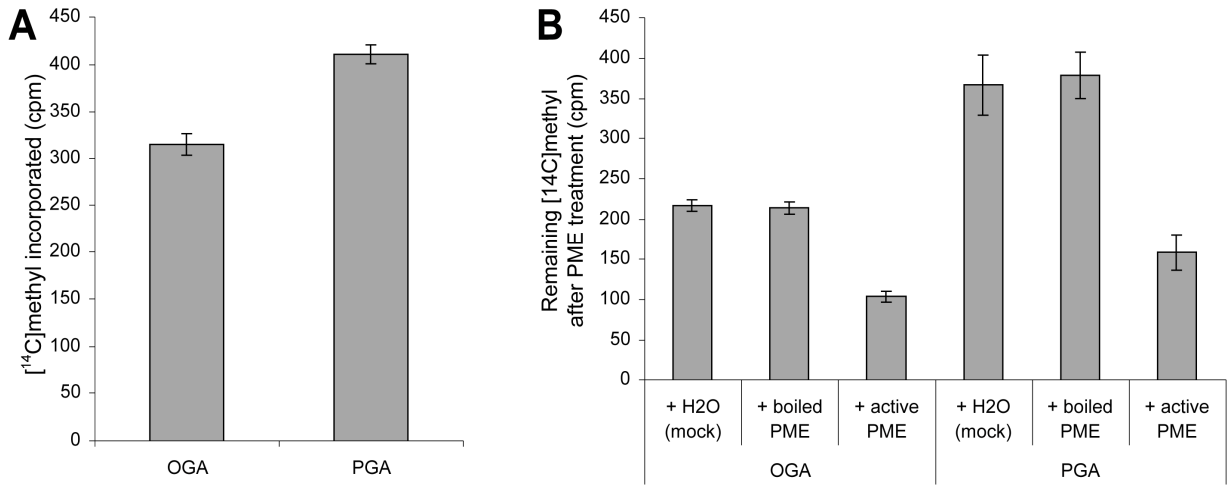


Figure D-5. Arabidopsis SP fraction has low levels of HG-MT activity.

MT reactions were carried out in 50 µl volume containing buffer, 25 µg DL-phosphatidylcholine, 50 µg OGA (DP 7-23) or PGA, 25 µM [^{14}C]methyl-*S*-adenosyl-L-methionine ([^{14}C]methyl-SAM), and 25 µl SP fraction, with incubation at 30°C for 4 hrs. Methylesterified products were precipitated and washed to remove unincorporated [^{14}C]methyl-SAM as described (Goubet and Mohnen, 1999). For treatment with PME (from *Aspergillus niger*, 2.2 µg/µl, 1 U/µg, 1 U = 1 µmol/min, a kind gift from Carl Bergmann, CCRC, Athens, GA), washed pellets were resuspended in 200 µl of 10 mM ammonium formate pH 4.5 and incubated with 0.5 µl of H_2O (mock), boiled PME, or active PME at room temperature overnight. PME treatment was terminated by addition of 250 µl 20% TCA, and the resulting pellet was washed once with 300 µl of 2% TCA. Washed pellets were resuspended in 200 µl H_2O and subjected to scintillation measurement. Data shown are the average of background-subtracted values ± standard deviation of duplicate reactions.

difference in the measured radioactivity between controls (H_2O or boiled enzyme) and sample treated with active PME is regarded as the amount of HG-MT activity, while the PME-resistant portion represents other methyltransferase activities (e.g. methyl ether synthesis) that may be present in the enzyme sample. In this experiment, Arabidopsis SP fraction showed low amounts of HG-MT activity, with average net HG-MT activity of 111.9 and 214.6 cpm per reaction for OGA and PGA substrates, respectively (Figure D-5B). This result indicates substrate preference for PGA rather than OGA, in agreement with the results reported previously with solubilized HG-MT from tobacco (Goubet and Mohnen, 1999). UDP-GalA has been shown to stimulate the activity of tobacco membrane-bound HG-MT, presumably because it allows membrane-bound HG:GalAT to produce more HG that in turn becomes substrate for the HG-MT activity (Goubet et al., 1998). However, Arabidopsis SP fraction MT activity did not increase when the reactions were supplemented with up to 100 μM UDP-GalA (Figure D-6), which is again similar to the response of tobacco solubilized HG-MT (Goubet and Mohnen, 1999).

To see whether the GAUT1:GAUT7 complex also has HG-MT activity, proteins immunoprecipitated by anti-GAUT1 and anti-GAUT7 antibodies were subjected to the HG-MT assay (Figure D-7). Arabidopsis SP fraction was used as the positive control, while proteins immunoprecipitated using preimmune sera were included as the negative controls. The methylesterified products yielded by all immunoprecipitated protein samples were resistant to PME, indicating the absence of HG-MT activity in any of these samples. This suggests that the GAUT1:GAUT7 GalAT complex does not seem to carry HG-MT activity despite the discovered association between putative HG-MTs and the holocomplex. However, as reflected by their NSAF values in the proteomics analysis, the identified putative HG-MTs were much lower in abundance compared to GAUT1 and GAUT7 in the immunoprecipitated proteins. Therefore the

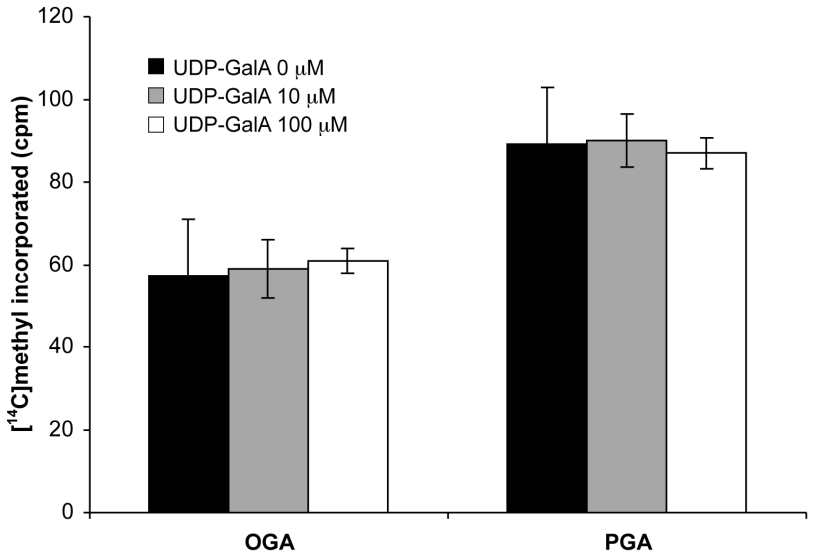


Figure D-6. UDP-GalA does not stimulate MT activity of Arabidopsis SP fraction.

MT reactions were carried out in a 50 μl volume containing buffer, 25 μg DL-phosphatidylcholine, 50 μg OGA (DP 7-23) or PGA, 8 μM [^{14}C]methyl-*S*-adenosyl-L-methionine ([^{14}C]methyl-SAM), 12 μM SAM, 25 μl SP fraction, and UDP-GalA as indicated. Reactions were incubated at 30°C for 4 hrs. Methylesterified products were subsequently processed as described (Goubet and Mohnen, 1999). Data are average of background-subtracted values ± standard deviation of duplicate reactions.

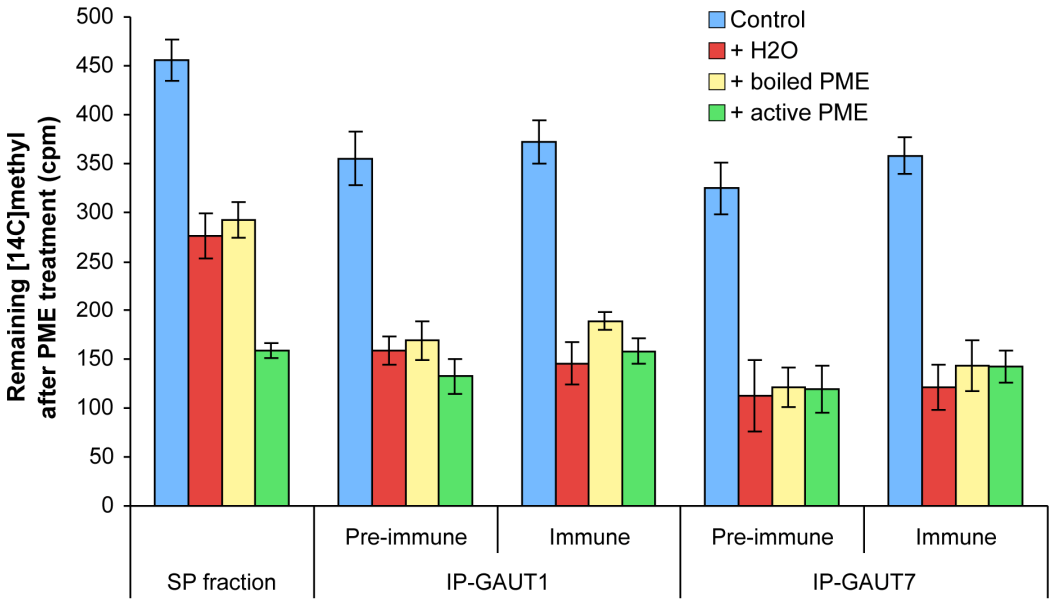


Figure D-7. Anti-GAUT1 and anti-GAUT7 immunoabsorbed proteins do not show HG-MT activity.

Immunoprecipitations of proteins from Arabidopsis SP fraction were carried out using anti-GAUT1 and anti-GAUT7 preimmune and immune antibodies conjugated to magnetic beads as previously described (Sterling et al., 2006).

MT reactions were carried out in a 50 µl volume containing buffer, 25 µg DL-phosphatidylcholine, 50 µg PGA, 25 µM [^{14}C]methyl-*S*-adenosyl-L-methionine ([^{14}C]methyl-SAM), and 25 µl enzyme sample. Reactions were incubated at 30°C for 4 hrs, and reaction products were subsequently processed as described (Goubet and Mohnen, 1999). Subsets of the washed, pelleted products were also subjected to PME treatment as described (see Figure D-5). Data are average of background-subtracted values ± standard deviation of triplicate reactions.

very low level of HG-MT activity in the GAUT1:GAUT7 complex may be due to the low levels of protein. Development of an optimized, more robust HG-MT assay is essential to allow detection of low amounts of HG:MT activity and to facilitate comparisons between samples with subtle differences in activity.

APPENDIX E

PURIFICATION OF RHAMNOGALACTURONAN-I (RG-I) OLIGOSACCHARIDES FROM ARABIDOPSIS SEED MUCILAGE

To study the pectic acceptor substrate specificity of the GalAT activity of Arabidopsis SP fraction and of the GAUT1:GAUT7 GalAT complex, it was essential to obtain and use purified RG-I oligomers of a DP comparable to the DP of the oligogalacturonides (OGA) acceptors, i.e. DP 7-23. It has been shown with GalAT preparations from several species that detergent-solubilized GalAT prefers OGA of DP ≥10 OGA for maximum activity (Doong and Mohnen, 1998; Akita et al., 2002; Ishii, 2002). An OGA mixture of DP 7-23 has routinely been used as acceptor substrate in GalAT activity assays in our laboratory. Data from experiments using RG-I polymer and oligomer preparations, as shown in Appendix D, presented possible false positive results in GalAT activity assays since these substrates may have contained not only RG-I, but also regions of OGA that can serve as an acceptor substrate for HG:GalAT. The much larger size of the RG-I polymer also renders it an inappropriate substrate when compared to OGAs of DP 7-23. Furthermore, the alternating Rha-GalA backbone of RG-I makes it necessary to test RG-I oligomers that have two different non-reducing ends: those with a Rha residue at the non-reducing end ($(R\text{-}G)_n$, here named RG-I-R), and those with a GalA residue at the non-reducing end ($(G\text{-}R)_n$, here named RG-I-G). A purification scheme was thus used to isolate these two

types of RG-I oligomers of DP 6-26 from Arabidopsis seed mucilage in sufficient quantity for the substrate specificity studies.

Preparation of RG-I crude polymer from Arabidopsis seed mucilage

Arabidopsis seed mucilage consists largely of un-branched, unsubstituted RG-I (Western et al., 2004; Macquet et al., 2007), making it an excellent starting material from which to obtain RG-I backbone oligomers. Seed mucilage was extracted from Arabidopsis wild-type seeds as described in Caffall (2008) with modifications to facilitate scaling-up. Arabidopsis seeds were incubated with warm water (~50°C) with stirring for 1 hour (100 ml nanopure water/ml seeds; ~0.7 g/ml seeds). This step was repeated five more times. The supernatants containing extracted mucilage from the six repeated extractions were pooled, filtered through two layers of Miracloth (Calbiochem, San Diego, CA), concentrated on a rotary evaporator to reduce volume, and lyophilized to dryness. The dry mucilage was weighed (~700 mg), reconstituted in 40 ml of 10 mM ammonium formate pH 4.5, and treated overnight at room temperature with stirring with 6 U of endopolygalacturonase (EndoPG; from *Aspergillus niger*, 1.2 U/μl, 0.5 μg/μl, 1 U = 1 μmol/min; a kind gift from Carl Bergmann, CCRC, Athens, GA). The treated solution was dialyzed against water (Spectra/Por 6; 1000 MWCO dialysis tubing) to remove digested HG domains from the RG-I polymer preparation, filtered through 0.45 μm nylon filters, and lyophilized. The resulting EndoPG-treated mucilage is referred to as the RG-I crude polymer.

The RG-I crude polymer preparation was monitored by analytical HPAEC-PAD (pulsed amperometric detection), using conditions and settings as summarized in Table E-1. The HPAEC chromatograms showed that the EndoPG digestion did not significantly change the peak profiles of the preparation (Figure E-1, A and B), and that there was no observable difference in

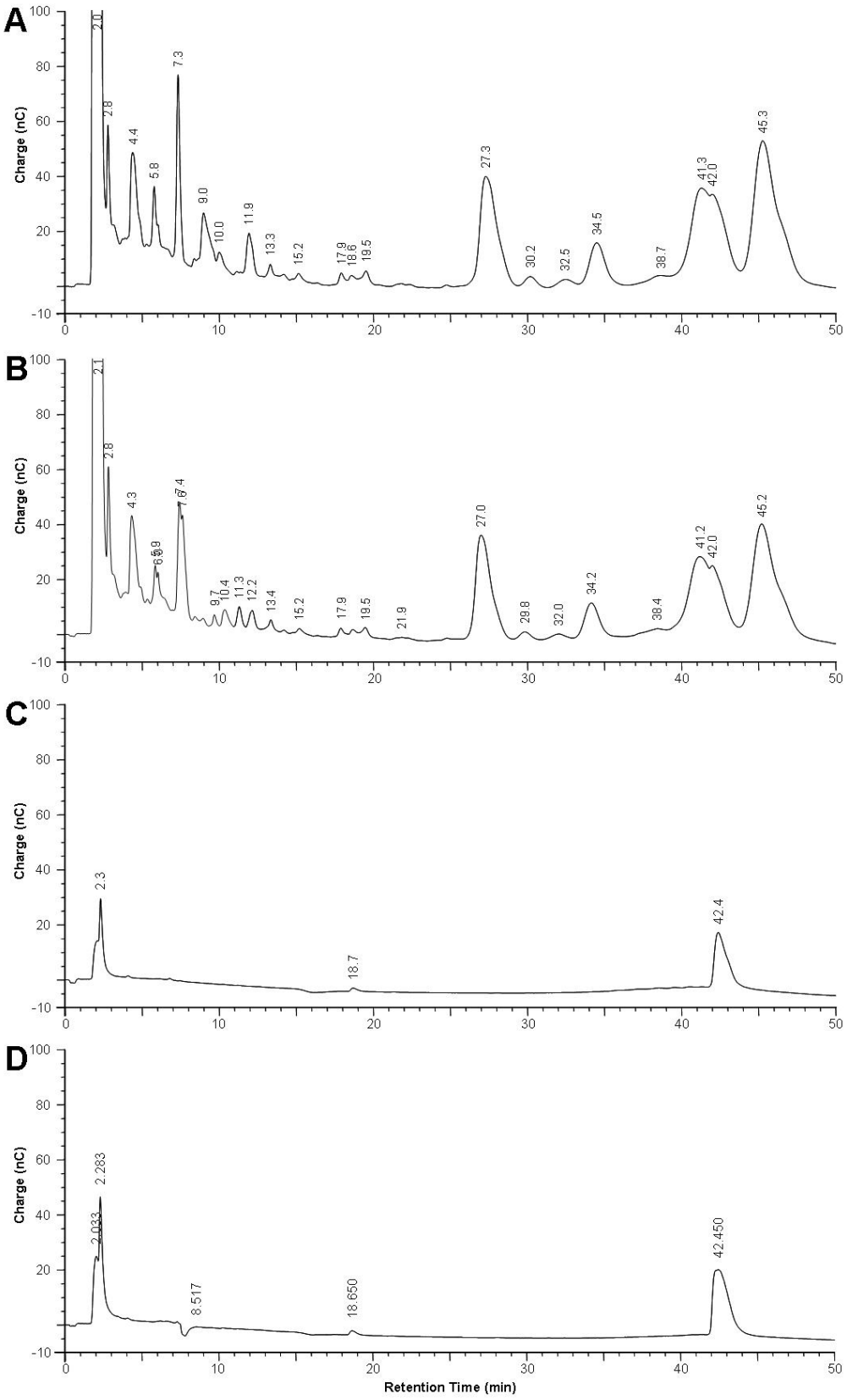
the chromatograms of samples taken after dialysis versus after filtration (Figure E-1, C and D). However, there was a dramatic change in the peak profiles before and after the dialysis step (Figure E-1, compare B and C) that seemingly removed not only salt and digested HG domains, but also material other than the main peak at 42.4 minutes, which is presumably the target RG-I polymer. For example, a major peak eluting at ~27 min., which was shown in preliminary trials to be resistant to the RG-hydrolase digestion and acid hydrolysis (see next section) and therefore interfered with downstream purification of the resulting RG-I oligomers (data not shown), was removed by this dialysis step.

Assuming that the preparation contains solely un-branched RG-I polymer, the amount of the RG-I polymer in the preparation was calculated based on total sugar (corresponds to Rha and GalA) and uronic acid (correspond to GalA) contents, which were measured by phenol sulfuric acid (DuBois et al., 1956) and uronic acid (*m*-hydroxybiphenyl; (Blumenkrantz and Asboe-Hansen, 1973) assays, respectively. The phenol sulfuric acid assay works rather poorly for acidic sugars, therefore two standard curves (for Rha and for GalA, respectively) were generated in this assay. The response signal attributed to GalA, calculated from the total GalA content from the uronic acid assay result, was subtracted from the total response signal of the phenol sulfuric acid assay of the RG-I sample to obtain the net response signal for Rha. This net value was then converted to the amount of Rha, which was subsequently combined with the amount of GalA to give the total amount of RG-I polymer in the sample. From two preparations using as starting material 4.2 and 12.6 grams of Arabidopsis wild-type seeds, the total amount of mucilage (i.e. RG-I crude polymer prior to EndoPG treatment) was quantified to be 403.2 and 1150.1 mg in total sugar content, respectively, while the final yields of the RG-I crude polymer were 108.0 and 282.5 mg in total sugar content, respectively. This corresponds to a ~25%

Table E-1. Conditions and settings for analytical and semi-preparative HPAEC-PAD on a Dionex system (Dionex Corp., Sunnyvale, CA).

	Analytical	**Semi-preparative**
Column(s)	CarboPac PA-1 (4 x 250 mm) with CarboPac Guard column (3 x 25 mm)	CarboPac PA-1 (9 x 250 mm)
Flow rate	1 ml / minute	2 ml / minute
Buffer system	Eluant A: H_2O Eluant B: 1 M sodium acetate buffer	
Gradient scheme	95% A, 5% B at 0 - 2 min.; 60% A, 40% B at 15 min.; 20% A, 80% B at 47 min.; 0% A, 100% B at 50 – 55 min.; 95% A, 5% B at 56 – 66 min.	
Post-column base	400 mM NaOH	1 M NaOH

Figure E-1. HPAEC-PAD chromatograms of RG-I crude polymer preparation from Arabidopsis seed mucilage separated over a CarboPac PA-1 column: (A) before and (B) after EndoPG treatment, (C) after dialysis, and (D) after filtration using 0.45 μm nylon filter. Numbers above the peaks indicate the peak retention times in minutes.


A
B
C
D
Charge (nC)
Retention Time (min)
2.0
2.8
4.4
5.8
7.3
9.0
10.0
11.9
13.3
15.2
17.9
18.6
19.5
27.3
30.2
32.5
34.5
38.7
41.3
42.0
45.3
2.1
2.8
4.3
6.9
7.4
9.7
10.4
11.3
12.2
13.4
15.2
17.9
19.5
21.9
27.0
29.8
32.0
34.2
38.4
41.2
42.0
45.2
2.3
18.7
42.4
2.283
2.033
8.517
18.650
42.450

average final yield. The ratios between GalA and sugar other than GalA, were approximately 1 : 2.75 and 1 : 2 in the mucilage and in the RG-I crude polymer, respectively.

Optimization of RG-hydrolase and acid hydrolysis digestions of RG-I crude polymer

Production of RG-I-R and RG-I-G oligomers from the RG-I crude polymer was carried out by digestion using RG-hydrolase enzyme and controlled acid hydrolysis, respectively, based on protocols as described (Mutter et al., 1998) with modifications. RG-hydrolase (RG α-D-galactopyranosyluronide-(1,2)-α-L-rhamnopyranosyl hydrolase; previously named rhamnogalacturonase (RGase)) was first isolated from *Aspergillus aculeatus* and shown to cleave the α-1,2-linkages between the GalA and Rha residues in RG-I backbone (Schols et al., 1990; Mutter et al., 1998). Controlled acid hydrolysis takes advantage of the different sensitivity of glycosidic linkages to acidic conditions: in the RG-I backbone the linkage →4)-α-D-GalA*p*-(1→2)-α-L-Rha*p*-(1→ is more stable than the linkage →2)-α-L-Rha*p*-(1→4)-α-D-GalA*p*-(1→ at pH<2, thus resulting in preferential cleavage between Rha and GalA (Renard et al., 1997). The use of both methods were previously described on wall preparations containing more complex RG-I, such as de-esterified sugar-beet pulp and modified hairy regions (MHR) from apple (Mutter et al., 1998). The protocols, thus, required optimization to be used on the relatively simpler Arabidopsis RG-I crude polymer in this experiment. The goal was to devise reaction conditions that would yield maximum amounts of RG-I oligomers of DP 6-26.

The RG-hydrolase digestion reaction was optimized into a 200 µl reaction containing 3 mg RG-I crude polymer and 0.256 µg of RG-hydrolase (from *Aspergillus aculeatus*, 2.05 mg/ml; a kind gift from Henk Schols, Wageningen University) in 50 mM sodium acetate buffer, pH 5.4. The reaction was incubated for 48 hrs at 40°C and terminated by boiling for 5 minutes. Reaction

products were loaded onto a 2 ml column of Dowex 50W X8 (Sigma, St.Louis, MO) pre-washed with sterile nanopure H_2O, from which the flow-through and 6 ml H_2O elution were collected and lyophilized to dryness.

The acid hydrolysis procedure was optimized into a 500 µl reaction containing 5 mg of RG-I crude polymer in 0.1 M HCl, which was incubated for 4 hrs at 80°C. The hydrolysis was terminated by addition of 100 µl 1 M NaOH. Reaction products were filtered over a Dowex column (see above) or dialyzed against water (Spectra/Por CE; 100-500 MWCO) to remove salt, and then lyophilized to dryness.

The lyophilized products from the above RG-hydrolase digestion and the acid hydrolysis reactions were dissolved, independently, in 500 µl sterile nanopure H_2O, and a 10% aliquot of each was separated on analytical HPAEC-PAD to evaluate the digestion profiles (Figure E-2). A standard mixture of authentic mono-, di-, and tri-GalA (Sigma, St. Louis, MO), and OGA of DP 7-23 were also separated by HPAEC and used as a comparison for the RG-I oligomers. The digested/hydrolyzed RG-I polymers showed peak patterns characteristic of a homologous series of oligomers (Figure E-2, B and C) that eluted at earlier retention times than the standard OGA mixture (Figure E-2A), which was as expected due to the higher net negative charges of the latter HG oligomers. Several of these RG-I oligomer peaks were collected, desalted, lyophilized, re-dissolved in sterile nanopure H_2O, and subjected to MALDI-TOF analysis to determine their masses. By comparing the expected masses of RG-I oligomers and the masses observed by MALDI-TOF (Table E-2), the HPAEC-PAD peaks of the RG-I oligomers were annotated (Figure E-2, B and C).

The optimized RG-hydrolase digestion and acid hydrolysis reactions yielded oligomer series ranging from tetramer up to observable ~60-mer, with the distances between two

Figure E-2. Representative HPAEC-PAD chromatograms of RG-I oligomers obtained from optimized RG-hydrolase digestion and controlled acid hydrolysis.

(A) Standard mixture of 5 μg each of mono-, di-, and tri-GalA, and 200 μg of OGA enriched for DP 7-23.

(B) RG-I oligomers resulting from RG-hydrolase digestion.

(C) RG-I oligomers resulting from controlled acid hydrolysis.

The numbers above the peaks indicate the peak retention times in minutes, followed by the peak identities, if known (e.g. $GalA_6$ = OGA of DP 6; $(Rha\text{-}GalA)_2$ = Rha-GalA-Rha-GalA with the first Rha representing the non-reducing end of the oligomer). Blue lines represent the concentration gradient of water in comparison to the eluent (1 M sodium acetate buffer).

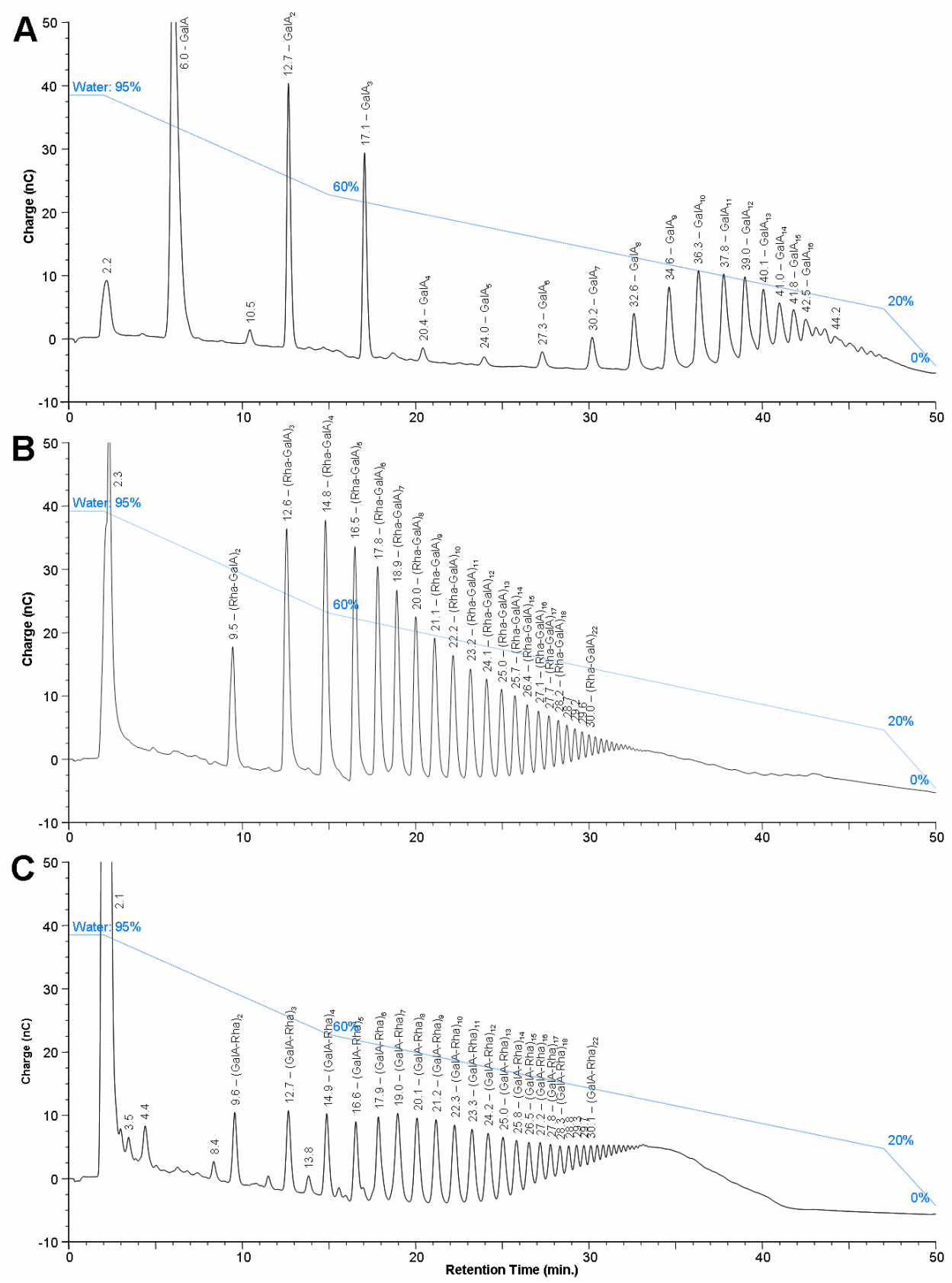

A
Water: 95%
60%
20%
0%
Charge (nC)
2.2
6.0 - GalA
10.5
12.7 – GalA2
17.1 – GalA3
20.4 – GalA4
24.0 – GalA5
27.3 – GalA6
30.2 – GalA7
32.6 – GalA8
34.6 – GalA9
36.3 – GalA10
37.8 – GalA11
39.0 – GalA12
40.1 – GalA13
41.0 – GalA14
41.8 – GalA15
42.5 – GalA16
44.2
B
Water: 95%
60%
20%
0%
Charge (nC)
2.3
9.5 – (Rha-GalA)2
12.6 – (Rha-GalA)3
14.8 – (Rha-GalA)4
16.5 – (Rha-GalA)5
17.8 – (Rha-GalA)6
18.9 – (Rha-GalA)7
20.0 – (Rha-GalA)8
21.1 – (Rha-GalA)9
22.2 – (Rha-GalA)10
23.2 – (Rha-GalA)11
24.1 – (Rha-GalA)12
25.0 – (Rha-GalA)13
25.7 – (Rha-GalA)14
26.4 – (Rha-GalA)15
27.1 – (Rha-GalA)16
27.7 – (Rha-GalA)17
28.2 – (Rha-GalA)18
29.2
29.6
30.0 – (Rha-GalA)22
C
Water: 95%
60%
20%
0%
Charge (nC)
2.1
3.5
4.4
8.4
9.6 – (GalA-Rha)2
12.7 – (GalA-Rha)3
13.8
14.9 – (GalA-Rha)4
16.6 – (GalA-Rha)5
17.9 – (GalA-Rha)6
19.0 – (GalA-Rha)7
20.1 – (GalA-Rha)8
21.2 – (GalA-Rha)9
22.3 – (GalA-Rha)10
23.3 – (GalA-Rha)11
24.2 – (GalA-Rha)12
25.0 – (GalA-Rha)13
25.8 – (GalA-Rha)14
26.5 – (GalA-Rha)15
27.2 – (GalA-Rha)16
27.8 – (GalA-Rha)17
28.3 – (GalA-Rha)18
29.3
29.8
30.1 – (GalA-Rha)22
Retention Time (min.)

Table E-2. Annotation of HPAEC-PAD peaks of the RG-I oligomers by MALDI-TOF analysis. RG-I-R and RG-I-G oligomers are represented by the structures $(Rha\text{-}GalA)_n$ and $(GalA\text{-}Rha)_n$, respectively. MALDI-TOF mass spectrometry analysis was performed using an Applied Biosystem 4700 Proteomics Analyzer (Applied Biosystem, Foster City, CA). Multiple masses detected by MALDI-TOF that are higher than the theoretical mass correspond to the sodium adducts of the respective RG-I oligomers. n/d = not detected.

RG-I oligomer		HPAEC retention time (min.)		Theoretical mass (Da)	Experimental mass(es) (MALDI-TOF m/z)	
DP	Structure	Analytical	Semi-preparative		Negative mode	Positive mode
4-mer	$(Rha\text{-}GalA)_2$	9.5	11.2	662.6	685.2	685.1; 707.1; 729.1
	$(GalA\text{-}Rha)_2$	9.6	11.3		662.3; 684.3	685.2; 707.1
6-mer	$(Rha\text{-}GalA)_3$	12.6	14.2	984.9	986.7; 1030.5	1007.2; 1051.2; 1073.2
	$(GalA\text{-}Rha)_3$	12.7	14.2		984.9	1007.2; 1029.2
8-mer	$(Rha\text{-}GalA)_4$	14.8	16.2	1307.2	1307.9; 1329.8; 1351.8; 1389.7	1329.4; 1351.2
	$(GalA\text{-}Rha)_4$	14.9	16.2		1307.9; 1329.8	1329.4; 1351.2
10-mer	$(Rha\text{-}GalA)_5$	16.5	17.8	1629.5	1629.2; 1651.1; 1673.1	1673.5
	$(GalA\text{-}Rha)_5$	16.6	17.8		1629.3; 1651.1; 1673.0	1673.4; 1695.4
12-mer	$(Rha\text{-}GalA)_6$	17.8	18.9	1951.8	1950.4; 1972.4; 1994.3	n/d
	$(GalA\text{-}Rha)_6$	17.9	18.9		1950.4; 1994.2	n/d
14-mer	$(Rha\text{-}GalA)_7$	18.9	20.0	2274.1	n/d	n/d
	$(GalA\text{-}Rha)_7$	19.0	19.9		2271.6; 2315.5	n/d

consecutive oligomer peaks equivalent to the dimer Rha-GalA or GalA-Rha, respectively (Figure E-2). A major peak at 2.1-2.2 min. likely corresponded to neutral sugar monomers, presumably Rha monomer as a by-product of the digestion/hydrolysis reactions. Acid hydrolysis seemed to give smaller amounts of oligomers DP 4-30 (retention times between 9 and 30 minutes) and higher amounts of large, partially digested polymer (retention times between 30 and 40 minutes), as inferred by the peak sizes, compared to products generated from the RG-hydrolase digestion. Acid hydrolysis also resulted in more additional, un-identified minor peaks (e.g. at retention times 3.5, 4.5, 8.4, and 13.8 minutes) compared to the RG-hydrolase digestion.

Scaled-up isolation of RG-I-R and RG-I-G oligomers of DP 6-26

The optimized protocols for RG-hydrolase digestion and controlled acid hydrolysis resulted in a majority of the product as RG-I oligomers of DP 4-30, from which DP 6-26 could be isolated for use as acceptor substrate for comparison in reactions to OGA (DP 7-23) acceptors. Multiple reactions of both RG-hydrolase digestion and acid hydrolysis were carried out, which were then pooled to make the scaled-up preparations of the RG-I-R and RG-I-G, respectively. After dialysis and lyophilization, the pooled preparations were resolved by semi-preparative HPAEC using conditions and settings as summarized in Table E-1.

A small aliquot of each preparation was first separated by HPAEC with post-column base addition and monitored by PAD. The peak patterns were obtained and used to guide the collection of the target peaks of DP 6-26 in the subsequent HPAEC separations of the remaining bulk of RG preparation. Initially, these runs were also monitored by PAD with post-column addition of base, with the target peak fractions collected in tubes containing 1 M HCl to immediately neutralize the solutions. The neutralized target peak fractions were pooled, desalted

by dialysis, and lyophilized. However, there was a concern that the high pH condition, due to the post-column addition of base, may have caused base-catalyzed modification of the molecular structure of the purified RG-I oligomers, despite the very short exposure time to the base (less than 5 minutes from post-column base addition to the neutralization by HCl in the receiving tubes). To avoid any unwanted modifications of the RG-I oligomers that could compromise their further use, all subsequent collections of target peaks were made from HPAEC separations performed without post-column base addition. Collected target peak fractions were pooled, dialyzed, and lyophilized. The dried DP 6-26 RG-I-R and RG-I-G preparations were dissolved in sterile, nanopure H_2O, and analyzed by phenol sulfuric acid and uronic acid assays to determine the RG-I content. The final yields of DP 6-26 RG-I-R and RG-I-G oligomers, were 10.9 and 4.95 mg, respectively, which were prepared from 48 and 85 mg, respectively, of starting RG-I crude polymer. These numbers clearly indicate that the production of RG-I-R oligomers is more efficient than the RG-I-G oligomers. In addition, the partially digested/hydrolyzed RG-I polymer could potentially be recycled as starting material for the next digestion/hydrolysis reactions, with some optimization of the procedure.

The procedures described above serve as a basis for further improvements to optimize the purification of RG-I-R and RG-I-G oligomers. While the protocols for RG-hydrolase digestion and acid hydrolysis reactions are more or less optimized, the current procedure to isolate the resulting RG-I-R and RG-I-G oligomers is time-consuming and inefficient, especially if there is a future need to isolate discreet sizes of the oligomers (e.g. DP 10, DP 12, and so forth). For example, the HPAEC-PAD system requires the use of sodium acetate eluant, which necessitates a desalting step, usually by dialysis, prior to lyophilization. Development of a HPAEC scheme using lyophilizable eluant, such as ammonium formate, and with a compatible detection method

such as evaporative light scattering detection (ELSD), would allow the collected fractions to be directly lyophilized. Further scaling-up of the HPAEC using a preparative column would help reduce the need to perform numerous chromatography runs to obtain a sufficient quantity of the target oligomers. As an alternative to chromatography, a selective EtOH precipitation may also be used to prepare a mixture of RG-I oligomers with a certain range of DP, similar to the method developed previously for OGA (Spiro et al., 1993). Such a method is currently being tested in our laboratory.

APPENDIX F

EFFORTS TO PROBE THE BIOLOGICAL FUNCTIONS OF GAUT1 AND GAUT7 BY PROMOTER:*GUS* EXPRESSION AND MUTANT ANALYSES

The biochemical characterization of GAUT1 and GAUT7, as described in previous publications (Sterling, 2004; Sterling et al., 2006) and in this dissertation, has provided valuable information on the enzymatic GalAT function of GAUT1 *in vitro* and demonstrated a protein-protein interaction and complex formation between GAUT1 and GAUT7 *in vitro* and *in vivo*. However, to fully understand the function of GAUT1 and GAUT7 in pectin synthesis *in planta,* it is important to study the biological significance of these two proteins. To address this issue, we have attempted to elucidate the biological functions of GAUT1 and GAUT7, i.e. by assessing the transcriptional expression of *GAUT1* and *GAUT7* through promoter:*GUS* expression analysis, and by examining T-DNA mutant lines available for these genes. Results and current progress from these efforts are herein described.

Promoter:GUS transcriptional expression analysis of Arabidopsis GAUT1 and GAUT7

The promoter-β-glucuronidase (GUS) reporter gene system is a valuable tool to study gene transcript expression *in planta*. This technique makes use of the *E. coli uidA* (*gusA*) gene encoding β-glucuronidase, which catalyzes the hydrolysis of glucuronides (Jefferson et al., 1987). The *GUS* gene driven by the promoter of a gene of interest can be introduced into plants,

so that its expression in transgenic plants will reflect the specific transcript expression of the gene of interest. The presence of the GUS enzyme in plant tissues/cells can be assessed by its ability to cleave the histochemical substrate 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc), which results in a visible blue color in the cells in which it is expressed.

The *GAUT1* and *GAUT7* promoter:*GUS* expression analyses were expected to show whether the expression of *GAUT1* and *GAUT7* was constitutive or regulated spatially or temporally. Since GAUT1 and GAUT7 have been shown to interact in a GalAT complex (see Chapter 3 in this dissertation), we hypothesized that their encoding genes should be co-expressed. This study was also expected to provide a more detailed expression pattern than that available from transcript data such as RT-PCR studies (Caffall et al., 2009), massive parallel signature sequencing (*Arabidopsis* MPSS database at http://mpss.udel.edu/at/), or microarray databases such as Genevestigator (Zimmermann et al., 2004). The results from the promoter:*GUS* reporter gene study would serve as a valuable resource to guide future studies of altered phenotypes in mutant and RNA interference (RNAi) plants.

The genomic sequences comprising the intergenic region and 5'-untranslated region (5'-UTR) upstream of the start codons of *GAUT1* and *GAUT7* (1932 bp and 868 bp, respectively), were considered as the promoter regions of the genes (Figure F-1). These regions were cloned (PCR primers are listed in Table F-1) and, upon sequence verification, inserted upstream of the *GUS* coding sequence in vector pBI101. The resulting constructs were introduced into *Agrobacterium tumefaciens* strain GV3101::pMP90, and the transformed Agrobacteria were subsequently used to transform wild-type Arabidopsis plants using the floral-dip method (Clough and Bent, 1998). Histochemical analyses were conducted on the T2 progenies of multiple lines

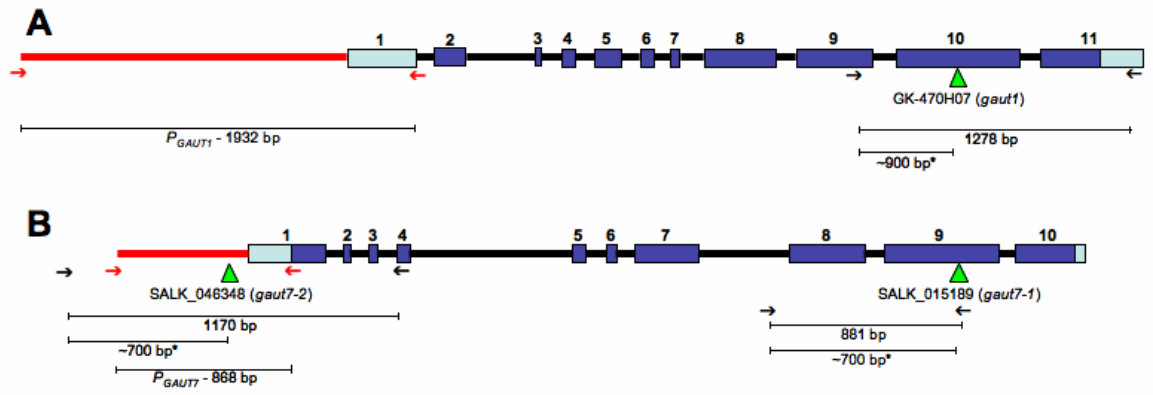


Figure F-1. The gene structures of (A) *GAUT1* and (B) *GAUT7*.

Numbered boxes represent exons, with light and dark blue colors representing un-translated regions and coding regions, respectively. Introns are shown as black thick lines, and intergenic regions as red thick lines. Positions of T-DNA insertions of mutant lines are indicated by green triangles. Arrows indicate positions of PCR primers used to amplify the promoter regions (red arrows) or to genotype the mutant lines (black arrows). The resulting PCR products are shown. (*) - the product sizes of PCR reactions using the gene specific and T-DNA specific primer pairs are based on estimation from gel electrophoresis results and not drawn to scale.

Table F-1. List of PCR primers used to clone *GAUT1* and *GAUT7* promoter regions for the promoter:*GUS* expression analysis and to perform genotyping in mutant studies.

The restriction sites for sub-cloning are underlined. F – Forward; R – Reverse.

Promoter:GUS Constructs			
Gene ID	**Primer Name**	**Primer Sequence (5' to 3')**	**Digest**
At3g61130 (*GAUT1*)	GAUT1 F	GTC GAC CTC TCA CTC GCT CTC TCT CTT CTT TCT ACG	SalI, BamHI
	GAUT1 R	CGA ACT AAA TCA AAA AAA AAA CTT AAA CTA TCC CCC TAG G	
At2g38650 (*GAUT7*)	GAUT7 F	GTC GAC ACT CAA AAC TAA AAG AAC AGT CAC	SalI, XbaI
	GAUT7 R	GTA AGT TAA GAT TTA GCC CTT AAG ATC T	

Genomic PCRs for Insertion Lines			
Gene ID	**Insertion lines**	**Primer Name**	**Primer Sequence (5' to 3')**
Insertion primer for SALK-lines (left border)		Left Border c F (LBcF)	GGT GAT GGT TCA CGT AGT GGG CCA TC
Insertion primer for GABI-KAT line (left border)		GKo8409 F	ATA TTG ACC ATC ATA CTC ATT GC
At3g61130 (*GAUT1*)	GABI-KAT 470H07 (*gaut1*)	GAUT1 F	TAC GCC CTC TTT TCC GAC AAT G
		GAUT1 R	TGT TTT TAA CCC GAA AGA AGA AGA
At2g38650 (*GAUT7*)	SALK_015189 (*gaut7-1*)	015189 F	ATA TCA AGG TCC CAA AGG GGA GAT AAG T
		015189 R	CTC AAG AGA AGC TTT GAT GTG TAG AAT CC
	SALK_046348 (*gaut7-2*)	046348 F	TTC GGA TAC ATC TCT CTG CAA AAC C
		046348 R	CTT GCA CCA GAT TGA ACC TAA ATG G

of positively transformed T1 plants. For a detailed description of the methodology, please see the Materials and Methods section of Chapter 3 in this dissertation.

Overall, the promoter:*GUS* expression patterns of *GAUT1* and *GAUT7* occur throughout the growth and development of the plant, and extensively overlap with each other (Figure 3.1 in Chapter 3 of this dissertation; also see below). In general, P_{GAUT1}*:GUS* seemed to give stronger expression than P_{GAUT7}*:GUS*, as the samples from plants transformed with the former needed shorter incubation times than those of the latter (0.5-3 hrs compared to 24-72 hrs) in the staining solution to develop the blue staining. However, due to the qualitative nature of the promoter:*GUS* analysis, this observation should not be interpreted conclusively to mean that the *GAUT1* promoter is stronger than *GAUT7* promoter. It is possible that *GAUT7* full expression *in vivo* requires other elements in addition to those included in the construct used in this experiment, such as the 3'UTR or even the exons and introns.

In seedlings, *GAUT1* and *GAUT7* expression occurs in cotyledons and hypocotyls, especially in the meristematic regions and vascular tissues (Figure 3.1, I and J, in Chapter 3). GUS staining was seen throughout the cotyledons and young leaves, with intense staining in the leaf primordia, at the tips of cotyledon and in developing young leaves, and in leaf vasculature. Strong *GAUT1* expression is also evident throughout the vasculature of the seedling roots except in the root elongation zone, and is especially strong at the root tips and in budding lateral roots. *GAUT7* has a similar, but at a much lower intensity, root expression pattern.

The co-expression of *GAUT1* and *GAUT7* is also apparent in the stem, roots, and inflorescence of mature plants (Figure 3.1, K through M, in Chapter 3; Figure F-2). In both P_{GAUT1}*:GUS* and P_{GAUT7}*:GUS* transformed plants, GUS staining was observed in the cortex, phloem, cambium, metaxylem, and protoxylem in cross-sections of stems, and in the phloem and

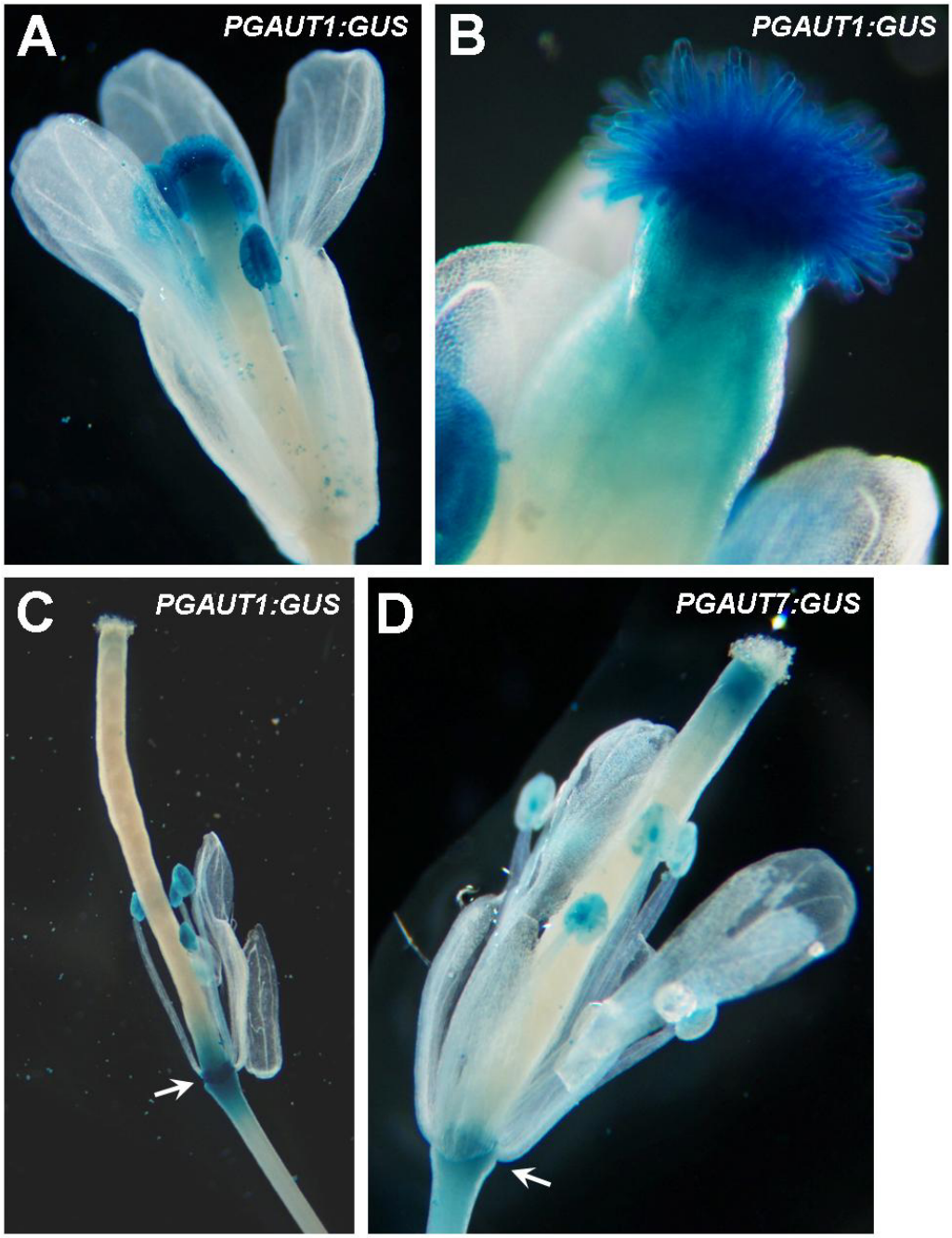


Figure F-2. Promoter:*GUS* expression of *GAUT1* and *GAUT7* in Arabidopsis inflorescence. (A) GAUT1 expression is evident in anthers, pollen grains, and stigma. (B) A close-up of the stigma shows GAUT1 expression in papillae. (C) and (D) GAUT1 and GAUT7 are expressed in the abscission zones (indicated by white arrows).

vascular cambium in cross-sections of tap roots. These results were quite surprising, since they suggest the involvement of GAUT1 and GAUT7 not only in primary wall synthesis as previously thought, but also in secondary wall synthesis. Indeed, the occurrence of pectin-like material was reported in the fibrillar wall deposits in the xylem fibers of poplar (Arend et al., 2008). In the flowers, strong expressions of both *GAUT1* and *GAUT7* are evident in anthers, pollen grains, stigmas, and abscission zones of the flower/developing siliques, while weak expression is seen in the sepals (Figure F-2). The presence of *GAUT1* and *GAUT7* expression in both the male and female reproductive tissues suggests roles of the encoded proteins in pollination and fertilization, and indicates their importance to ensure plant fertility and reproduction. Intriguingly, strong expression of *GAUT1*, but not of *GAUT7*, is also present in the papillae of the stigmas (Figure F-2, A and B). This exclusive expression suggests a tissue-specific function of GAUT1 apart from GAUT7, and may help explain the as yet unsuccessful effort to isolate a homozygous mutant of *GAUT1* (see below).

The widespread occurrence of *GAUT1* and *GAUT7* co-expression is in agreement with previous RT-PCR (Caffall, 2008; Caffall et al., 2009) and publicly available microarray data (Yamada et al., 2003; Zimmermann et al., 2004), and in support of the biochemical data on the protein-protein interaction between GAUT1 and GAUT7 in a protein complex (see Chapter 3 in this dissertation). The fact that there are tissues where *GAUT1* is expressed in the absence of *GAUT7* (e.g. in the papillae), however, suggests that while GAUT1 is dependent on complex formation for its Golgi retention (see Chapter 3 in dissertation), it may not always be partnering with GAUT7, and thus opens up the possibility either that other protein partners exist for GAUT1 or that GAUT1 may not be proteolytically processed in all cell types.

Arabidopsis GAUT7 mutant

Two Arabidopsis mutant lines with T-DNA insertions in *GAUT7* have previously been obtained and screened for homozygosity by Dr. Kerry Caffall (Caffall, 2008; Caffall et al., 2009). Mutant lines SALK_015189 and SALK_046348 carry the insertions in exon 9 and in the promoter region of *GAUT7*, respectively, and were named *gaut7-1* and *gaut7-2* mutants, respectively (Figure F-2B). These lines were phenotypically characterized for seed germination rate, growth phenotype, and glycosyl residue composition. However, they did not show any significant difference from the wild-type (Caffall, 2008; Caffall et al., 2009).

In an effort to find phenotypic features of the *gaut7-1* and *gaut7-2* mutants that differentiate them from wild-type, the growth rate of mutant seedlings *in vitro* was measured. Mutant seeds were sterilized by immersion in 95% EtOH for one min. and 50% bleach/0.05% Tween 20 solution for 10 min, followed by four times rinsing with sterile water. Sterilized seeds were plated along-side wild-type seeds on half-strength Murashige and Skoog media (Sigma, St. Louis, MO) solidified with 0.7% (w/v) of either Bacto Agar (BD, Franklin Lakes, NJ) or phytagel (Sigma, St. Louis, MO). Plated seeds were incubated for 2-3 days at 4°C in the dark, before being transferred to a growth cabinet maintained at 24°C either with a 16-hr light/8-hr dark period or with no lighting (i.e. in the dark). At 10 days of age, *gaut7-1* mutants displayed shorter roots in light-grown seedlings and shorter hypocotyls in dark-grown seedlings in comparison to wild-type (Figure F-3), suggesting that *gaut7-1* seedlings grew more slowly. In contrast, *gaut7-2* seedlings did not differ from the wild-type. No significant difference was observed in mature plants grown on soil, however, between wild-type and either *gaut7* mutants, similar to previous data reported by Dr. Kerry Caffall (Caffall, 2008). Although the results on seedlings are preliminary and need to be repeated, the observation suggests that the seedling

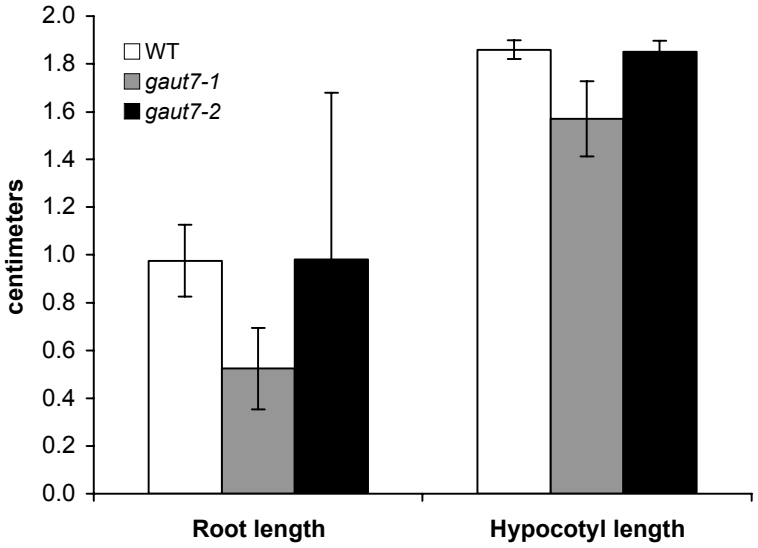


Figure F-3. The growth of Arabidopsis wild-type (WT), *gaut7-1,* and *gaut7-2* seedlings. Root length was measured on light-grown seedlings, and hypocotyl length was measured on etiolated, dark-grown seedlings, each at 10 days old. The data shown are averages from 30 seedlings ± standard error.

phenotype may be a characteristic of the *gaut7-1* mutant. Further research will be carried out on seedlings, instead of mature plants, to study the biological function(s) of GAUT7.

Based on the above results from seedlings, it is possible that the effect of mutation in *gaut7-1* is more severe than in *gaut7-2*. Both *gaut7-1* and *gaut7-2* were reported to be knocked-down mutants by semi-quantitative RT-PCR analysis (Caffall, 2008). However, the relative reduction in expression level between the two mutant lines is unknown. Our preliminary attempt to reproduce the semi-quantitative RT-PCR data using inflorescence samples gave conflicting results, where *gaut7-1* seemed to be knocked-down but *gaut7-2* appeared knocked-up (data not shown). Reproducible RT-PCR and quantitative real time PCR will be necessary to resolve this discrepancy.

Arabidopsis GAUT1 mutant

An Arabidopsis mutant line GK-470H07 that carries a single T-DNA insertion in exon 10 of the *GAUT1* gene was generated by GABI-Kat (http://www.gabi-kat.de) and obtained through the Nottingham Arabidopsis Stock Center (NASC; http://arabidopsis.info/). The gaut1 mutant was received as a set of T3 seeds from eighteen T2 plants. Since the T2 generation is supposed to segregate into wild-type, heterozygotes, and homozygotes, attempts were made to screen for a homozygous line among the T3 plants.

The initial effort was to germinate sterilized seeds (4-7 T3 seeds from each of the T2 plants) on plates containing half-strength MS solidified with 0.7% of phytagel or agar. Only 21 seeds germinated out of the 95 seeds plated, and only 8 got transferred to soil and grew to maturity. Genotyping of the 8 mature plants revealed 3 heterozygous (originated from two T2 plants) and 5 wild-type. There was no observable difference between the heterozygous and wild-

type plants. Harvested seeds (T4 generation) from the heterozygous plants were subsequently germinated on plates, the seedlings transferred to soil, and the plants genotyped. The expectation was that 25% of the plants would be homozygous. However, thus far, out of 48 plants tested only wild-type (18 plants, 37.5%) and heterozygous lines (30 plants, 62.5%) were identified. No homozygous plants were recovered.

An attempt was also made to germinate the seeds on moistened filter paper, instead of on a medium plate, to avoid the necessity of sterilizing the seeds. The fear was that the sterilization treatment may have had a negative effect on the germination of homozygous *gaut1* seeds and rendered them not viable. Using this approach for over four separate sets of experiments, the germination rates of the wild-type and heterozygous *gaut1* seeds were observed to be 96.5% (111/115 seeds) and 85.2% (345/405 seeds), respectively, demonstrating ~12% lower germination of the heterozygous *gaut1* seeds relative to wild-type. Whole 10-day-old seedlings (~1 cm in length) were used directly for genotyping using the REDExtract-N-Amp Plant PCR kit (Sigma, St. Louis, MO). Out of 184 seedlings genotyped, 81 (44.0%) were found to be wild-type, 100 (54.3%) heterozygous, and 3 (1.6%) homozygous. Intriguingly, over the last two batches of the experiments it is noted that a considerable proportion of the heterozygous *gaut1* seeds look smaller in size while the rest appear comparable to wild-type seeds, and that the small seeds have a much lower germination rate (56.7%, 17/30 seeds) compared to the large seeds (93.8%, 30/32 seeds). No homozygous *gaut1* individuals were found in this particular experiment. It is possible that homozygous *gaut1* individual seeds were among the smaller seeds and failed to germinate upon sowing. In previous experiments, this phenomenon may have gone unnoticed and thus larger seeds may have been, unintentionally, selectively picked for sowing, resulting in an overall higher germination rate (82.5%) of the heterozygous *gaut1* seeds.

Unfortunately, since the whole seedlings were sacrificed for genotyping, no homozygous *gaut1* plant has been isolated to date.

While its identification indicates a potential for the isolation of homozygous *gaut1* mutant plants, it is clear that a large number of the plants grown from the heterozygous *gaut1* seeds need to be screened to obtain a homozygous individual. The very low rate of homozygous *gaut1* occurrence (1.6%) and the much lower germination rate of the smaller seeds (56.7%) suggests that the mutation in *GAUT1* likely affects embryo development and causes lethality. Indeed, the ubiquitous expression of *GAUT1,* as shown by the above promoter:*GUS* expression analyses, demonstrates the importance of GAUT1 throughout plant growth and development, with *GAUT1* expression detected relatively strongly in developing seedlings. It would be interesting to see if homozygous *gaut1* are indeed among the small seeds and if there is any phenotypic difference between the small and large seeds that can be observed by microscopy. On the other hand, strong GAUT1 expression was also observed in both pollen and stigma, thus, the possibility remains that homozygous *gaut1* seeds are generated at much lower rate due to defects in fertility. More experiments are currently underway to try to isolate homozygous *gaut1* plants and/or to determine the phenotypic consequences of the mutation.